

DOI: https://doi.org/10.1093/eep/dvad005 Advance Access Publication 15 July 2023 Review Article

Epigenetic clocks and research implications of the lack of data on whom they have been developed: a review of reported and missing sociodemographic characteristics

Sarah Holmes Watkins^{1,2,*}, Christian Testa³, Jarvis T. Chen³, Immaculata De Vivo^{4,5}, Andrew J. Simpkin⁶, Kate Tilling^{1,2}, Ana V. Diez Roux⁷, George Davey Smith^{1,2}, Pamela D. Waterman³, Matthew Suderman^{1,2}, Caroline Relton^{1,2} and Nancy Krieger³

¹Population Health Sciences, Bristol Medical School, University of Bristol, Bristol BS8 2BN, UK, ²Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol BS8 2BN, UK, ³Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA 02115, USA, ⁴Program in Genetic Epidemiology and Statistical Genetics, Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA, ⁵Department of Medicine, Harvard Medical School, Brigham and Women's Hospital, Boston, MA 02115, USA, ⁶School of Medicine, National University of Ireland Galway, Galway H91 TK33, Ireland, ⁷Department of Epidemiology and Biostatistics and Urban Health Collaborative, Dornsife School of Public Health, Drexel University, Philadelphia, PA 19104, USA

*Correspondence address. Population Health Sciences, Bristol Medical School, University of Bristol, Bristol BS8 2BN, UK. E-mail: s.h.watkins@bristol.ac.uk

Abstract

Epigenetic clocks are increasingly being used as a tool to assess the impact of a wide variety of phenotypes and exposures on healthy ageing, with a recent focus on social determinants of health. However, little attention has been paid to the sociodemographic characteristics of participants on whom these clocks have been based. Participant characteristics are important because sociodemographic and socioeconomic factors are known to be associated with both DNA methylation variation and healthy ageing. It is also well known that machine learning algorithms have the potential to exacerbate health inequities through the use of unrepresentative samples – prediction models may underperform in social groups that were poorly represented in the training data used to construct the model. To address this gap in the literature, we conducted a review of the sociodemographic characteristics of the participants whose data were used to construct 13 commonly used epigenetic clocks. We found that although some of the epigenetic clocks were created utilizing data provided by individuals from different ages, sexes/genders, and racialized groups, sociodemographic characteristics are generally poorly reported. Reported information is limited by inadequate conceptualization of the social dimensions and exposure implications of gender and racialized inequality, and socioeconomic data are infrequently reported. It is important for future work to ensure clear reporting of tangible data on the sociodemographic and socioeconomic characteristics of all the participants in the study to ensure that other researchers can make informed judgements about the appropriateness of the model for their study population.

Keywords: DNA methylation; epigenetic age; health equity; socioeconomic

Introduction

DNA methylation (DNAm) is an epigenetic modification to DNA that is involved in a number of aspects of genome regulation. DNAm is dynamic, changing as we age [1], in the course of disease [2] and in the presence of environmental exposures such as smoking [3] and childhood adversity [4, 5]. Ageing is a particularly well-established influence on DNAm, with changes occurring at many DNAm sites across the genome as humans get older [6]. These changes are so reliable that numerous research groups have developed 'epigenetic clocks', mathematical models that use DNAm measurements to predict chronological age and other age-related characteristics [7].

The first generation of epigenetic clock methods (e.g. Horvath and Hannum [1, 8]) aimed simply to predict chronological

age as accurately as possible. It was observed early on that errors in age prediction, subsequently known as 'age acceleration' or 'age deceleration' (depending on the direction of the error), were associated with a variety of exposures and diseases as well as mortality risk [9]. These associations suggested that, beyond chronological ageing, these clocks may actually provide a measure of biological ageing (which refers to the progressive decline of the body's physiological functions [10]). To enhance this feature, a second generation of epigenetic clocks was developed; these incorporated health biomarkers and phenotypes alongside chronological age in the model. Increasing numbers of studies are now considering how social determinants of health might impact epigenetic ageing, with evidence indicating positive associations between accelerated epigenetic aging and social adversity measured in relation to socioeconomic position [11],

Received 20 February 2023; revised 17 June 2023; accepted 13 July 2023

© The Author(s) 2023. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

education [12], neighbourhood characteristics [13–15], and racial discrimination [15].

Although the epigenetic clocks themselves have been quite extensively reviewed [7, 16–18], relatively little attention has been paid - in both these review articles and also the empirical studies that use these clocks - to the sociodemographic characteristics of the individuals whose data have been used to develop these clocks. This sociodemographic information is important in order to assess the generalizability of the clocks to different populations, given robust evidence that (i) variation in DNAm is altered by genetic, social, and environmental factors [3, 19–24], (ii) adverse social and biophysical exposures are socially patterned and affect the risk of disease and also processes of healthy ageing [25-27], and (iii) study selection pressures (which may be related to sociodemographic characteristics) can induce biased study estimates [28-31], which could lead to problems with clock generalizability. Our concern is that for these reasons an algorithm developed to estimate ageing using DNAm in one population may not translate to other populations with divergent sociodemographic or socioeconomic profiles. For example, a clock may miss environmental impacts on the methylome that contribute to ageing in some populations if it is developed in a population with low exposure to those impacts, e.g. pollution or social adversity [32].

This is not a new concern, as there is existing literature describing the potential for machine learning algorithms to exacerbate health inequities through the use of unrepresentative samples this may happen when prediction models underperform in social groups that were poorly represented in the training data used to construct the model [33-35]. If prediction models underperform, then estimates will be less accurate, which could bias associations in either direction. Because of the nature of epigenetic clocks, the direction of bias is likely to be quite complicated. Epigenetic clocks are made up of many CpG sites (often hundreds) scattered across the genome. The dynamic and regulatory nature of DNAm therefore implies that clocks will integrate biological signals from a wide variety of biological functions and pathways (which for the most part are unknown). The responsive nature of DNAm implies that they will capture variation due to a multitude of endogenous biological factors and external exposures that differ between populations. As such this bias may either over- or under-estimate the effects, and we do not think it would be easy to predict which direction the bias might go, even for specific studies.

The possibility of bias has already been described for PhenoAge, one measure of epigenetic age [36]; it is a particular concern for epigenetic data as two recent reviews show that, at least in regard to racialized groups, epigenetic data are predominantly from individuals of 'European' heritage [37, 38] (although no such review has been conducted for other sociodemographic characteristics). Wide-spread inconsistencies in the literature about associations with epigenetic age acceleration support this concern, as inconsistencies are likely driven by differences in characteristics between study populations. For example, inconsistent associations have been reported with education [12, 15, 39-42], socioeconomic status [15, 39, 41, 43–50], and racialized group [51–54]; as well as inconsistent associations between epigenetic age acceleration and health and social outcomes when analyses are stratified by sociodemographic characteristics, such as education level [55], country of birth [40], and racialized group [42, 55, 56]. One recent study reported that when testing the association between epigenetic clocks and healthspan-related characteristics, smaller effect sizes were found for Black American participants in comparison to white American participants [54], suggesting it is possible that associations could be biased towards the null in some study populations. These inconsistencies may be due in part to some clocks

including loci known to be differentially methylated by country of birth [40] and racialized group [36, 45].

