A Section 508–conformant HTML version of this article is available at https://doi.org/10.1289/EHP8793.

# Surrounding Greenness and Biological Aging Based on DNA Methylation: A Twin and Family Study in Australia

*Rongbin Xu,[1] Shuai Li,[2,3,4] Shanshan Li,[1] Ee Ming Wong,[4,5] Melissa C. Southey,[4,5,6] John L. Hopper,[2] Michael J. Abramson,[1] and Yuming Guo[1]*

[1]School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia
[2]Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Victoria, Australia
[3]Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK
[4]Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, Victoria, Australia
[5]Department of Clinical Pathology, Melbourne Medical School, University of Melbourne, Melbourne, Victoria, Australia
[6]Cancer Epidemiology Division, Cancer Council Victoria, Victoria, Australia

**Background:** High surrounding greenness has many health benefits and might contribute to slower biological aging. However, very few studies have evaluated this from the perspective of epigenetics.

**Objectives:** We aimed to evaluate the association between surrounding greenness and biological aging based on DNA methylation.

**Methods:** We derived Horvath's DNA methylation age (DNAmAge), Hannum's DNAmAge, PhenoAge, and GrimAge based on DNA methylation measured in peripheral blood samples from 479 Australian women in 130 families. Measures of DNAmAge acceleration (DNAmAgeAC) were derived from the residuals after regressing each DNAmAge metric on chronological age. Greenness was represented by satellite-derived Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI) metrics within 300-, 500-, 1,000-, and 2,000-m buffers surrounding participant addresses. Greenness-DNAmAgeAC associations were estimated using a within-sibship design fitted by linear mixed effect models, adjusting for familial clustering and important covariates.

**Results:** Greenness metrics were associated with significantly lower DNAmAgeAC based on GrimAge acceleration, suggesting slower biological aging with higher greenness based on both NDVI and EVI in 300–2,000 m buffer areas. For example, each interquartile range increase in NDVI within 1,000 m was associated with a 0.59 (95% CI: 0.18, 1.01)–year decrease in GrimAge acceleration. Greenness was also inversely associated with three of the eight components of GrimAge, specifically, DNA methylation-based surrogates of serum cystatin-C, serum growth differentiation factor 15, and smoking pack years. Associations between greenness and biological aging measured by Horvath's and Hannum's DNAmAgeAC were less consistent, and depended on neighborhood socioeconomic status. No significant associations were estimated for PhenoAge acceleration.

**Discussion:** Higher surrounding greenness was associated with slower biological aging, as indicated by GrimAge age acceleration, in Australian women. Associations were also evident for three individual components of GrimAge, but were inconsistent for other measures of biological aging. Additional studies are needed to confirm our results. https://doi.org/10.1289/EHP8793

## Introduction

As a result of decreased fertility and improved survival, the population is aging rapidly across the world (United Nations 2017). The global population aged ≥60 y has increased from 382 million in 1980 to 962 million in 2017 and is predicted to double again by 2050, reaching nearly 2.1 billion (United Nations 2017). With this rapid demographic shift, it is important to better understand the determinants of the biological aging process and how to achieve healthy aging (WHO 2015). Several biomarkers of biological aging have been proposed to predict the functional capability of a person, and the DNA methylation age (DNAmAge) has been suggested as the most robust (Ferrucci et al. 2020; Jylhävä et al. 2017).

DNAmAge is a measurement of biological aging based on DNA methylation, a biological process that regulates gene expression by adding methyl groups to the cytosine bases of the DNA (Horvath and Raj 2018). Four main algorithms are used to calculate DNAmAge: Horvath's age (Horvath 2013), Hannum's age (Hannum et al. 2013), PhenoAge (Levine et al. 2018), and GrimAge (Lu et al. 2019). The DNAmAge acceleration (DNAmAgeAC) can be estimated as the residuals after regressing individual measures of DNAmAge on chronological age (the calendar time that has passed since birth). Increased DNAmAgeAC has been positively associated with mortality and conditions related to aging, such as cancer and cardiovascular diseases (Horvath and Raj 2018; Levine et al. 2018; Lu et al. 2019). DNAmAgeAC calculated from the most recently developed estimator (GrimAge) showed the strongest association (GrimAge > PhenoAge > Hannum > Horvath) with deaths, cancers, coronary heart disease, and early age at menopause in an analysis of longitudinal data from 6,395 adults (mean age: 63.0 y) in four cohorts (Framingham Heart Study, Women's Health Initiative, Jackson Heart Study, and InCHIANTI Study) followed for 13.7 y on average (Lu et al. 2019).

An important feature of DNAmAge is that DNA methylation is modifiable by interventions and environmental factors, making it a useful tool for identifying or validating healthy aging determinants (Horvath and Raj 2018). One of the potentially important determinants is surrounding greenness, also called residential green space or surrounding vegetation (Reid et al. 2018), referring to a measurement of the density of vegetated land (e.g., gardens, parks, grassland) in surrounding areas (Taylor and Hochuli 2017). Greenness is a critical component of a healthy built environment. A high level of surrounding greenness may decrease mental stress, facilitate social interaction, boost physical activity, and reduce harm from noise, air pollution, and heat exposure (James et al. 2015; Markevych et al. 2017). Numerous epidemiological studies have reported evidence supporting health benefits of high surrounding greenness, including

Address correspondence to Yuming Guo, School of Public Health and Preventive Medicine, Monash University, Level 2, 553 St. Kilda Rd., Melbourne, VIC 3004 Australia. Email: yuming.guo@monash.edu

Supplemental Material is available online (https://doi.org/10.1289/EHP8793).

M.J.A. holds investigator-initiated grants from Pfizer and Boehringer-Ingelheim for unrelated research. He has undertaken an unrelated consultancy for and received assistance with conference attendance from Sanofi. He also received a speaker's fee from GlaxoSmithKline. All other authors declare they have no actual or potential competing financial interests.

Received 9 December 2020; Revised 6 August 2021; Accepted 10 August 2021; Published 30 August 2021.

slower cognitive decline and reduced risks of mental illness, obesity, cardiovascular diseases, type 2 diabetes, and death (de Keijzer et al. 2018, 2020; Fong et al. 2018; James et al. 2015; Jia et al. 2018; Lachowycz and Jones 2011; Rojas-Rueda et al. 2019; Sarkar 2017).

To the best of our knowledge, only one study of infants, which was reported as a conference abstract, has evaluated the relationship between greenness and DNA methylation-based measures of biological aging (Sbihi et al. 2018). The study estimated gestational age (GA) based on DNA methylation measured in cord blood samples and reported a borderline association between greenness surrounding the maternal residence and DNA methylation-based GA acceleration (i.e., older DNA methylation GA compared with chronological GA) {0.6 d, 95% [confidence interval (CI)]: −0.6 to 1.8, for one interquartile range (IQR) exposure increase}. This suggested that high greenness may accelerate development *in utero* (Sbihi et al. 2018).

Given the evidence supporting the benefits of high surrounding greenness on health, we aimed to evaluate whether high surrounding greenness can slow biological aging in adults. This can be indicated by inverse associations between greenness exposures and DNA methylation-based measures of biological aging.

## Methods

### *Study Design and Population*

We analyzed data from the Australian Mammographic Density Twins and Sisters Study (AMDTSS), a twin and family study population (Stone et al. 2007). The AMDTSS was originally designed to investigate genetic, environmental, and lifestyle factors associated with mammographic density, a major risk factor for breast cancer (Boyd et al. 2002). In 1995–1999, female twins in three Australian cities (Melbourne, Sydney, and Perth) were recruited through the Australian Twin Registry. In 2004–2009, the twins were invited to participate further, and their sisters were invited to participate as well. The study included participants 40–70 years of age and excluded pregnant women or nursing mothers and women with a history of breast cancer. Written informed consent was obtained from all participants. Participants were asked to complete a telephone-administered questionnaire and, if they consented to provide a blood sample, were sent a blood collection kit with a shipping container, tubes, and detailed instructions and asked to visit a nearby contracted pathology laboratory where a 27-mL peripheral whole-blood sample was collected. In the case a participant was unable to travel to a pathology laboratory, a trained phlebotomist visited their home to perform the blood collection. All blood samples were delivered to the Genetic Epidemiology Laboratory at The University of Melbourne within 48 h after collection. These samples were then processed to generate Guthrie cards, which were used to measure DNA methylation.

The AMDTSS recruited over 2,700 participants during 2004–2009. All AMDTSS participants with DNA methylation measurements were eligible for the present study, which comprised 479 women from 130 families, including 66 monozygotic twin pairs, 66 dizygotic twin pairs, and 215 sisters of these twins. Each family had 3–6 participants. Each eligible participant had one DNA methylation measurement. The AMDTSS was approved by the Human Research Ethics Committee of The University of Melbourne. The present analysis was approved by the Monash University Human Research Ethics Committee.

### *DNA Methylation Data*

DNA methylation was measured in peripheral blood samples using the Illumina Infinium HumanMethylation450 BeadChip array, as detailed in our previous study (Li et al. 2015). We processed the raw methylation data with the Bioconductor *minfi* package (Aryee et al. 2014). This process included normalization of data using Illumina's reference factor-based normalization methods (*preprocessIllumina*) and correction of type I and II probe bias using subset-quantile within array normalization (SWAN) (Maksimovic et al. 2012). We used an empirical Bayes batch-effects removal method, ComBat (Johnson et al. 2007), to minimize the technical variation across batches. After excluding 65 probes corresponding to known single nucleotide polymorphisms, and probes with detection $p > 0.01$ in one or more samples, 479,957 probes were left for further analyses (Li et al. 2015). The methylation level of each cytosine–phosphate–guanine site (CpG) was represented by a probe's beta value ranging from 0 (not methylated) to 1 (fully methylated).

The processed DNA methylation data were used to calculate DNAmAge and cell-type proportions (i.e., proportions of leukocyte cells) in peripheral blood. Four measures of biological age based on methylation of specific CpGs were derived using an online calculator (http://dnamage.genetics.ucla.edu/new): Horvath's age, based on 353 CpGs (Horvath 2013); Hannum's age, based on 71 CpGs (Hannum et al. 2013); Pheno age, based on 513 CpGs (Levine et al. 2018); and Grim age, based on 1,030 CpGs (Lu et al. 2019). In our processed DNA methylation data set, methylation data were missing for 1 CpGs used to calculate Horvath's age, 2 CpGs used to calculate PhenoAge, and 21 CpGs used to calculate GrimAge, whereas methylation data were complete for all CpGs used to derive Hannum's age. The methylation levels for the missing CpGs were imputed by the online calculator using the default $k$ nearest neighbor (KNN) method through the *impute.knn* function, with default parameters embedded in the *impute* R package (Di Lena et al. 2019; Hastie et al. 2018; Horvath 2013; Troyanskaya et al. 2001). In a simulation study using data from 1,495 healthy control samples in the Gene Expression Omnibus, values of Horvath's DNAmAge derived using KNN to impute methylation data that were missing completely at random for 10% of the CpGs used to derive the metric were consistent with Horvath's DNAmAge estimates based on the complete data (Pearson correlation coefficient = 0.99, mean absolute error = 1.8 y) (Di Lena et al. 2019). Proportions of $CD8^+$ T cells, $CD4^+$ T cells, natural killer (NK) cells, B cells, monocytes, and granulocytes were estimated using a widely used published algorithm (Houseman et al. 2012). The proportions of exhausted CD8 T cells, naïve $CD8^+$ T cells, naïve $CD4^+$ T cells, and plasma cells (effector B cells) were estimated using the algorithm embedded in the online DNAmAge calculator (Horvath 2013). The online calculator also provides separate estimates for the eight individual components that contribute to GrimAge, including smoking pack years and the concentration of seven plasma proteins [adrenomedullin, beta-2-microglobulin, cystatin-C, growth differentiation factor 15 (GDF-15), leptin, plasminogen activator inhibitor 1 (PAI-1), and tissue inhibitor metalloproteinases 1 (TIMP-1)], which were all estimated by DNA methylation data based on the model trained in Framingham Heart Study samples (Lu et al. 2019).

### *Exposure Assessment*

Participants who agreed to donate a blood sample provided residential address information (state, city, suburb, street name, and house number) for shipment of the blood collection kit. We transformed participant addresses into longitudes and latitudes using the Google Map Application Programming Interface (API) through the *geocode* function embedded in *ggmap* R package (version 3.0.0; R Development Core Team). These longitudes and latitudes were then used for exposure assessment. According to the 2011 Australian Statistical Geography Standard, 84% (i.e., 403/479) of the residential addresses were in urban areas (ABS 2011, 2020). Because residential histories were not collected, we

assumed that the participants lived at the same residence during the 12 months before their blood draw. This assumption is supported by a national survey conducted in 2007–2008 that reported that nearly 80% of Australians 40–79 years of age had not changed residences in the last 5 y (ABS 2010). At the time of blood draw, 12 (2.5%) participants (two monozygotic twin pairs, two dizygotic twin pairs, and two non-twin sister pairs) lived at the same residential address with her twin or sister.

Residential surrounding greenness was measured using the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI), which were derived from Moderate resolution Imaging Spectroradiometer (MODIS) images collected from the National Aeronautics and Space Administration's Terra satellite (Didan 2015). Green vegetation strongly absorbs visible red light (wavelengths from 0.6 to 0.7 μm) for photosynthesis, whereas it reflects most near-infrared light (wavelengths from 0.7 to 1.1 μm). According to this characteristic, NDVI is calculated as $\mathrm{NDVI} = (\mathrm{NIR} - \mathrm{Red})/(\mathrm{NIR} + \mathrm{Red})$, where Red and NIR represent land surface reflectance of red light and near-infrared light measured by satellite sensors, respectively. The EVI is calculated similarly to the NDVI, but it corrects for distortions in the reflected light due to atmospheric aerosols and the land surface background beneath the vegetation. The EVI has a higher sensitivity to the vegetation difference in high biomass regions (e.g., forests) than the NDVI. The MODIS provides gridded NDVI and EVI values at a $250 \times 250$ m spatial resolution for every 16 d since 2000. The 16-d composites were created on a per-pixel basis, using the Constrained Maximum Value Composite algorithm. Briefly, the algorithm selects the best observations from all images within 16 d, minimizing the effects of image quality, cloud, and view angle. Both NDVI and EVI values range from −1 to 1, with negative values representing ice, water, and nonvegetated soil and higher positive values corresponding to higher greenness.

Based on the 16-d gridded NDVI/EVI data set, we calculated the residential surrounding greenness in several steps. First, we selected all 16-d NDVI/EVI gridded observations at a 250 m resolution during 2003–2009 across Australia. Second, we converted negative NDVI/EVI values in all pixels to zero so that negative values would not offset positive values in calculating the areal average (Reid et al. 2018). Third, we calculated the surrounding NDVIs/EVIs within different radial buffers (300, 500, 1,000, and 2,000 m) centered on each participant's home address for every 16 d, by averaging the values of all pixels weighted by area size covered by the buffer areas. We chose 300-, 500-, 1,000-, and 2,000-m buffers, because the health impacts of surrounding greenness are most likely to be evident within these buffers (Browning and Lee 2017; Reid et al. 2018). Finally, to reduce the influence of seasonal variations across different locations, we classified exposures for each index and radial buffer using the highest 16-d NDVI (or EVI) during the year (365 d) before each participant's blood draw.

### *Covariate Data*

Through a telephone-administered questionnaire, we collected data on birth date, educational level, marital status, smoking behavior, and alcohol use behavior (see the section "Variables measured by telephone administrated questionnaire" in the Supplemental Material for details). For 95% of the participants (457/479), the telephone interview was conducted within 1 y before the blood draw, and the other participants' telephone interviews were conducted within 1–2 y before the blood draw or up to 3 months after the blood draw (Figure S1). Chronological age was defined as the calendar time period between birth date and the day when the blood sample was collected. Area-level socioeconomic status (SES) was represented by the percentile of Index of Relative Socioeconomic Advantage and Disadvantage (IRSAD) in 2006 for the population census collection district (CD) in which the participant lived. The IRSAD is a linear combination of multiple indicators of community-level SES (e.g., household income, educational level, occupation, housing conditions) collected from the population census in Australia (ABS 2006). A higher IRSAD represents a higher area-level SES. The CD was the smallest geographic area for which the IRSAD data were available. In our study, the participants were distributed across 366 CDs, with the usual population size ranging from 82 to 1,686.

We collected daily ambient temperature and relative humidity from a gridded climate data set at $0.05^\circ \times 0.05^\circ$ (about $5 \times 5$ km) spatial resolution (Jeffrey et al. 2001). Each participant was assigned the values of the pixel where her home address was located. For each participant, we used the average ambient temperature and relative humidity within 1 y (365 d) prior to the day blood was taken in order to represent the participant's exposure to meteorological factors.

### *Statistical Analysis*

***Main model analyses.*** We used the same procedure to estimate associations between each DNAmAgeAC metric and surrounding greenness metric. First, we calculated the DNAmAgeAC as residuals after regressing DNAmAge on chronological age (Horvath and Raj 2018; Levine et al. 2018; Lu et al. 2019). Then, we estimated associations between greenness and DNAmAgeAC using a within-sibship design (Li et al. 2017; Stone et al. 2010) fitted by a linear mixed-effect model, with random family-specific intercepts to capture familial clustering:

$$E(DNAmAgeAC_{ij}) = \beta_0 + \beta_{within} \times (Green_{ij} - \overline{Green_i}) + \beta_{between} \times \overline{Green_i} + \beta \times X_{ij} + \mu_i, \quad (1)$$

where $DNAmAgeAC_{ij}$ is the DNAmAge acceleration of subject $j$ from family $i$; $Green_{ij}$ $j$ is the surrounding greenness index for subject $j$ from family $i$ (e.g., NDVI within 1,000 m); $\overline{Green_i}$ is the average surrounding greenness of all subjects in family $i$, with each subject's surrounding greenness defined according to their residential address and blood draw date; and $\mu_i$ is the random intercept for family $i$. We modeled greenness as a linear term because the linear model had a lower Bayesian information criterion (BIC) value than models of greenness as a natural cubic spline with 2–4 degrees of freedom (df) (Table S1).

$\beta_{between}$ and $\beta_{within}$ represent effect estimates (difference in mean DNAmAgeAC per IQR increase in greenness) for between- and within-sibship associations, respectively. The between-sibship association represents a cross-sectional comparison between families that may be biased by familial factors and genetic background that vary between families. By contrast, the within-sibship association represents intra-family differences between twins and sisters, thus reducing or eliminating confounding by genetic background and shared familial factors (Li et al. 2017; Stone et al. 2010). Therefore, we report $\beta_{within}$ and its 95% CI throughout this article.