Whatever the reason for these inconsistencies, it is clear that any interpretation of associations with epigenetic age acceleration should take into account any differences between the populations in which the epigenetic clock was derived and the populations where the associations were observed. To assist in these comparisons, we collate for the first time the basic social characteristics of the participants utilized in the development of 13 commonly used epigenetic clocks.

Results

Table 1 contains extracted information from each of the clock papers. Supplementary Table S1 reproduces Table 1 along with additional details of the conceptualization of racialized groups and genders. Supplementary Table S2 provides details of how each clock was constructed and the specific datasets used by each paper.

Age, Biological Specimen Collection Date, and Birth Cohort

Eleven of the 13 papers report the chronological age of all their training and test dataset participants. Of these, three models included children in their training and test datasets: two included participants from birth to adulthood, with no information as to the biological specimen collection dates or birth cohorts of participants (aged 0-100 years [1]; aged 0-78 years [57]), and one model included participants aged 2–104 years old, with 1/14 cohorts born in 1936 and one in 1921 [58]. Four papers reporting the chronological age of all their participants report using only adult participants (aged between 16 and 101 years) without reporting specimen collection date or birth cohort [8, 59-61]. One paper reported using only adult participants (ages 31-82 years) and reported specimen collection date (2000-08) and birth cohort (1924-75) for all of their participants [62]. One paper reported age for all participants (adults, aged 20-91 years) and reported the specimen collection date (1988-2014) and birth cohort (1907-94) for the majority of their participants (4/5 datasets) [63]. Two papers reported using adult participants (aged 21-93 years), with information on specimen collection date (1998-2007) and birth cohort (1906-86) for some of their participants [55, 64].

One paper reported age for their training data plus 81% of their test data, plus biological specimen collection date for all participants (1998–2013) and birth cohort for the majority of participants (1917–95) [65]. One paper did not report age but included it as the Y-axis on plots, illustrating that participants were ~20–90 years old, with no information on biological specimen collection date or birth cohort [66].

Sex or Gender

Biological sex or gender were the most frequently reported characteristics, with seven out of 13 clock papers reporting data for all participants, and three out of 13 reporting partial data. Of the seven reporting full data, four report 'sex' using the biological sex terms 'males' and 'females'. These papers included between 48% and 83% females in their training and test datasets [55, 62, 64, 65]. One paper reports 'gender' using the gendered terms 'women' and 'men', including 52% women [8], and two report using the terms 'female' and 'male' without specifying sex or gender, including between 0% and 58% female participants [1, 60]. Of the three studies reporting partial data on sex/gender, one reported on one of eight datasets (100% women, with no use of the terms sex or gender) [61]; one reported on two of five test datasets (46% women, no

	Sociode- mographic characteristics used in model Report validation across constructing sociodemographic clock? groups?	No No No	Gender and No. Compare the age- ethnicity (no ing rate between men definition of and women in train- either; ethnic ing data; but no test groups included dataset is used so not are 'Caucasian considered validation and Hispanic')
D	s Education	paper	Not stated paper
	Economic measures	Not stated in paper	Not stated in paper
	Nativity	Not stated in paper	Not stated in paper
	Sex or gender	No specification of termiology; terms 'male' and 'female' only Training: 37% female. Test: 8% female	Terms 'gender', 'women', and 'men' with no explicit defini- tion. Training: 52% women
	Country	Training: reported for 3/39 datasets - two datasets from one study include data from the Netherlands and one dataset includes data from the USA. Test: reported for 2/32 datasets - two datasets drawn from one study include data from the UK	Not stated in paper
	Racialized groups	No specification of terminology; the term 'ethnic- ity' is used with regard to one of the studies. Train- ing: reported for 5/39 datasets; one study com- prising four datasets includes individuals of 'non-Hispanic Caucasian ethnic- ity' (4 brain regions from the same individuals, and one study includes 'Taiwanese' indi- viduals. Test: reported for 1/32 datasets; one study includes 'Cambian'	Term 'ethnicity' with no explicit definition. Train- ing: '426 Caucasian and 230 Hispanic individuals'
1 0	Age (years) and year at time of specimen collection; birth cohort	Age: 0–100; year: not stated in paper; birth cohort: cannot determine	Age: 19–101; year: not stated in paper; birth cohort: cannot determine
	Clock	Horvath [1]	Hannum [8]

Report validation across sociodemographic groups?	No	°Z.	oN	(continued)
Sociode- mographic characteristics used in model constructing clock?	Sex (no defini- tion)	° Z	°Z	
Education	Not stated in paper	Not stated in paper	Not stated in paper	
Economic measures	Not stated in paper	Not stated in paper	Not stated in paper	
Nativity	Not stated in paper	Not stated in paper	Not stated in paper	
Sex or gender	Not stated in paper	Not stated in paper	Not stated in paper	
Country	Not stated in paper	Training: Reported for 4/14 cohorts - three cohorts from Scotland and one from Australia	Not stated in paper	
Racialized groups	Not stated in paper	No specification of terminology. Train- ing: reported for one of 14 cohorts; 'the MND cohort is from a systems genomics study of Motor Neuron Disease in Chinese subjects'	Not stated in paper	
Age (years) and year at time of specimen collection; birth cohort	Age: 19–101; year: not stated in paper; birth cohort: cannot determine	Age: 2–104; year: not stated in paper; birth cohort: reported for two cohorts – one born in 1921 and the other born in 1936	Age: not stated in paper, included only as Y-axis on some plots (~20-90 years old); year: not stated in paper; birth cohort: cannot determine	
Clock	epiTOC [59]	Zhang's elastic net [58]	MiAge [66]	

4 | Environmental Epigenetics, 2023, Vol. 00, No. 00

Table 1: (Continued)

Clock	Age (years) and year at time of specimen collection; birth cohort	Racialized groups	Country	Sex or gender	Nativity	Economic measures	Education	Sociode- mographic characteristics used in model constructing clock?	Report validation across sociodemographic groups?
DNAmTL [64]	Age – training: 22–93, Test 1: 22–82, and Test 2: 21–100; year – training: stated for 1/2 cohorts: 2000–04, Test 1: stated for 1/3 cohorts: 2000–04, and Test 2: stated for 4/9 cohorts: 1998–2007; birth cohort – train– ing: 1/2 cohorts: 1906–81, Test 1: 1/3 cohorts: 1915–86 cohorts: 1915–86	Terms 'ethnicity', 'race/ethnicity', and 'ancestry' used with no def- inition. Training: 19% 'Euro- pean ancestry' (text)/'European' (tables) and 81% 'African Ances- try' (text)/'African American' (tables), Test 1: 14% 'African Ances- try' (text)/'African American' (tables), 86% 'Euro- pean ancestry' (text)/'European' (text)/'European' (text)/'European' (text)/'European' (text)/'European' (text)/'European' (text)/'European' (text)/'European' (text)/'European' (text)/'European' (text)/'European' (text)/'European' (text)/'African American' (text)/'African American' (text)/'African American' (text)/'African American' (text)/'African American' (text)/'African	Training: reported for 1/2 - one in the USA. Test 1: reported for 2/3 - two in the USA. Test 2: one in the UK, two in Scotland, one in ftaly in Italy	Term 'sex' used with no defi- nition; terms males/females and men/women used. Training: 75% female. Test 1: 56% female. Test 2: 74% female	Not stated in paper	Not stated in paper	Reported for 5417 of 9875 (54.9%) of test dataset participants: 16.6% less than high school, 16.7% high school degree, 33.6% college, 33.2% college degree and beyond	°N N	Yes. Compare the DNAm/TL correlation with chronological age between men and women – in Framing- ham Heart Study (FHS) (test dataset), higher correlation in men (0.39 vs. 0.26) even though training data were 75% women. In Bogalusa Heart Study (BHS) (test dataset), higher correlation in women (0.48 vs. 0.34). Also compare between the racialized groups in one test dataset – in BHS no difference between 'black' and 'white' participants (0.36 vs. 0.35)

(continued)

Report validation across sociodemographic groups?	Yes. Compare the cor- relation between predicted age and chronological age in males and females in test dataset. The prest data	Yes: State that 'Gender- related differences in the age-predictions were not observed using this signature (data not shown).'	Yes. Compare the difference between predicted and chronological age for males and females in test dataset in plot in figure 3, showing no difference at $P = 0.05$, but do not provide statistics or elabo- rate on comparison. In the Supplementary material, compare the difference between predicted and chrono- logical age for differing numbers of years in education, showing no difference at $P = 0.05$
Sociode- mographic characteristics used in model constructing clock?	° Z	° Z	Ŝ
Education	Not stated in paper	Not stated in paper	Not stated in paper
Economic measures	Not stated in paper	Not stated in paper	Not stated in paper
Nativity	Not stated in paper	Not stated in paper	Not stated in paper
Sex or gender	Terms 'male' and 'female' used with no definition. Training: 100% male. Test: 48% female	Term 'women' only used in reference to one dataset. Training: not reported. Test: reported for 1/8 datasets = 100% 'women'	Termi 'gender' used, with 'males' and 'females'. Train- ing 1: 81% female. Test 1: unclear. Test 2: unclear. Test 2: unclear.
Country	Not stated in paper	Not stated in paper	Training 1 and Test 1: not reported. Train- ing 2: Germany. Test 2: Germany
Racialized groups	Not stated in paper	Not stated in paper	Not stated in paper
Age (years) and year at time of specimen collection; birth cohort	Age: 21–55; year: not stated in paper; birth cohort: cannot determine	Age: 16–72; year: not stated in paper; birth cohort: cannot determine	Age: 0–78; year: not stated in paper; birth cohort: cannot determine
Clock	Bocklandt [60]	Koch/Wag- ner [61]	Weidher [57]

(continued)

Table 1: (Continued)

Report validation across sociodemographic groups?	Yes. Downstream asso- ciation analyses in both training and test stratified by sex/gen- der. Most associations stronger in women in training data; most associations stronger in men in test data (similar proportion of men and women in training and test datasets)	Yes, across racilaized groups in the test dataset (Black, white, and Hispanic groups). Test via the correla- tion between DNAm age and chronologi- cal age showed similar correlations for Black and white partici- pants, with a slightly higher correlation for Hispanic participants
Sociode- mographic characteristics used in model constructing clock?	Sex (not defined)	° Z
Education	Not stated in paper	Not stated in paper
Economic measures	Not stated in paper	Not stated in paper
Nativity	Not stated in paper	Not stated in paper
Sex or gender	Term 'sex' used with terms males/females and women/men (males/females in table where % is taken from). Training: 52% female. Test: 50% female	No mention for DNAm data other than cohort descrip- tion and no definition. Biomarker selection: not stated in paper. Training: no data presented. Test: data pre- sented for 2/5 datasets (27% of participants) – 54% 'men' and 46% 'women'
Country	Training: Ger- many: Test: Germany	Biomarker selec- tion: the USA. Training: no data presented. Test: data pre- sented for 4/5 datasets, all based in the USA
Racialized groups	Not stated in paper	Terms used: 'race/ethnicity' and 'ancestry', no definition Biomarker selec- tion: not stated in paper. Training: no data presented. Test: data pre- sented for 4/5 datasets (72% datasets (72% 'African Arnerican', 11% 'Hispanic' 41% 'White'
Age (years) and year at time of specimen collection; birth cohort	Age - training: 50-75, test: 31-82; year - training: 2000-02, test: 2006-08; birth cohort - training: born 1925-50, test: born 1924-75	Age – biomarker selection: 20+, training: 21–91 (baseline) and 30–100 (follow- up), test: 50–80; year – biomarker selection: 1988–94 and 1999–2014, training: 1998 (paseline) and 2007 (follow-up), test: 193–2008; birth cohort – biomarker selection: before 1974–94, train- ing: 1907–77, test: 1913–43
Clock	Zhang's mortality [62]	PhenoAge [63]

(continued)

Table 1: (Cc	ntinued)								
Clock	Age (years) and year at time of specimen collection; birth cohort	Racialized groups	Country	Sex or gender	Nativity	Economic measures	Education	Sociode- mographic characteristics used in model constructing clock?	Report validation across sociodemographic groups?
Duned- inPoAm [65]	Age – biomarker selection: 26, 32, and 38, training: 38, test: reported for 3/4 datasets (81% participants): 18–95; year – biomarker selec- tion: 1999, 2004, and 2010, train- ing: 2010, test: 1999–2013; birth cohort – biomarker selection and training: April 1972–March 1973, test: (3/4 datasets) 1917–95	Terms used: only 'white' and 'of white Euro- pean descent'. Biomarker selection and training: 93% were 'white'/'of white European descent'; not specified for the remaining 7% of participants. Test: 'mostly of white European descent'; data reported for 1/4 datasets: 77% 'white'	Biomarker selection and training: New Zealand. Test: reported for 3/4 datasets – two from the UK and one from the USA	Terms 'sex', 'male'/'men', and 'female/'women' used, no defini- tion. Biomarker selection and training: 48% female. Test: reported for 3/4 datasets - 61.5% male	Biomarker selection and train- ing: 100% born in Dunedin, New Zealand. Test: not stated	Biomarker selection and training: 'The cohort represents the full range of socioeconomic status on NZ's South Island'. Test: reported for 1/4 datasets (41% of test participants) – 25.6% of E-Risk families lived in 'wealthy achiever' neighbourhoods compared to 25.3% nationwide; 5.3% vs. 11.6% lived in 'urban prosperity' neigh- bourhoods; 29.6% vs. 26.9% lived in 'comfortably off neighbourhoods; 13.9% lived in 'moderate means' neighbour- hoods; and 26.1% vs. 20.7% lived in 'hard-pressed' neighbourhoods	paper	°Z	Yes. Includes slopes for 'men' and 'women' for one of the test datasets in figure 3 of the DunedinPoAm paper, but no statistics are reported. Com- pare clock estimates between participants from low, middle, and high social class cate- gories, with those from the lower socioeco- nomic groups having faster clock ageing
									(continued)