$X_{ij}$ represents a set of covariates that might bias greenness–DNAmAgeAC associations (James et al. 2015; Markevych et al. 2017), including educational level (secondary or below; vocational training; university), marital status (married or de facto; never married; widowed or separated or divorced), smoking behavior (current smoker, former smoker, never smoked), alcohol use (current drinker, former drinker, never drank), and area-level SES (four categories according to quartiles in Australia); plus continuous variables for chronological age, annual mean temperature, annual mean relative humidity, and the proportions of seven types of leukocyte cells (naïve $CD8^+$ T cells, $CD8^+$ T cells,

plasma cells, $CD4^+$ T cells, NK cells, monocytes, and granulocytes), as suggested by (Lu et al. 2019). Given adjustment for cell-type proportions, effect estimates from our main model should be interpreted as associations between greenness and differences in intrinsic DNAmAgeAC (independent of differences in leukocyte proportions) (Horvath and Raj 2018).

Missing values for educational level (two participants) and area-level SES (one participant), were addressed by five multiple imputations using the default method embedded in the *mice* R package (van Buuren and Groothuis-Oudshoorn 2011), with the four DNAmAgeAC metrics, the eight greenness metrics, and the main model covariates (other than the target variable of the imputation) used as predictors.

***Subgroup analyses.*** We performed subgroup analyses by educational level (secondary or below vs. higher than secondary, including vocational training and university), area-level SES (lowest 50% vs. highest 50%), age group (<60 vs. ≥60 y), smoking behavior (never smoked vs. current or former smoker), DNA methylation-based smoking index (described below, ≤median vs. >median) (Gao et al. 2017), and alcohol use (never drank vs. current or former drinker) to evaluate the potential modification of the greenness-DNAmAgeAC association by these factors. We tested the statistical difference in effect estimates between subgroups by adding an interaction term (e.g., education–greenness) to the main model.

***Sensitivity analyses.*** We performed several sensitivity analyses. First, we further adjusted for season (four categories: spring, summer, fall, winter) and survey year (continuous variable) at the time of the blood draw to account for potential seasonal variation and long-term trends in DNA methylation. Second, we modeled annual mean temperature and relative humidity as nonlinear (vs. linear) terms using natural cubic splines with 3 df (Xu et al. 2020). Third, we repeated models without adjusting for smoking and alcohol use because there was no clear evidence that these two variables were related to greenness. Fourth, we repeated the models without adjusting for cell-type proportions in order to generate effect estimates that reflected extrinsic differences in DNAmAgeAC associated with greenness. The extrinsic differences reflect both differences in DNAmAgeAC among different leukocyte subpopulations and intrinsic differences in DNAmAgeAC independent of leukocyte type (Horvath and Raj 2018). Finally, instead of adjusting for self-reported smoking as a categorical variable (i.e., never, former, or current smoker) we adjusted for a continuous DNA methylation-based smoking index based on nine CpGs, using the algorithm provided by Gao et al. (2017), which the authors reported to be a stronger predictor of frailty in a cross-sectional sample of German adults than self-reported smoking. In our study population, the smoking index tended to be higher in current smokers than in former or never smokers, although smoking index distributions overlapped among the three smoking groups based on self-report (Figure S2). We used fixed-effects meta-regressions to test whether effect estimates from sensitivity analyses differed significantly from corresponding estimates based on the main model (Xu et al. 2019).

***Associations with components of GrimAge.*** To evaluate the association between surrounding greenness and individual components of GrimAge, we fit separate models using *Z*-scores for each component as the outcome instead of GrimAge. *Z*-score was calculated as (actual value − mean value)/ standard deviation.

All data analyses were performed using R software (version 3.5.1; R Development Core Team) with packages *nlme* (linear mixed effect model), *splines* (nature cubic splines), *mvmeta* (meta-regression), and *mice* (multiple imputation by chained equations). In this analysis, a two-sided $p < 0.05$ was considered as statistically significant.

## Results

We included 479 participants who resided in many different locations across Australia (Figure 1). All of them were female, including 132 monozygotic twins, 132 dizygotic twins, and 215 sisters of the twins from 130 families (Table 1). Their chronological ages ranged from 39.7 to 77.5 y (median = 55.7 y; IQR: 11.75 y). Among the four DNAmAgeAC estimates, the GrimAge acceleration showed the smallest variation [standard deviation (SD) = 3.12 y; minimum to maximum: −7.10 to 13.74 y; median = −0.44 y; IQR: 3.91 y], whereas the PhenoAge acceleration showed the largest variation (SD = 5.78 y; minimum to maximum: −15.45 to 21.25 y; median = 0.31 y; IQR: 8.17 y) (Table 2). The median NDVI and EVI values within 1,000 m were 0.55 (IQR: 0.164) and 0.35 (IQR: 0.119), respectively. The distributions of NDVI and EVI were similar across the four buffer zones for each metric (Table 2), and the greenness indicators were highly correlated (Pearson correlation coefficients = 0.85–0.98 and 0.82–0.97 among NDVI and EVI measures, respectively, and 0.77–0.92 between NDVI and EVI measures) (Figure S3). Among the four DNAmAge estimates, the GrimAge showed the strongest association with chronological age (Pearson correlation coefficients = 0.88 for GrimAge, 0.72 for Hannum's age, 0.66 for Horvath's age, 0.62 for PhenoAge) (Figures S3–S4). The four DNAmAge and DNAmAgeAC estimates were also correlated with each other (Pearson correlation coefficients = 0.67–0.80 between DNAmAge estimates, and 0.26–0.64 between DNAmAgeAC estimates). GrimAge acceleration was positively associated with the eight individual components of GrimAge, with Pearson correlation coefficients ranging from 0.19 for DNA methylation-based serum TIMP-1 to 0.80 for DNA methylation-based smoking pack years (Figure S5).

GrimAge acceleration showed consistent significant (all $p < 0.05$) inverse associations with NDVI and EVI metrics across all buffers (Figure 2). Each IQR increase in NDVI value within 1,000 m was associated with a 0.59-y decrease (95% CI: −1.01, −0.18; $p = 0.005$) in GrimAge acceleration, whereas the effect estimates were slightly weaker for other greenness metrics. However, none of the greenness exposure metrics were significantly associated with DNAmAgeAC based on Horvath's age, Hannum's age, or PhenoAge (all $p > 0.10$) in the study sample as a whole.

There were significant interactions between surrounding greenness and area-level SES on Horvath's and Hannum's DNAmAgeAC. In general, IQR increases in surrounding greenness were associated with increased Horvath's and Hannum's DNAmAgeAC among participants whose area-level SES was below the median of Australia [e.g., for an IQR increase in EVI within 300 m, 1.21 y (95% CI: 0.14, 2.28) and 0.87 y (95% CI: 0.05, 1.70), respectively], whereas corresponding estimates suggested slower age acceleration with increased greenness among those with a higher area-level SES [i.e., −0.43 y (95% CI: −1.35, 0.50; $p_{interaction} = 0.03$) for Horvath's DNAmAgeAC and −0.86 y (95% CI: −1.64, −0.08; $p_{interaction} = 0.004$) for Hannum's DNAmAgeAC] (Figure 3; Table S2). For GrimAge acceleration, there was a consistent pattern of inverse associations with greenness among those with a higher area-level SES, and null or weaker inverse associations among those with a lower area-level SES, although the differences were not statistically significant ($p_{interaction} = 0.10$–$0.66$). Associations with PhenoAge acceleration varied, without consistent patterns by greenness metric or area-level SES ($p_{interaction} \geq 0.28$).

Differences in associations with Horvath's and Hannum's DNAmAgeAC by education were somewhat consistent with differences by area-level SES, with increased greenness tending to have positive or null associations with age acceleration among those with lower education, and inverse or null associations among those with a higher education, although the differences