Table 1: (C	Continued)								
Clock	Age (years) and year at time of specimen collection; birth cohort	Racialized groups	Country	Sex or gender	Nativity	Economic measures	Education	Sociode- mographic characteristics used in model constructing clock?	Report validation across sociodemographic groups?
GrimAge [55]	Age – biomarker selection and training: 59–73, test: 46.5–78; year – biomarker selec- tion and training: not reported, test: reported for 2/5 test datasets: 1998–2007; birth cohort – biomarker selection and training: not reported, test: for 2/5 datasets: 1920–57	Terms used: 'race/ethnicity' and 'ancestry'. Biomarker selec- tion and training: not stated in paper. Test: 50% 'European ances- try (Caucasians)', 40% 'African Americans', 10% 'Hispanic ancestry'	Biomarker selection and USA. Test: the USA (2/5) and Italy (1/5)	Terms 'sex' and 'female' used, no definition. Biomarker selection and training: 54% female. Test: 83% female	Not stated in paper	Not stated in paper	Stated for 3079 (41.7%) of 7375 test dataset par- ticipants: 8.8% less than high school, 14.8% high school degree, 52% some college degree and beyond	Sex (not defined)	Yes. Stratified test datasets by racialized group membership to test whether the mortality predic- tors applied to each group. However, do not make direct com- parisons in the text. Stratify by educational attainment in a Sup- plementary analysis and find that AgeAc- celGrim remains a significant predictor of time-to-death for all education levels (although the low- est HR for less than high school education); same for time to Coro- nary Heart Disease. AgeAccelGrim is strati- fied by racialized group in a Supplementary analysis with differ- ences between groups, with higher SDs in the black participants, although this is not diantoned this is not diantoned

use of the terms sex or gender) [63]; and one paper reports 'gender' using the terms 'female' and 'male', including 81% female participants in the training data, with no reporting for the test datasets [57]. Three studies use 'sex' and one used 'gender' as a covariate in the age estimation model, but none conceptualize the reason for this [8, 55, 59, 62].

None of the papers defined what they meant by the terms sex or gender beyond three studies identifying it as a 'covariate' or 'confounder', so where terminology is mixed it is unclear which specific characteristic is being reported. Being specific about the characteristic that is reported is important because biological sex and gender identity have independent and interacting effects on health [67–70].

Racialized Groups and Country Context

Only two papers report data on racialized group membership for all the participants in their study, using descriptive terms pertaining to 'race', 'ethnicity', and 'ancestry' (but without defining what these terms mean, whether these data were self-reported by participants or obtained from medical records, and with limited reporting of country context). The paper developing the Hannum clock [8] was developed using a training dataset including '426 Caucasian and 230 Hispanic individuals', with no country context reported. We note that there is scientific recognition that the origins of the term 'Caucasian' mean its use should be discontinued this is because the term was derived from a scientifically fallacious typology that presumed humanity originated in the Caucuses and was 'white', with other 'racial' groups framed as 'degenerate' lineages that branched off of the 'Caucasian' trunk [71–73]. The paper developing the DNAmTL clock [64] utilized training and test datasets; the training dataset comprised a total of 2256 individuals from two datasets, of whom 81% were categorized as being 'African ancestry' in the text and 'African American' in the tables; and 19% as 'European ancestry' in the text and 'European' in the tables. The country context was provided for one of these datasets (one was based in the US), so it is not possible to fully contextualize the reported racialized groups based on information reported in the paper. Their first test data comprised 1078 individuals from three datasets (Test and Training 1 comprise unique sets of individuals from two datasets). They report that 86% were categorized as being 'European ancestry' in the text and 'European' in the tables, and 14% as 'African ancestry' in the text and 'African American' in the tables. The country context was provided for 2/3 of these datasets (two were based in the USA), so again it is not possible to fully contextualize the groups. A second test dataset comprised 9359 individuals from five additional datasets, of whom 77% were categorized as 'European ancestry' in the text and 'European' in the tables, 15% as 'African ancestry' in the text and 'African American' in the tables, and 8% as 'Hispanic ancestry' in the text and 'Hispanic' in the tables. The country context was reported for all five datasets: one in the UK, two in Scotland, one in the USA, and one in Italy.

Five papers that developed an epigenetic clock reported partial data on racialized group membership but did not report clear numbers for all participants in the paper. The paper developing the Horvath clock [1] reported racialized group membership for 5/39 training datasets – participants from four datasets (four brain regions from the same individuals) were categorized as 'non-Hispanic Caucasian ethnicity' (country context not reported), and participants from one dataset were categorized as 'Taiwanese' (country context not reported). Racialized group membership data was reported for 1/32 test datasets, where participants were categorized as 'Gambian' (country context not reported). The paper developing the DunedinPoAm clock [65] reported that 93% of their training dataset participants, all born and residing in New Zealand, were 'white' but did not explicitly report data on racialized groups for the remainder of their training data participants. They state that test dataset participants were 'mostly of white European descent', reporting that 77% of participants in one of the four test datasets were 'white' with no country context; the country context was reported for the other three test datasets with no information on racialized groups (one recruited in the USA and two in the UK). The paper developing the GrimAge clock [55] did not report the racialized group memberships of their US-derived biomarker selection and training dataset (70% of the Framingham Heart Study offspring cohort). Among individuals in their five test datasets, 50% were categorized as 'European ancestry (Caucasians)', 40% as 'African American', and 10% as 'Hispanic'; the country context was not integrated, with 2/5 of the test datasets based in the USA and 1/5 in Italy. The paper developing the PhenoAge clock [63] did not report the racialized group membership of the participants in which the biomarkers were selected the National Health and Nutrition Examination Survey III and IV; both are US national datasets), or in whom the DNAm predictor was trained (InChianti, for which they do not report the country context); they reported the racialized group membership of participants in four of the five test datasets, all of which were based in the USA, comprising 72% of test participants, of whom 18% were categorized as being 'Black', 26% as 'African American', 41% as 'White', and 11% as 'Hispanic'. No data on racialized group membership was provided for the remaining 36% of test participants from two other datasets. Finally, Zhang's et al. paper developing an elastic net predictor [58] reported data on racialized group membership for one of 14 datasets, where they state that the 695 participants from the motor neuron disease (MND) cohort dataset were 'Chinese' but did not provide further information or country context.