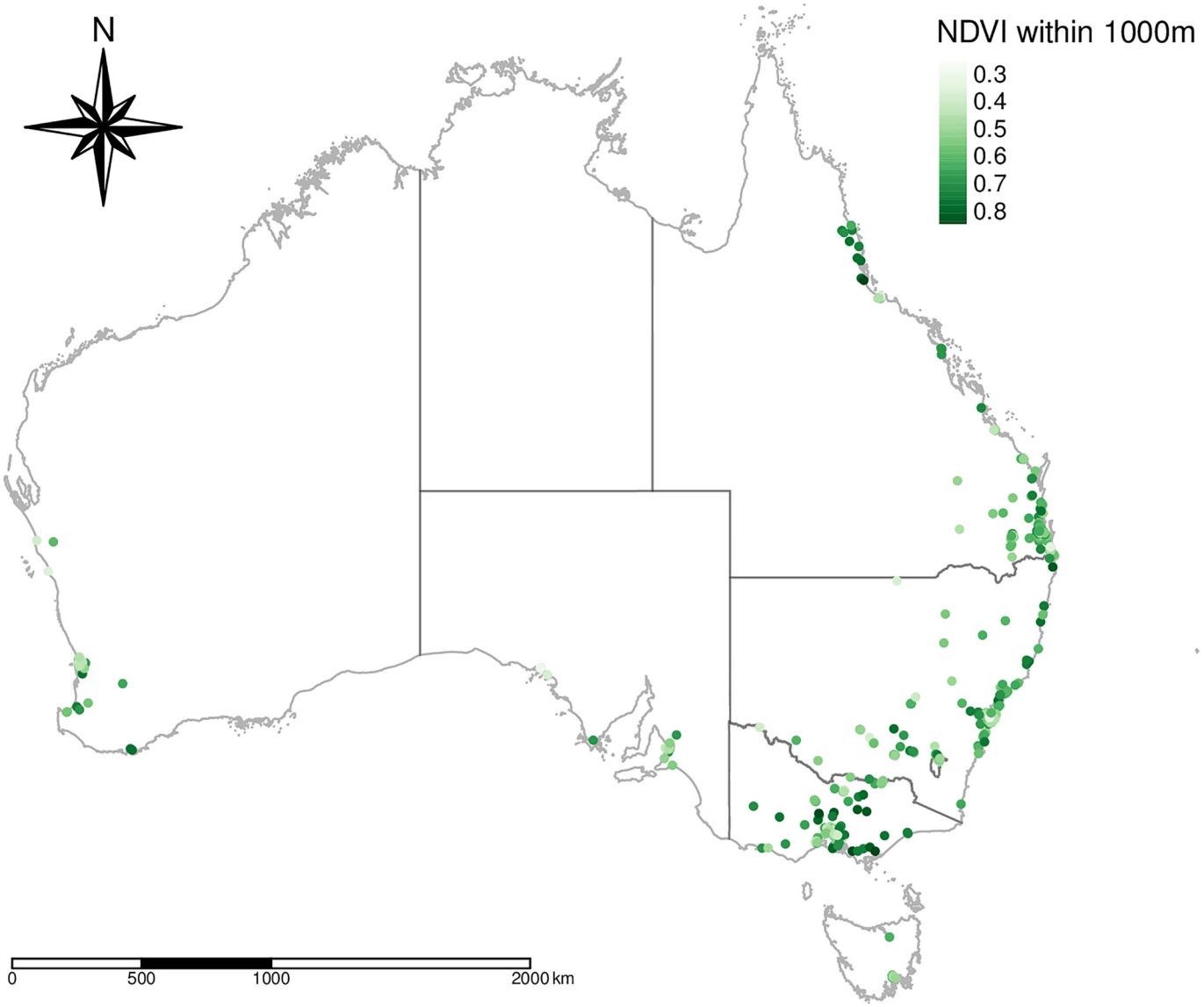


**Figure 1.** The geographical distribution and level of surrounding greenness of the 479 participants in Australia during 2004–2009. Note: NDVI, Normalized Difference Vegetation Index.

were not statistically significant ($p_{interaction} \geq 0.08$ and $\geq 0.23$, respectively) (Figure S6, Table S2). For GrimAge acceleration, associations were very similar between those with lower and higher education ($p_{interaction} \geq 0.26$), whereas associations with PhenoAge acceleration varied, without clear differences by education ($p_{interaction} \geq 0.17$).

Differences in associations between greenness and DNAmAgeAC by smoking varied depending on the DNAmAgeAC metric and the variable used to classify smoking. For Horvath's and Hannum's DNAmAgeAC, associations tended to be positive or null among current and former smokers (combined) and inverse or null among never smokers ($p_{interaction} =$ 0.04–0.13 and 0.03–0.24 for Horvath's and Hannum's metric, respectively) (Figure S7, Table S2). However, associations followed an opposite pattern for the DNA methylation-based smoking index, such that associations tended to be positive or null among those with a higher smoking index (above the median), and inverse or null for those with a lower smoking index ($p_{interaction} =$ 0.07–0.39 and 0.08–0.26, respectively) (Figure S8). In contrast, associations with GrimAge DNAmAgeAC were inverse and similar between current or former smokers and never smokers ($p_{interaction} \geq 0.28$) and between those with a high or low smoking index ($p_{interaction} \geq 0.28$) (Figures S7–S8, Table S2.) Associations with PhenoAge DNAmAgeAC varied by greenness metric but were similar by smoking status ($p_{interaction} \geq 0.19$) and high or low smoking index ($p_{interaction} \geq 0.50$).

Associations between greenness metrics and DNAmAgeAC varied somewhat by age (<60 or ≥60 y), but the differences were not statistically significant (all $p_{interaction} \geq 0.15$) and there were no consistent patterns by age among the DNAmAgeAC outcomes (Figure S9, Table S2). In general, associations between greenness and DNAmAgeAC were similar between those who never used alcohol and those who were current or former users, regardless of the DNAmAgeAC metric used (all $p_{interaction} \geq 0.18$) (Figure S10, Table S2).

*Z*-scores for five of the individual components used to derive the overall GrimAge metric (DNA methylation-based predictors of smoking pack years, serum cystatin-C, serum GDF-15, serum beta-2-microglobulin, and serum PAI-1) were inversely associated with all of the greenness metrics (Figure 4; Table S3). The associations were strongest for DNA methylation-based surrogates of smoking pack years (significant for all exposures), followed by serum cystatin-C and serum GDF-15. Of the remaining

**Table 1.** Basic characteristics of the 479 participants in Australia during 2004–2009.

| Characteristics | *n* (%) |
|---|---|
| Women | 479 (100.0) |
| Age group (y) | |
| <60 | 323 (67.4) |
| ≥60 | 156 (32.6) |
| Twin pairs | |
| Dizygotic twins | 132 (27.6) |
| Monozygotic twins | 132 (27.6) |
| Sisters of twins | 215 (44.9) |
| Education | |
| Secondary or below | 198 (41.3) |
| Vocational training | 135 (28.2) |
| University | 144 (30.1) |
| Missing | 2 (0.4) |
| Marital status | |
| Married or de facto | 369 (77.0) |
| Never married | 28 (5.8) |
| Widowed/separated/divorced | 82 (17.1) |
| Area-level SES percentile in Australia | |
| 0, 25 | 88 (18.4) |
| 25, 50 | 103 (21.5) |
| 50, 75 | 136 (28.4) |
| 75, 100 | 151 (31.5) |
| Missing | 1 (0.2) |
| Smoking | |
| Current | 41 (8.6) |
| Former | 147 (30.7) |
| Never | 291 (60.8) |
| Alcohol use | |
| Current | 235 (49.1) |
| Former | 52 (10.9) |
| Never | 192 (40.1) |

Note: SES, socioeconomic status.

components, associations with exposures were mixed (inverse, null, and positive) for methylation-based estimators of serum adrenomedullin, close to the null for serum TIMP-1, and null or positive for serum leptin.

Effect estimates were similar to the main model ($p_{\text{differences}} \geq 0.45$) in sensitivity analyses of models adjusted for season and survey year, for annual mean temperature and relative humidity as nonlinear terms, and for a DNA methylation-based smoking index instead of self-reported smoking, as well as for models that were not adjusted for self-reported smoking and alcohol use or for cell-type proportions (Figure S11, Table S4).

## Discussion

To the best of our knowledge, this is the first study to evaluate the association between surrounding greenness and measures of biological aging based on DNA methylation. Higher surrounding greenness was associated with slower biological aging, as measured by GrimAge acceleration. Of the eight components used to derive GrimAge, DNA methylation-based estimators of smoking pack years, serum GDF-15, and serum cystatin-C showed the strongest associations across both greenness metrics in all buffers when analyzed as separate outcomes. The association between higher surrounding greenness and slower biological aging measured by Horvath's, Hannum's, and PhenoAge DNAmAgeAC were close to the null in the population as a whole. Some significant positive associations were estimated for Horvath's and Hannum's age acceleration metrics among those with lower area-level SES, and some significant associations were estimated for Hannum's metric among those with a higher area-level SES, whereas no significant associations were estimated for PhenoAge acceleration.

Given the well-documented associations between higher surrounding greenness and a lower risk of many aging related conditions (e.g., cognitive decline, cardiovascular diseases, type 2 diabetes) (de Keijzer et al. 2018, 2020; Fong et al. 2018; James et al. 2015; Jia et al. 2018; Lachowycz and Jones 2011; Rojas-Rueda et al. 2019; Sarkar 2017), we hypothesized that high surrounding greenness would slow the rate of biological aging. Our findings for GrimAge support this hypothesis in our study population. This is generally consistent with a cross-sectional study

**Table 2.** Descriptive statistics of continuous variables the 479 participants in Australia during 2004–2009.

| | | | Percentiles | | | | |
|---|---|---|---|---|---|---|---|
| Variables | Mean ± SD | Min | 25th | 50th | 75th | Max | IQR |
| Chronological age (y) | 56.41 ± 7.90 | 39.69 | 50.16 | 55.66 | 61.91 | 77.53 | 11.75 |
| DNAmAge (y) | | | | | | | |
| Horvath's age | 55.54 ± 6.46 | 38.62 | 50.89 | 55.19 | 60.02 | 72.89 | 9.13 |
| Hannum's age | 57.28 ± 6.37 | 40.89 | 52.91 | 57.11 | 61.50 | 74.88 | 8.59 |
| PhenoAge | 53.04 ± 7.37 | 34.92 | 47.96 | 52.75 | 57.73 | 72.79 | 9.78 |
| GrimAge | 54.20 ± 6.46 | 39.64 | 49.51 | 53.86 | 58.54 | 72.76 | 9.04 |
| DNAmAgeAC (y) | | | | | | | |
| Horvath's age acceleration | 0.00 ± 4.88 | −16.33 | −3.23 | −0.15 | 3.02 | 15.40 | 6.26 |
| Hannum's age acceleration | 0.00 ± 4.41 | −14.98 | −2.75 | 0.21 | 2.66 | 18.42 | 5.40 |
| PhenoAge acceleration | 0.00 ± 5.78 | −15.45 | −4.20 | 0.31 | 3.96 | 21.25 | 8.17 |
| GrimAge acceleration | 0.00 ± 3.12 | −7.10 | −2.17 | −0.44 | 1.74 | 13.74 | 3.91 |
| Surrounding greenness | | | | | | | |
| NDVI (buffer) | | | | | | | |
| 300 m | 0.56 ± 0.12 | 0.25 | 0.47 | 0.55 | 0.64 | 0.87 | 0.175 |
| 500 m | 0.56 ± 0.12 | 0.26 | 0.47 | 0.55 | 0.64 | 0.87 | 0.165 |
| 1,000 m | 0.56 ± 0.12 | 0.26 | 0.48 | 0.55 | 0.64 | 0.87 | 0.164 |
| 2,000 m | 0.57 ± 0.12 | 0.27 | 0.48 | 0.56 | 0.66 | 0.85 | 0.178 |
| EVI (buffer) | | | | | | | |
| 300 m | 0.37 ± 0.10 | 0.15 | 0.30 | 0.36 | 0.42 | 0.76 | 0.122 |
| 500 m | 0.36 ± 0.10 | 0.14 | 0.30 | 0.35 | 0.42 | 0.77 | 0.121 |
| 1,000 m | 0.36 ± 0.10 | 0.14 | 0.30 | 0.35 | 0.42 | 0.74 | 0.119 |
| 2,000 m | 0.37 ± 0.10 | 0.13 | 0.29 | 0.35 | 0.43 | 0.70 | 0.134 |
| Annual mean temperature (°C) | 17.68 ± 2.81 | 11.35 | 15.26 | 17.76 | 20.26 | 25.07 | 5.00 |
| Annual mean relative humidity (%) | 69.81 ± 3.76 | 52.97 | 67.92 | 70.13 | 72.13 | 80.80 | 4.21 |