One study reports only the country context of training and test datasets (both based in Germany) [62]. Another reported only country context for 1/2 training datasets and 1/2 test datasets (both based in Germany) [57]. The remaining four studies [59–61, 66] report no information about the racialized group membership or country context of their participants. None of the papers that did report data categorized by 'racial' or 'ethnic' terms that they employed explicitly explained or justified their conceptualization or usage of racialized groups, beyond three studies conceptualizing use as a 'covariate' or 'confounder'. One study [8] includes 'ethnicity' as a covariate in the age estimation model but does not conceptualize the reasons for this.

Nativity

Only one study explicitly reported the nativity status of their training dataset participants – 100% of the Dunedin PoAm [65] training dataset study participants were born in New Zealand (with no report as to the nativity status of the test dataset).

Economic Measures

Only the Dunedin PoAm study [65] included data for some participants on economic status in the paper. For the biomarker selection and training dataset, they state that 'The cohort represents the full range of socioeconomic status on NZ's South Island', but they do not provide descriptive statistics that might enable comparison across datasets. For one of four of their test datasets (41% of test participants), they provide neighbourhood-level socioeconomic data alongside national statistics, illustrating that the sample is broadly representative of the UK in terms of neighbourhood SES.

Education

Two papers developing epigenetic clocks included data on educational attainment for some of their participants. The first paper, developing the GrimAge clock [55], reported education data for 41.7% of their test dataset participants; they were generally highly educated, with a quarter (24.2%) having a college degree or higher, and only 8.8% having less than high school education (14.8% had a high school degree, and 52% had some college education). The second paper, developing the DNAmTL clock [64], reported education data for 54.9% of their test dataset participants. A higher proportion had less than high school education (16.6%), with the remainder of the sample generally highly educated: 16.7% had a high school degree, 33.6% had some college education, and 33.2% had a college degree or higher. Neither paper reported information on the education level of their training dataset, which is critical to the transportability of the predictor.

Validation across Social Groups

Eight of the 13 papers reporting epigenetic clock methods report some form of validation of their model in their test data stratified by one or more sociodemographic characteristics. One (developing the DNAmTL clock) compared model performance between men and women in the two test datasets, and additionally between racialized groups in one of these test datasets, by correlating the clock estimate with chronological age [64]. In the first test dataset, they found a lower correlation for women even though the training data comprised 75% women. In the second test dataset, they found a higher correlation for women. We suggest this shows the importance of the need to consider multiple participant characteristics when validating predictive algorithms. There was no difference in clock correlation with chronological age between the two racialized groups ('black' and 'white' participants). The paper developing Zhang et al.'s mortality clock [62] stratified downstream analyses by sex/gender; in their test dataset, they found stronger associations in men despite a slightly higher proportion of women in the training data. The paper reporting the Bocklandt clock [60] used 100% 'males' in their training data and tested clock performance in males and females in their test dataset. They found a lower correlation between chronological and predicted age for females compared to males, as well as a higher error between predicted and chronological age for females, suggesting reduced accuracy of the clock. The paper developing the PhenoAge clock [63] looked at the correlation between the age estimator and chronological age stratified by the racialized group they found slightly higher correlations in 'Hispanic' participants, with similar correlations for 'black' and 'white' participants. The paper developing the GrimAge clock [55] stratified all analyses in the main text by the racialized group, conceptualizing this as testing whether the ageing predictor applied to each group; no conclusions from this were presented in the paper, but in a supplementary analysis shows higher standard deviations of the age estimator for 'Black' participants and lower standard deviations for 'Hispanic' participants, with 'white' participants having standard deviations in the middle of these groups. However, racialized group membership was not reported for the training dataset participants. Additionally, in the supplement they stratify model predictions by educational attainment and find that the model performed for participants of all education levels (although the lowest hazard ratio was found for less than high school education). One paper compared the difference between predicted and chronological age for participants with differing numbers of years in education, showing no difference at P = 0.05 [57] and presented plots comparing clock performance in 'males' and 'females' but did not provide statistics or elaborate on the comparison in the text. One paper developing an epigenetic clock stated that no difference was observed in clock performance between genders but did not present data [61]. One paper presented plots comparing clock performance in 'males/men' and 'females/women' but did not provide statistics or elaborate on the comparison in the text [65]. One paper compared the ageing rate between men and women but only in the test dataset (which we did not consider to be validation) [8].

Discussion

The basic sociodemographic characteristics of participants are generally poorly reported in the 13 papers which developed the most popular epigenetic clocks. This makes it challenging for researchers to judge whether the clock is likely to accurately transport to the population they want to study, where the estimation of epigenetic age may be inaccurate in populations with different characteristics, introducing uncertainty to the relationship between epigenetic age and health and social outcomes, therefore biasing estimates in uncertain directions. This is important because different populations are likely to have different socially patterned social, economic, and biophysical exposures that affect their methylomes, and so clocks developed in a socially homogeneous population may not transport well to a population with different social characteristics and different exposures and experiences.

Chronological age was reported for all study participants in 11 of 13 papers. Of these, three include infants and children in model development, meaning that researchers should be cautious about the application of other models to data from children. The majority of other studies included a wide range of adult ages. However, only six of the papers report the biological specimen collection dates, some with information reported on birth cohort (although birth cohort can be derived easily from age and specimen collection date). Biological specimen collection date and birth cohort are important because this means researchers can ascertain whether participants lived through events or situations that might have had impacts on health and health equity. To properly analyse issues related to health equity, it is crucial to combine this with data on place and other sociodemographic characteristics. Participants in the six epigenetic clock models where birth cohort could be derived were born from as long ago as 1906 to as recent as 1995, meaning a range of historical and social events (such as war, economic crisis, and changes in social environment) may have been experienced by participants, depending on their other characteristics.