Notes: The chronological age was calculated as the difference between the date blood was taken and birth date. 2-y average, the average of the current year and previous year prior to blood being taken; DNAmAge, DNA methylation age; DNAmAgeAC, DNA methylation age acceleration; EVI, Enhanced Vegetation Index. IQR, interquartile range; max, maximum; min, minimum; NDVI, Normalized Difference Vegetation Index; SD, standard deviation.

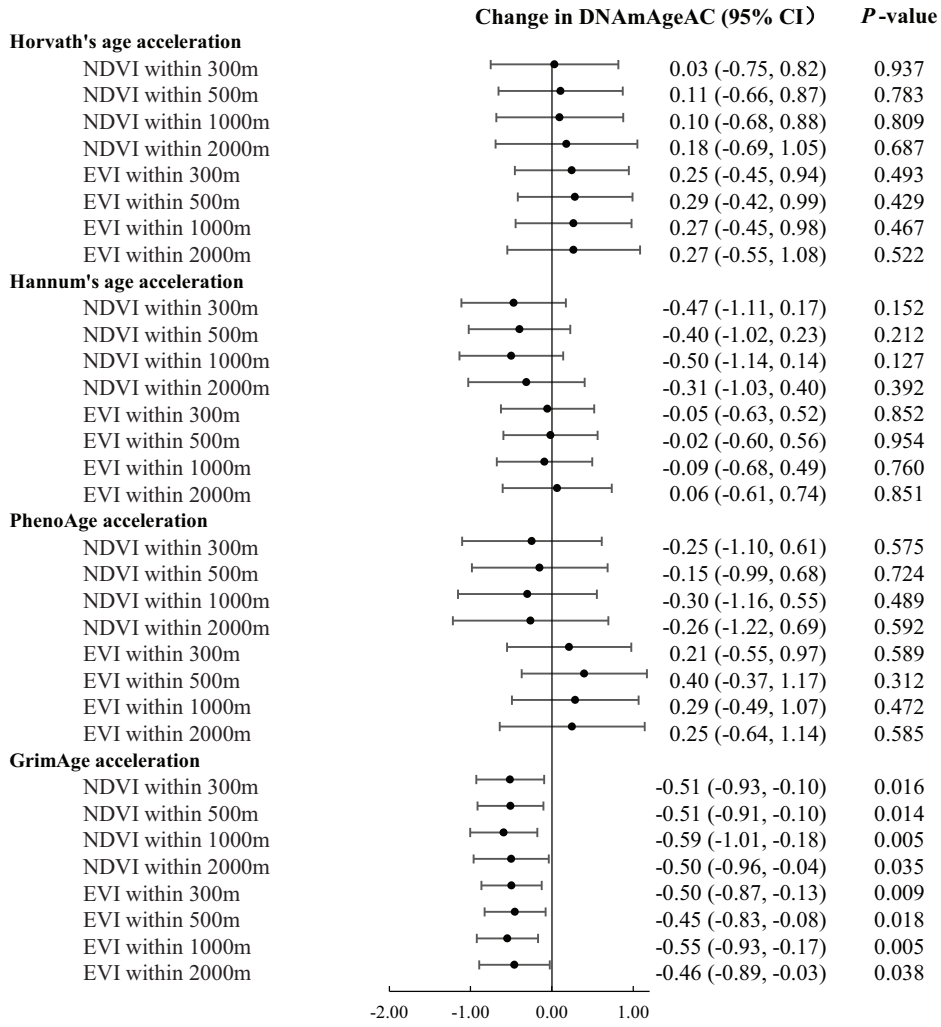


**Figure 2.** The association between every interquartile range increase in surrounding greenness and change in DNAmAge acceleration among 479 women in Australia during 2004–2009. Points and error bars represent point estimates and 95% CI, respectively. The effect estimates were from within-sibship analyses fitted by a linear mixed effect model adjusting for individual's chronological age, educational level, marital status and area-level socioeconomic status, smoking behavior, alcohol use, annual mean temperature, annual mean relative humidity, and proportions of seven type of blood cells (naïve $CD8^+$ T cells, $CD8^+$ T cells, plasma cells, $CD4^+$ T cells, NK cells, monocytes, and granulocytes). Missing values of covariates were addressed by multiple imputations. Note: CI, confidence interval; DNAmAgeAC, DNA methylation age acceleration; EVI, Enhanced Vegetation Index; NDVI, Normalized Difference Vegetation Index; NK, natural killer.

from Hong Kong, China, which reported that average telomere length was longer in residents of a 69-$km^2$ region with more green space than in residents of four highly populated regions (area size: 7–10 $km^2$) with little green space after adjusting for age, smoking, SES, and physical activity level (Woo et al. 2009). However, although telomere length has been interpreted as a marker of biological aging, DNAmAge is more strongly associated with mortality and morbidity (Jylhävä et al. 2017) and is thought to provide a more accurate measure of age acceleration, particularly in differentiated cells (Horvath and Raj 2018).

Based on meta-analyses of longitudinal data from four cohorts (6,395 adults followed for 13.7 y on average), Lu et al. (2019) estimated that each 1-y increase in GrimAge acceleration was associated with a 10% [hazard ratio (HR) = 1.10 (95% CI: 1.09, 1.12)] increase in all-cause mortality. Using this information and the estimated difference in GrimAge acceleration with an IQR increase in NDVI within 500 m in our study population (–0.51 y per 0.165-unit increase in NDVI), we estimate that a 0.1-unit increase in NDVI within 500 m would be associated with a 3% (i.e., $HR = 1.10^{-0.51 \times 0.1/0.165} = 0.97$) decrease in all-cause mortality. This estimate is consistent with a recent meta-analysis that reported a pooled HR of 0.96 (95% CI: 0.94, 0.97) for all-cause mortality associated with every 0.1-unit increase in NDVI within 500 m (Rojas-Rueda et al. 2019).

Although all measures of greenness exposure were associated with significantly slower biological aging based on the GrimAge DNAmAgeAC metric, associations with the other measures of biological age acceleration were variable (depending on the greenness measure or study subpopulation) or null. These apparent inconsistencies were not unexpected given recent evidence suggesting that GrimAge is a better indicator of biological aging. For example, among the four DNAmAge estimators, DNAmAgeAC estimated by GrimAge showed the strongest association (GrimAge > PhenoAge > Hannum > Horvath) with aging-related conditions (death, cancer, coronary heart disease, and age at menopause) based on an analysis of longitudinal data from four cohorts (6,395 participants followed for 13.7 y on average) (Lu et al. 2019). Similarly, a longitudinal analysis of data from two cohorts of older individuals ($n = 2,462$, mean follow-up of ~9 y) reported that, of the four DNAmAgeAC metrics assessed, GrimAge acceleration had the strongest and most consistent associations with mortality, myocardial infarction, and stroke (Wang et al. 2021).