Where gender or sex was reported, with one exception clocks were trained on both male and female/women and men participants; however, gender or sex was not reported for all participants in five of the 13 studies, and none included any discussion of differences between the influences of sex-related biology and societal gender, or their potential interaction. The two clocks reporting data on the racialized group membership of all participants in their training dataset did not provide country context for all of their participants. Each included individuals from two racialized groups: one included participants predominantly categorized as being 'Caucasian' (with no country context) and the other included participants predominantly categorized as being of 'African ancestry' or 'African American' (with at least some of these participants living in the USA). Only two studies (the papers developing DNAmTL and GrimAge) reported the racialized group membership of all participants in their test dataset (this

information is not as pertinent as the training dataset); neither fully integrated country context, with one study including persons located in the USA, the UK, Scotland, and Italy, as well as some participants from unspecified countries, who were categorized as being 'European' or 'European ancestry', 'African American' or 'African ancestry', and 'Hispanic' or 'Hispanic ancestry'; the other study included persons living in the USA and Italy, as well as unspecified countries, who were categorized as being 'European ancestry (Caucasians)', 'African American', and 'Hispanic'. Both of these studies included a majority of participants categorized as being 'European' or 'European ancestry', some of whom were located in Europe, some in the USA, and for some participants, there was no country context. Only one paper reported the nativity of their participants alongside the country context.

We note that none of the papers that we reviewed explicitly explained or justified their conceptualization or usage of racialized groups, even when they reported data on these characteristics, utilizing them as covariates in the model in one case, or stratified their analysis according to categories they employed for 'race', 'ethnicity', or 'ancestry'. This mirrors the findings of a recent systematic review examining how a large number of epigenetic studies poorly incorporate data on social groups and social determinants of health [74]. The inclusion of sociodemographic characteristics as features in models such as epigenetic clocks needs to be thoroughly conceptualized and justified because the inclusion of these characteristics could exacerbate inequities by adjusting away inequalities experienced by individuals with these characteristics [75, 76], meaning inequalities in biological age would be masked by the inclusion in the clock algorithm.

Crucially, only three of the papers we reviewed presented tangible data pertaining to the socioeconomic circumstances of their participants. Two reported education levels for 41.7% and 54.9% of their test dataset participants, where participants were generally highly educated (with the majority having at least some college education). One reported neighbourhood-level economic data for 41% of test dataset participants alongside national figures, illustrating that the sample was broadly representative of the UK. However, none of these papers reported education or economic data for their training data, which are the critical dataset to report. This information is essential to assess the transportability of these clocks to other datasets; it is also important to ensure that health inequalities are not masked or perpetuated in epigenetic research (this may happen when prediction models underperform in social groups that are poorly represented in their training data). The lack of reporting that we find is likely to be due to at least in part the absence of social characteristics in publicly available datasets such as those on GEO; biological data repositories have previously been criticized for a lack of social characteristics of their participants because this prevents the investigation of health inequities that exist between social groups [77]. We would like to reiterate this need for socioeconomic data in the context of epigenetic datasets, as well as the importance of obtaining and reporting these data from cohort studies that have epigenetic data.

Eight of the epigenetic clock models make efforts to validate their models in participants stratified by one or more sociodemographic characteristics, including sex/gender, racialized groups, and education level. Some suggested that there may be little difference between the groups tested, whereas some suggested lower accuracy in groups dissimilar in some way to the training population. However, we suggest that validation methods ought to be improved beyond simple testing of correlation between the clock model and chronological age, or testing downstream associations and that papers should consider multiple sociodemographic characteristics in these validation analyses and ensure to give them due consideration as an important part of the manuscript. None of the eight papers followed up any differences they found in the discussion, or relate differences they find to the characteristics of their training dataset, missing important opportunities to delve into whom these clocks may and may not apply to.

In conclusion, we find that although some of the epigenetic clocks were created utilizing data from datasets including individuals from different sexes/genders and racialized groups, this information is limited by inadequate conceptualization of the social dimensions and exposure implications of gender and racialized inequality, the absence of any socioeconomic data, or any consideration of interactive effects involving these social groups, along with a frequent failure to be clear on the countries from which the data were obtained and also the nativity of the participants. As a result, it is difficult to conclude how transportable the epigenetic clocks with poorly characterized sociodemographic data may be and which social groups they might apply to. Future epigenetic research should ensure to report these important participant characteristics, in combination, to contextualize their work; to properly investigate health inequities, we recommend that at a minimum researchers should collect and report both individual-level and structural-level data as one of our authors has previously suggested [78]. Researchers working with existing methods should ensure they check (where possible) the characteristics of the participants used to generate the clocks against their own population of study. They should also be mindful of the possibility of inaccurate prediction if the population the clock was developed in does differ (or is unknown), and ensure to report this as part of any published work. With the increasing use of epigenetic clocks to conduct work into social determinants of health, an important piece of future work would be to obtain primary data study (where available) to ascertain a more complete picture of the populations in which these epigenetic clocks were developed and what impacts this may have had on the conclusions of subsequent studies using the clocks. This is particularly important to enable studies to address inequalities in health.

Methods

We included the 13 epigenetic clocks discussed in a recent review that either provide the CpG sites used to construct the clock or provide the means to calculate it [7] – this includes all clocks commonly used in the literature. We extracted participant sociodemographic and socioeconomic characteristics, as reported in the original clock papers and all associated Supplementary material. Where applicable, we extracted information separately for training and test data, as the DNAm data in which the clocks were trained are the most pertinent information. For the secondgeneration clocks, if biomarkers were selected in a separate cohort, we also extracted information about that cohort. We extracted a number of social characteristics of participants that are important for understanding inequalities in health. The participant characteristics we extracted, as characterized in the studies, pertained to age, biological specimen collection date, and birth cohort; sex or gender; racialized groups (including 'race', 'ethnicity', and 'ancestry'); nativity (whether an individual was born in the country of recruitment); country context (the country in which the participants were recruited, identifying the societal structures in which people live); socioeconomic position (e.g. as measured by income, social class); and education; as well as reported validation across social groups. Extracting the data, we use the terms used

in the original paper and note where terminology is problematic. Where age and either specimen collection date or birth cohort were reported, we calculated the missing value using the available data.

Data availability

All data used in this manuscript are available in the original papers that developed the 13 epigenetic clock methods (all references are contained in Supplementary Table S2).

Supplementary data

Supplementary data are available at EnvEpig online.

Funding

This work was funded by the National Institutes of Health/ National Institute on Minority Health and Health Disparities (SR01MD014304). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. G.D.S., C.R., K.T., M.S., and S.H.W. work within the Medical Research Council Integrative Epidemiology Unit at the University of Bristol, which is supported by the Medical Research Council (MC_UU_00011/1, MC_UU_00011/3, and MC_UU_00011/5).

Conflict of interest statement. The authors declare no conflict of interest.

Author contributions

N.K., C.R., J.T.C., I.D.V., A.V.D.R., M.S., S.H.W., P.D.W., C.T., and A.J.S. conceived the study. All authors contributed to the design of the review. S.H.W. conducted the review and extracted the data. S.H.W. and N.K. summarized the results. S.H.W. wrote the manuscript. All authors edited and approved the manuscript.