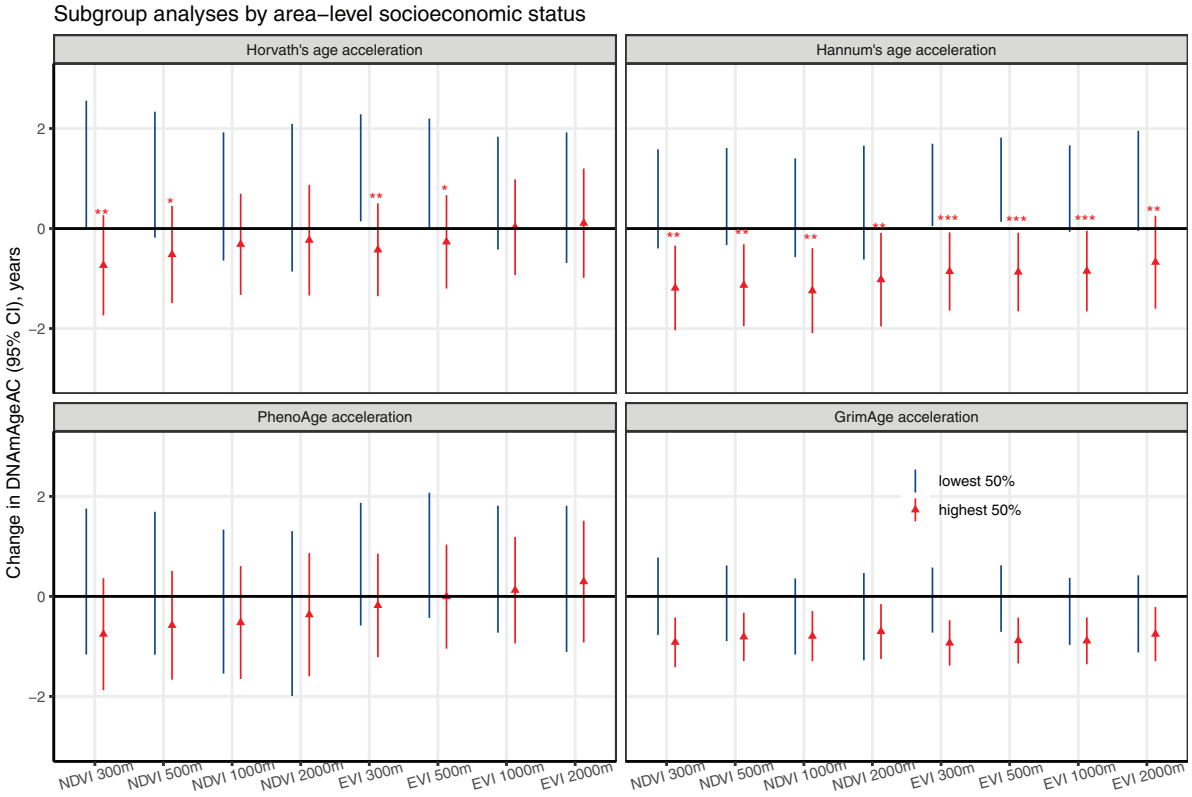


**Figure 3.** The association between every interquartile range increase in surrounding greenness and change in DNAmAge acceleration among 479 women in Australia during 2004–2009, subgroup analyses by area-level socioeconomic status (191 women at lowest 50% vs. 288 women at highest 50%). The effect estimates were from within-sibship analyses fitted by a linear mixed effect model adjusting for individual's chronological age, educational level, marital status, smoking behavior, alcohol use, and proportions of seven type of blood cells (naïve $CD8^+$ T cells, $CD8^+$ T cells, plasma cells, $CD4^+$ T cells, NK cells, monocytes, and granulocytes). Missing values of covariates were addressed by multiple imputations. Points and error bars represent point estimates and 95% CIs, respectively. Corresponding numeric data, including interaction *p*-values for all comparisons, are provided in Table S2. *, $p_{interaction} < 0.10$. **, $p_{interaction} < 0.05$. ***, $p_{interaction} < 0.01$. Note: CI, confidence interval; DNAmAgeAC, DNA methylation age acceleration; EVI, Enhanced Vegetation Index; NDVI, Normalized Difference Vegetation Index; NK, natural killer.

GrimAge acceleration also showed the strongest association with fatty liver and excess visceral fat in 1,177 adults from the Framingham Heart Study (Lu et al. 2019). A recent analysis of data from three cross-sectional surveys of UK adults 45–87 years of age ($n = 1{,}388$–$1{,}685$, depending on the outcome) also reported that GrimAge acceleration was the strongest predictor of physical and cognitive functioning, as assessed by grip strength, forced expiratory volume in 1 s ($FEV_1$), mental speed, and episodic memory (Maddock et al. 2020). For example, a 1-y increase in GrimAge acceleration was associated with a 0.03-mL (95% CI: –0.05, –0.01) lower mean $FEV_1$, whereas corresponding estimates for Horvath's, Hannum's, and PhenoAge DNAmAgeAC metrics were 0.00 (95% CI: –0.01, 0.01), 0.00 (95% CI: –0.02, 0.01), and –0.01 (95% CI: –0.02, –0.01), respectively (Maddock et al. 2020). GrimAge acceleration has also shown stronger associations with potential determinants of aging than other metrics, including smoking and alcohol use in a cross-sectional analysis of 1,100 adults in the United States (Zhao et al. 2019) and physical activity and neighborhood SES in cross-sectional analyses of 2,700 women in the United States (Kresovich et al. 2020; Lawrence et al. 2020). Taken together, these findings suggest that GrimAge DNAmAgeAC may be a better measure of biological aging than the other DNAmAgeAC metrics examined in our analysis.

Of the eight individual DNA methylation-based components of GrimAge, three had the strongest and most consistent associations with greenness when modeled as individual outcomes: *a*) DNA methylation-based markers of smoking pack years; *b*) GDF-15; and *c*) cystatin-C. DNA methylation-based smoking pack years represents a DNA methylation signature resulting from prolonged exposure to cigarette smoke (Lu et al. 2019), and this biomarker can be reversed following smoking cessation (McCartney et al. 2018). A plausible explanation for our results is that high greenness may help reverse the DNA methylation alterations resulting from exposure to cigarette smoke, given that high greenness is related to reduced risk of many adverse outcomes (e.g., cardiovascular diseases, cognitive decline, mental illness) that can be caused by smoking (Banks et al. 2019; Deal et al. 2020; Fong et al. 2018; Wootton et al. 2020).

GDF-15 is a promising biomarker for disease prognosis and an emerging target for cancer immunotherapy because of its important immune-regulatory function (Wischhusen et al. 2020). Serum levels of GDF-15 often increase with age, and highly elevated serum levels of GDF-15 are mostly linked to inflammation, myocardial ischemia, and cancer (Wischhusen et al. 2020). GDF-15 is also a potential target in the regulation of body weight and energy expenditure because high levels of GDF-15 contributes to appetite loss (Tsai et al. 2016). The DNA methylation-based GDF-15 was built to predict serum levels of GDF-15 (Lu et al. 2019). In our study, its inverse association with greenness may reflect the documented associations between high greenness and improved immune function (Kuo 2015) or improved metabolic health (e.g., lower adiposity) (Sarkar 2017).

High serum levels of cystatin-C is a widely used clinical indicator of reduced kidney function or declined glomerular filtration

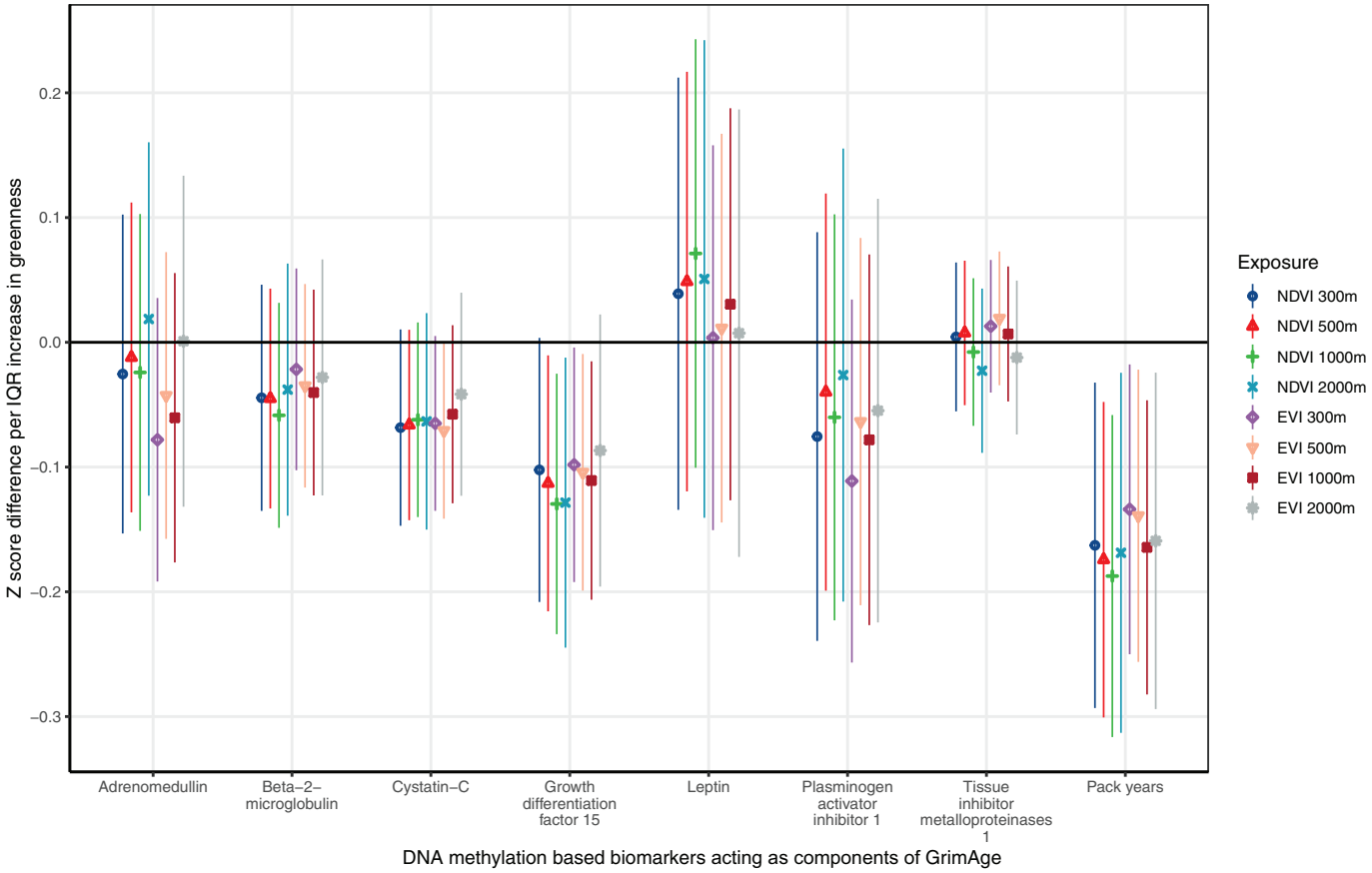


**Figure 4.** The association between every IQR increase in surrounding greenness and change in *Z*-score of eight DNA methylation-based biomarkers predicting GrimAge among 479 women in Australia during 2004–2009. The effect estimates were from within-sibship analyses fitted by a linear mixed effect model adjusting for each individual's chronological age, educational level, marital status, smoking behavior, alcohol use, and proportions of seven type of blood cells (naïve $CD8^+$ T cells, $CD8^+$ T cells, plasma cells, $CD4^+$ T cells, NK cells, monocytes, and granulocytes). Missing values of covariates were addressed by multiple imputations. *Z*-score was calculated as (actual value – sample mean value)/sample standard deviation. Points and error bars represent point estimates and 95% CIs, respectively. Corresponding numeric data, including interaction *p*-values for all comparisons, are provided in Table S3. Note: CI, confidence interval; DNAmAgeAC, DNA methylation age acceleration; EVI, Enhanced Vegetation Index; IQR, interquartile range; NDVI, Normalized Difference Vegetation Index; NK, natural killer.

rate (Di Somma and Marino 2019). In our study, the DNA methylation-based cystatin-C, a predictor of serum level of cystatin-C (Lu et al. 2019), was inversely associated with greenness, although the association was only statistically significant for EVI within 300 m. This may suggest that high greenness may be associated with improved kidney function. However, current evidence about the relationship between greenness and kidney diseases remains scarce (Kondo et al. 2018). A prospective cohort study in 108,630 women from the U.S.-based Nurses' Health Study followed up between 2000 and 2008 found a relationship between high greenness and reduce mortality risk from kidney diseases, but the association was not statistically significant [HR = 0.63 (95% CI: 0.38, 1.04) for each 0.1-unit increase in NDVI within 250 m], possibly due to the small number of cases (139 cases) (James et al. 2016). The other potential explanation for our results for cystatin-C might be that high greenness is associated with lower adiposity (Sarkar 2017). Adipose tissue was found to be among the tissues expressing the highest level of cystatin-C mRNA, which contribute to the higher serum levels cystatin-C seen in obese people compared with nonobese people (Naour et al. 2009).

Patterns of associations between greenness and age acceleration according to neighborhood SES varied among the DNAmAgeAC metrics, with some significant differences for Horvath's and Hannum's metrics resulting from positive associations among those with lower neighborhood SES and inverse associations among those with a higher neighborhood SES. There were no significant differences by neighborhood SES for the GrimAge metric, although associations were consistently inverse for those with higher neighborhood SES and closer to the null for those with lower SES. Associations between greenness and Hannum's and GrimAge DNAmAgeAC were similar between those with higher and lower education, a proxy measure of individual-level SES, whereas associations with Horvath's DNAmAgeAC tended to be positive for those with lower education and null for those with higher education (with a significant difference for only one exposure metric.) In a Swiss cohort, the association between surrounding greenness and lower mortality was stronger among participants living in communities with a higher area-level SES (Vienneau et al. 2017), whereas in residents of Rome, Italy, residential greenness was associated lower mortality in the population as a whole but with higher mortality among residents in the lowest quintile of area-level SES (Orioli et al. 2019). Another study reported that a positive association between neighborhood socioeconomic disadvantage and a DNA methylation-based mortality risk index was attenuated in areas with high neighborhood greenness (Ward-Caviness et al. 2020). Differences in associations between greenness and biological age acceleration or other health outcomes could be related to differences in greenspace quality or accessibility between communities with higher SES or

lower SES (Vienneau et al. 2017). Higher neighborhood SES might also be associated with a stronger social cohesion (Feldman and Steptoe 2004) that could boost the health benefits of green space (Jennings and Bamkole 2019).

The twin and family design is a strength of our study. This design compares each participant with her siblings, thus reducing potential confounding by genetic background and shared familial factors (e.g., childhood living environment, familial dietary behavior) (Li et al. 2017; Stone et al. 2010). This could be an important advantage given that familial factors explain a substantial proportion of variation in DNAmAge (Horvath and Raj 2018; Li et al. 2020; Lu et al. 2019).

Several limitations also should be acknowledged. This is a cross-sectional analysis, and further longitudinal studies are warranted to verify our findings. Our sample size may have limited statistical power to estimate precise and stable greenness–DNAmAgeAC associations and potential effect modifiers. In addition, the study included only women. Previous studies reported that associations between surrounding greenness and decreased mortality and body fat were weaker in men than women (Ji et al. 2019; Sarkar 2017). Associations between greenness and DNAmAgeAC also might be weaker in men.

In conclusion, higher surrounding greenness was associated with slower biological aging measured by GrimAge acceleration in Australian women. The association was also evident for three individual components of GrimAge and seemed to be stronger among those with higher vs. lower neighborhood SES, although the differences were not statistically significant. Associations between greenness and biological aging measured by Horvath's and Hannum's DNAmAgeAC were less consistent, and varied depending on neighborhood SES. Our study and previous epidemiological evidence suggests that GrimAge may be a more reliable measure of biological aging than other DNAmAge metrics, but additional studies are needed to confirm our findings in other populations.

## Acknowledgments

R.X. is supported by China Scholarship Council (201806010405). Shuai Li is supported by an Early Career Fellowship of the Australian National Health and Medical Research Council (NHMRC; APP1109193). Y.G. is supported by a Career Development Fellowship of the Australian National Health and Medical Research Council (APP1163693). Shuai Li is supported by an Early Career Research Fellowship of the Victorian Cancer Agency (ECRF 19020). M.C.S. is a NHMRC Senior Research Fellow (APP1155163). J.L.H. is a NHMRC Senior Principal Research Fellow. The Australian Mammographic Density Twins and Sisters Study (AMDTSS) was facilitated through access to Twins Research Australia, a national resource supported by a Centre of Research Excellence Grant (1079102) from the NHMRC. The AMDTSS was supported by NHMRC (1050561 and 1079102), Cancer Australia and National Breast Cancer Foundation (509307). The raw and processed DNA methylation data set are open accessed or free on Gene Expression Omnibus (accession number GSE100227). As required by the ethics approval, the authors are not allowed to open other data (e.g., data on covariates) used in this study in order to protect the privacy of participants. If anyone wants to use the data to repeat our analyses or to perform other research, please contact J.L.H. (j.hopper@unimelb.edu.au), who is happy to engage with external collaborators. This person would have to be added to the ethics approval to access the data. Working with the AMDTSS group as collaborators would likely result in better science given this is family data and the nuances of this, as well as the sampling issues, need to be understood to guard against making false conclusions from a naïve use of data without such knowledge.

## References

ABS (Australian Bureau of Statistics). 2006. Information paper: an introduction to Socio-Economic Indexes for Areas (SEIFA), 2006. Australian Capital Territory, Australia: ABS. https://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/2039.0Main%20Features42006 [accessed 19 August 2021].

ABS. 2010. Australian Social Trends, Dec 2010. https://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/4102.0Main+Features1Dec%202010?OpenDocument [accessed 15 February 2021].

ABS. 2011. Australian Statistical Geography Standard (ASGS): Volume 4 - Significant Urban Areas, Urban Centres and Localities, Section of State, July 2011. https://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/1270.0.55.004Main+Features1July%202011?OpenDocument [accessed 10 January 2020].

ABS. 2020. Frequently asked questions. How does the ABS define urban and rural? https://www.abs.gov.au/websitedbs/D3310114.nsf/home/Frequently+Asked+Questions#Anchor7 [accessed 19 August 2021].

Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. 2014. Minfi: a flexible and comprehensive bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics 30(10):1363–1369, PMID: 24478339, https://doi.org/10.1093/bioinformatics/btu049.

Banks E, Joshy G, Korda RJ, Stavreski B, Soga K, Egger S, et al. 2019. Tobacco smoking and risk of 36 cardiovascular disease subtypes: fatal and non-fatal outcomes in a large prospective Australian study. BMC Med 17(1):128, PMID: 31266500, https://doi.org/10.1186/s12916-019-1351-4.

Boyd NF, Dite GS, Stone J, Gunasekara A, English DR, McCredie MRE, et al. 2002. Heritability of mammographic density, a risk factor for breast cancer. N Engl J Med 347(12):886–894, PMID: 12239257, https://doi.org/10.1056/NEJMoa013390.

Browning M, Lee K. 2017. Within what distance does "greenness" best predict physical health? A systematic review of articles with GIS buffer analyses across the lifespan. Int J Environ Res Public Health 14(7):675, PMID: 28644420, https://doi.org/10.3390/ijerph14070675.

de Keijzer C, Bauwelinck M, Dadvand P. 2020. Long-term exposure to residential greenspace and healthy ageing: a systematic review. Curr Environ Health Rep 7(1):65–88, PMID: 31981136, https://doi.org/10.1007/s40572-020-00264-7.

de Keijzer C, Tonne C, Basagaña X, Valentín A, Singh-Manoux A, Alonso J, et al. 2018. Residential surrounding greenness and cognitive decline: a 10-year follow-up of the Whitehall II cohort. Environ Health Perspect 126(7):077003, PMID: 30028296, https://doi.org/10.1289/EHP2875.

Deal JA, Power MC, Palta P, Alonso A, Schneider ALC, Perryman K, et al. 2020. Relationship of cigarette smoking and time of quitting with incident dementia and cognitive decline. J Am Geriatr Soc 68(2):337–345, PMID: 31675113, https://doi.org/10.1111/jgs.16228.

Di Lena P, Sala C, Prodi A, Nardini C. 2019. Missing value estimation methods for DNA methylation data. Bioinformatics 35(19):3786–3793, PMID: 30796811, https://doi.org/10.1093/bioinformatics/btz134.