References

- Horvath S. DNA methylation age of human tissues and cell types. Genome Biol 2013;14:R115.
- Locke WJ, Guanzon D, Ma C et al. DNA methylation cancer biomarkers: translation to the clinic. Front Genet 2019;10:1150.
- Joehanes R, Just AC, Marioni RE et al. Epigenetic signatures of cigarette smoking. Circ Cardiovasc Genet 2016;9:436–47.
- Houtepen LC, Hardy R, Maddock J et al. Childhood adversity and DNA methylation in two population-based cohorts. Transl Psychiatry 2018;8:266.
- Dunn EC, Soare TW, Zhu Y et al. Sensitive periods for the effect of childhood adversity on DNA methylation: results from a prospective, longitudinal study. Biol Psychiatry 2019;85:838–49.
- Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. Nat Rev Genet 2018;19:371–84.
- Bergsma T, Rogaeva E. DNA methylation clocks and their predictive capacity for aging phenotypes and healthspan. *Neurosci Insights* 2020;15:2633105520942221.
- Hannum G, Guinney J, Zhao L et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol Cell 2013;49:359–67.
- 9. Oblak L, van der Zaag J, Higgins-Chen AT *et al*. A systematic review of biological, social and environmental factors associated with epigenetic clock acceleration. *Ageing Res Rev* 2021;**69**:101348.

- Adams JM, White M. Biological ageing: a fundamental, biological link between socio-economic status and health? *Eur J Public Health* 2004;**14**:331–4.
- George A, Hardy R, Castillo Fernandez J et al. Life course socioeconomic position and DNA methylation age acceleration in mid-life. J Epidemiol Community Health 2021;75:1084–90.
- Fiorito G, McCrory C, Robinson O et al. Socioeconomic position, lifestyle habits and biomarkers of epigenetic aging: a multi-cohort analysis. Aging (Albany NY) 2019;11:2045–70.
- Martin CL, Ward-Caviness CK, Dhingra R et al. Neighborhood environment, social cohesion, and epigenetic aging. Aging (Albany NY) 2021;13:7883–99.
- 14. Lawrence KG, Kresovich JK, O'Brien KM et al. Association of neighborhood deprivation with epigenetic aging using 4 clock metrics. JAMA Netw Open 2020;**3**:e2024329.
- 15. Simons RL, Lei MK, Klopack E *et al.* The effects of social adversity, discrimination, and health risk behaviors on the accelerated aging of African Americans: further support for the weathering hypothesis. Soc Sci Med 2021;**282**:113169.
- 16. Bell CG, Lowe R, Adams PD et al. DNA methylation aging clocks: challenges and recommendations. *Genome* Biol 2019;**20**:249.
- Field AE, Robertson NA, Wang T et al. DNA methylation clocks in aging: categories, causes, and consequences. Mol Cell 2018;71:882–95.
- 18. Liu Z, Leung D, Thrush K et al. Underlying features of epigenetic aging clocks in vivo and in vitro. Aging Cell 2020;**19**:e13229.
- Barcelona de Mendoza V, Huang Y, Crusto CA et al. Perceived racial discrimination and DNA methylation among African American women in the InterGEN study. Biol Res Nurs 2018;20:145–52.
- 20. Cerutti J, Lussier AA, Zhu Y et al. Associations between indicators of socioeconomic position and DNA methylation: a scoping review. Clin Epigenetics 2021;13:221.
- 21. Wu Y, Qie R, Cheng M et al. Air pollution and DNA methylation in adults: A systematic review and meta-analysis of observational studies. Environ Pollut 2021;284:117152.
- 22. Colicino E, Just A, Kioumourtzoglou MA et al. Blood DNA methylation biomarkers of cumulative lead exposure in adults. J Expo Sci Environ Epidemiol 2021;**31**:108–16.
- 23. Galanter JM, Gignoux CR, Oh SS *et al*. Differential methylation between ethnic sub-groups reflects the effect of genetic ancestry and environmental exposures. *eLife* 2017;**6**:e20532.
- 24. Karlsson Linner R, Marioni RE, Rietveld CA et al. An epigenomewide association study meta-analysis of educational attainment. Mol Psychiatry 2017;22:1680–90.
- Diez Roux AV. Social epidemiology: past, present, and future. Annu Rev Public Health 2022;43:79–98.
- 26. Krieger N. Ecosocial Theory, Embodied Truths, and the People's Health. New York, NY: Oxford University Press, 2021.
- 27. Noren Hooten N, Pacheco NL, Smith JT et al. The accelerated aging phenotype: the role of race and social determinants of health on aging. Ageing Res Rev 2022;**73**:101536.
- Tyrrell J, Zheng J, Beaumont R et al. Genetic predictors of participation in optional components of UK Biobank. Nat Commun 2021;12:886.
- 29. Taylor AE, Jones HJ, Sallis H et al. Exploring the association of genetic factors with participation in the Avon Longitudinal Study of Parents and Children. Int J Epidemiol 2018;47:1207–16.
- Munafo MR, Tilling K, Taylor AE et al. Collider scope: when selection bias can substantially influence observed associations. Int J Epidemiol 2018;47:226–35.

- 31. Lamers F, Hoogendoorn AW, Smit JH et al. Sociodemographic and psychiatric determinants of attrition in the Netherlands Study of Depression and Anxiety (NESDA). Compr Psychiatry 2012;53:63–70.
- 32. Nwanaji-Enwerem JC, Jackson CL, Ottinger MA et al. Adopting a "compound" exposome approach in environmental aging biomarker research: a call to action for advancing racial health equity. Environ Health Perspect 2021;129:45001.
- 33. Chen IY, Pierson E, Rose S et al. Ethical machine learning in healthcare. Annu Rev Biomed Data Sci 2021;4:123–44.
- 34. Chen IY, Szolovits P, Ghassemi M. Can AI help reduce disparities in general medical and mental health care? AMA J Ethics 2019;21:E167–79.
- 35. Wiens J, Saria S, Sendak M et al. Do no harm: a roadmap for responsible machine learning for health care. Nat Med 2019;25:1337–40.
- 36. Philibert R, Beach SRH, Lei MK et al. Array-based epigenetic aging indices may be racially biased. Genes (Basel) 2020;11:685.
- 37. Breeze CE, Beck S, Berndt SI et al. The missing diversity in human epigenomic studies. Nat Genet 2022;**54**:737–9.
- Breeze CE, Wong JYY, Beck S et al. Diversity in EWAS: current state, challenges, and solutions. Genome Med 2022;14:71.
- 39. Hughes A, Smart M, Gorrie-Stone T et al. Socioeconomic position and DNA methylation age acceleration across the life course. Am J Epidemiol 2018;187:2346–54.
- 40. Dugue PA, Bassett JK, Joo JE et al. Association of DNA methylationbased biological age with health risk factors and overall and cause-specific mortality. Am J Epidemiol 2018;187:529–38.
- 41. Stevenson AJ, McCartney DL, Hillary RF et al. Childhood intelligence attenuates the association between biological ageing and health outcomes in later life. Transl Psychiatry 2019;9:323.
- 42. Zhao W, Ammous F, Ratliff S et al. Education and lifestyle factors are associated with DNA methylation clocks in older African Americans. Int J Environ Res Public Health 2019;16:3141.
- 43. Fiorito G, Polidoro S, Dugue PA et al. Social adversity and epigenetic aging: a multi-cohort study on socioeconomic differences in peripheral blood DNA methylation. Sci Rep 2017;7:16266.
- 44. Robinson O, Chadeau Hyam M, Karaman I *et al.* Determinants of accelerated metabolomic and epigenetic aging in a UK cohort. *Aging Cell* 2020;**19**:e13149.
- 45. Simons RL, Lei MK, Beach SR et al. Economic hardship and biological weathering: the epigenetics of aging in a U.S. sample of black women. Soc Sci Med 2016;150:192–200.
- 46. Austin MK, Chen E, Ross KM et al. Early-life socioeconomic disadvantage, not current, predicts accelerated epigenetic aging of monocytes. Psychoneuroendocrinology 2018;97:131–4.
- 47. Lawn RB, Anderson EL, Suderman M et al. Psychosocial adversity and socioeconomic position during childhood and epigenetic age: analysis of two prospective cohort studies. *Hum Mol Genet* 2018;**27**:1301–8.
- 48. Ryan CP, Hayes MG, Lee NR et al. Reproduction predicts shorter telomeres and epigenetic age acceleration among young adult women. Sci Rep 2018;8:11100.
- 49. Miller GE, Yu T, Chen E et al. Self-control forecasts better psychosocial outcomes but faster epigenetic aging in low-SES youth. Proc Natl Acad Sci USA 2015;112:10325–30.
- 50. Schmitz LL, Zhao W, Ratliff SM et al. The socioeconomic gradient in epigenetic ageing clocks: evidence from the multi-ethnic study of atherosclerosis and the health and retirement study. *Epigenetics* 2022;**17**:589–611.
- 51. Crimmins EM, Thyagarajan B, Levine ME et al. Associations of age, sex, race/ethnicity, and education with 13 epigenetic clocks in a

nationally representative U.S. sample: the health and retirement study. J Gerontol A Biol Sci Med Sci 2021;**76**:1117–23.

- 52. Horvath S, Gurven M, Levine ME et al. An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. *Genome Biol* 2016;**17**:171.
- 53. Tajuddin SM, Hernandez DG, Chen BH et al. Novel age-associated DNA methylation changes and epigenetic age acceleration in middle-aged African Americans and whites. Clin Epigenetics 2019;11:119.
- 54. Graf GH, Crowe CL, Kothari M et al. Testing Black-White disparities in biological aging among older adults in the United States: analysis of DNA-methylation and blood-chemistry methods. Am J Epidemiol 2022;191:613–25.
- 55.Lu AT, Quach A, Wilson JG et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. Aging (Albany NY) 2019;11:303–27.
- 56. Liu Z, Chen BH, Assimes TL et al. The role of epigenetic aging in education and racial/ethnic mortality disparities among older U.S. women. Psychoneuroendocrinology 2019;104:18–24.
- 57. Weidner CI, Lin Q, Koch CM *et al*. Aging of blood can be tracked by DNA methylation changes at just three CpG sites. *Genome Biol* 2014;**15**:R24.
- 58. Zhang Q, Vallerga CL, Walker RM et al. Improved precision of epigenetic clock estimates across tissues and its implication for biological ageing. Genome Med 2019;11:54.
- 59. Yang Z, Wong A, Kuh D et al. Correlation of an epigenetic mitotic clock with cancer risk. *Genome* Biol 2016;**17**:205.
- 60. Bocklandt S, Lin W, Sehl ME et al. Epigenetic predictor of age. PLoS One 2011;**6**:e14821.
- Koch CM, Wagner W. Epigenetic-aging-signature to determine age in different tissues. Aging (Albany NY) 2011;3:1018–27.
- 62. Zhang Y, Wilson R, Heiss J et al. DNA methylation signatures in peripheral blood strongly predict all-cause mortality. Nat Commun 2017;8:14617.
- 63. Levine ME, Lu AT, Quach A et al. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY) 2018;10:573–91.
- 64. Lu AT, Seeboth A, Tsai PC et al. DNA methylation-based estimator of telomere length. Aging (Albany NY) 2019;11:5895–923.
- 65. Belsky DW, Caspi A, Arseneault L et al. Quantification of the pace of biological aging in humans through a blood test, the DunedinPoAm DNA methylation algorithm. eLife 2020;9:e54870.
- 66. Youn A, Wang S. The MiAge calculator: a DNA methylationbased mitotic age calculator of human tissue types. *Epigenetics* 2018;**13**:192–206.
- 67. Mauvais-Jarvis F, Bairey Merz N, Barnes PJ et al. Sex and gender: modifiers of health, disease, and medicine. *Lancet* 2020;**396**:565–82.
- 68. Krieger N. Genders, sexes, and health: what are the connections and why does it matter? Int J Epidemiol 2003;32:652–7.
- 69. Krieger N. Measures of racism, sexism, heterosexism, and gender binarism for health equity research: from structural injustice to embodied harm – an ecosocial analysis. Annu Rev Public Health 2020;**41**:37–62.
- 70. Nielsen MW, Stefanick ML, Peragine D et al. Gender-related variables for health research. Biol Sex Differ 2021;**12**:23.
- 71. Krieger N. Stormy weather: race, gene expression, and the science of health disparities. *Am J Public Health* 2005;**95**:2155–60.
- 72. Augstein HF. From the land of the Bible to the Caucasus and beyond: the shifting ideas of the geographical origin of humankind. In: Ernst W, Harris B (eds.), Race, Science and Medicine, 1700-1960. London: Routledge, 1999, 58–79.

- 73. Painter NI. The History of White People. New York, NY: W.W. Norton, 2011.
- 74. Evans L, Engelman M, Mikulas A et al. How are social determinants of health integrated into epigenetic research? A systematic review. Soc Sci Med 2021;273: 113738.
- 75. Vyas DA, Eisenstein LG, Jones DS. Hidden in plain sight reconsidering the use of race correction in clinical algorithms. N Engl J Med 2020;383:874–82.
- 76. Obermeyer Z, Powers B, Vogeli C et al. Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019;**366**:447–53.
- 77. Krieger N, Jahn JL. Tumor specimen biobanks: data gaps for analyzing health inequities – the case of breast cancer. JNCI Cancer Spectr 2018;**2**:ky011.
- 78. Krieger N. Structural racism, health inequities, and the two-edged sword of data: structural problems require structural solutions. Front Public Health 2021;9:655447.