Di Somma S, Marino R. 2019. Diagnosis and management of acute kidney injury in the emergency department. In: *Critical Care Nephrology*. Ronco C, Bellomo R, Kellum JA, Ricci Z, eds. 3rd ed. Philadelphia, PA: Elsevier, 1296–1301.

Didan K. 2015. MOD13Q1—MODIS/Terra Vegetation Indices 16-Day L3 Global 250m Sin Grid v006. Sioux Falls, SD: NASA EOSDIS Land Processes DAAC. https://lpdaac.usgs.gov/products/mod13q1v006/ [accessed 19 August 2021].

Feldman PJ, Steptoe A. 2004. How neighborhoods and physical functioning are related: the roles of neighborhood socioeconomic status, perceived neighborhood strain, and individual health risk factors. Ann Behav Med 27(2):91–99, PMID: 15026293, https://doi.org/10.1207/s15324796abm2702_3.

Ferrucci L, Gonzalez-Freire M, Fabbri E, Simonsick E, Tanaka T, Moore Z, et al. 2020. Measuring biological aging in humans: a quest. Aging Cell 19(2):e13080, PMID: 31833194, https://doi.org/10.1111/acel.13080.

Fong KC, Hart JE, James P. 2018. A review of epidemiologic studies on greenness and health: updated literature through 2017. Curr Environ Health Rep 5(1):77–87, PMID: 29392643, https://doi.org/10.1007/s40572-018-0179-y.

Gao X, Zhang Y, Saum K-U, Schöttker B, Breitling LP, Brenner H. 2017. Tobacco smoking and smoking-related DNA methylation are associated with the development of frailty among older adults. Epigenetics 12(2):149–156, PMID: 28001461, https://doi.org/10.1080/15592294.2016.1271855.

Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, et al. 2013. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol Cell 49(2):359–367, PMID: 23177740, https://doi.org/10.1016/j.molcel.2012.10.016.

Hastie T, Tibshirani R, Narasimhan B, Chu G. 2018. impute: imputation for microarray data. [R package], https://git.bioconductor.org/packages/impute [accessed 25 August 2021].

Horvath S. 2013. DNA methylation age of human tissues and cell types. Genome Biol 14(10):R115, PMID: 24138928, https://doi.org/10.1186/gb-2013-14-10-r115.

Horvath S, Raj K. 2018. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. Nat Rev Genet 19(6):371–384, PMID: 29643443, https://doi.org/10.1038/s41576-018-0004-3.

Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. 2012. DNA methylation arrays as surrogate measures of cell mixture

distribution. BMC Bioinformatics 13:86, PMID: 22568884, https://doi.org/10.1186/1471-2105-13-86.
James P, Banay RF, Hart JE, Laden F. 2015. A review of the health benefits of greenness. Curr Epidemiol Rep 2(2):131–142, PMID: 26185745, https://doi.org/10.1007/s40471-015-0043-7.
James P, Hart JE, Banay RF, Laden F. 2016. Exposure to greenness and mortality in a nationwide prospective cohort study of women. Environ Health Perspect 124(9):1344–1352, PMID: 27074702, https://doi.org/10.1289/ehp.1510363.
Jeffrey SJ, Carter JO, Moodie KB, Beswick AR. 2001. Using spatial interpolation to construct a comprehensive archive of Australian climate data. Environ Model Softw 16(4):309–330, https://doi.org/10.1016/S1364-8152(01)00008-1.
Jennings V, Bamkole O. 2019. The relationship between social cohesion and urban green space: an avenue for health promotion. Int J Environ Res Public Health 16(3):452, PMID: 30720732, https://doi.org/10.3390/ijerph16030452.
Ji JS, Zhu A, Bai C, Wu C-D, Yan L, Tang S, et al. 2019. Residential greenness and mortality in oldest-old women and men in China: a longitudinal cohort study. Lancet Planet Health 3(1):e17–e25, PMID: 30654864, https://doi.org/10.1016/S2542-5196(18)30264-X.
Jia X, Yu Y, Xia W, Masri S, Sami M, Hu Z, et al. 2018. Cardiovascular diseases in middle aged and older adults in China: the joint effects and mediation of different types of physical exercise and neighborhood greenness and walkability. Environ Res 167:175–183, PMID: 30029039, https://doi.org/10.1016/j.envres.2018.07.003.
Johnson WE, Li C, Rabinovic A. 2007. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8(1):118–127, PMID: 16632515, https://doi.org/10.1093/biostatistics/kxj037.
Jylhävä J, Pedersen NL, Hägg S. 2017. Biological age predictors. EBioMedicine 21:29–36, PMID: 28396265, https://doi.org/10.1016/j.ebiom.2017.03.046.
Kondo MC, Fluehr JM, McKeon T, Branas CC. 2018. Urban green space and its impact on human health. Int J Environ Res Public Health 15(3):445, PMID: 29510520, https://doi.org/10.3390/ijerph15030445.
Kresovich JK, Garval EL, Martinez Lopez AM, Xu Z, Niehoff NM, White AJ, et al. 2020. Associations of body composition and physical activity level with multiple measures of epigenetic age acceleration. Am J Epidemiol 190(6):984–993, PMID: 33693587, https://doi.org/10.1093/aje/kwaa251.
Kuo M. 2015. How might contact with nature promote human health? Promising mechanisms and a possible central pathway. Front Psychol 6:1093, PMID: 26379564, https://doi.org/10.3389/fpsyg.2015.01093.
Lachowycz K, Jones AP. 2011. Greenspace and obesity: a systematic review of the evidence. Obes Rev 12(5):e183–e189, PMID: 21348919, https://doi.org/10.1111/j.1467-789X.2010.00827.x.
Lawrence KG, Kresovich JK, O'Brien KM, Hoang TT, Xu Z, Taylor JA, et al. 2020. Association of neighborhood deprivation with epigenetic aging using 4 clock metrics. JAMA Netw Open 3(11):e2024329, PMID: 33146735, https://doi.org/10.1001/jamanetworkopen.2020.24329.
Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, et al. 2018. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY) 10(4):573–591, PMID: 29676998, https://doi.org/10.18632/aging.101414.
Li S, Nguyen TL, Wong EM, Dugue P-A, Dite GS, Armstrong NJ, et al. 2020. Genetic and environmental causes of variation in epigenetic aging across the lifespan. Clin Epigenetics 12(1):158, PMID: 33092643, https://doi.org/10.1186/s13148-020-00950-1.
Li S, Wong EM, Joo JE, Jung C-H, Chung J, Apicella C, et al. 2015. Genetic and environmental causes of variation in the difference between biological age based on DNA methylation and chronological age for middle-aged women. Twin Res Hum Genet 18(6):720–726, PMID: 26527295, https://doi.org/10.1017/thg.2015.75.
Li S, Wong EM, Southey MC, Hopper JL. 2017. Association between DNA methylation at *SOCS3* gene and body mass index might be due to familial confounding. Int J Obes (Lond) 41(6):995–996, PMID: 28239163, https://doi.org/10.1038/ijo.2017.56.
Lu AT, Quach A, Wilson JG, Reiner AP, Aviv A, Raj K, et al. 2019. DNA methylation GrimAge strongly predicts lifespan and healthspan. Aging (Albany NY) 11(2):303–327, PMID: 30669119, https://doi.org/10.18632/aging.101684.
Maddock J, Castillo-Fernandez J, Wong A, Cooper R, Richards M, Ong KK, et al. 2020. DNA methylation age and physical and cognitive aging. J Gerontol A Biol Sci Med Sci 75(3):504–511, PMID: 31630156, https://doi.org/10.1093/gerona/glz246.
Maksimovic J, Gordon L, Oshlack A. 2012. SWAN: subset-quantile within array normalization for Illumina Infinium HumanMethylation450 BeadChips. Genome Biol 13(6):R44, PMID: 22703947, https://doi.org/10.1186/gb-2012-13-6-r44.
Markevych I, Schoierer J, Hartig T, Chudnovsky A, Hystad P, Dzhambov AM, et al. 2017. Exploring pathways linking greenspace to health: theoretical and methodological guidance. Environ Res 158:301–317, PMID: 28672128, https://doi.org/10.1016/j.envres.2017.06.028.
McCartney DL, Stevenson AJ, Hillary RF, Walker RM, Bermingham ML, Morris SW, et al. 2018. Epigenetic signatures of starting and stopping smoking. EBioMedicine 37:214–220, PMID: 30389506, https://doi.org/10.1016/j.ebiom.2018.10.051.
Naour N, Fellahi S, Renucci JF, Poitou C, Rouault C, Basdevant A, et al. 2009. Potential contribution of adipose tissue to elevated serum cystatin C in human obesity. Obesity (Silver Spring) 17(12):2121–2126, PMID: 19360013, https://doi.org/10.1038/oby.2009.96.
Orioli R, Antonucci C, Scortichini M, Cerza F, Marando F, Ancona C, et al. 2019. Exposure to residential greenness as a predictor of cause-specific mortality and stroke incidence in the Rome Longitudinal Study. Environ Health Perspect 127(2):27002, PMID: 30775931, https://doi.org/10.1289/EHP2854.
Reid CE, Kubzansky LD, Li J, Shmool JL, Clougherty JE. 2018. It's not easy assessing greenness: a comparison of NDVI datasets and neighborhood types and their associations with self-rated health in New York City. Health Place 54:92–101, PMID: 30248597, https://doi.org/10.1016/j.healthplace.2018.09.005.
Rojas-Rueda D, Nieuwenhuijsen MJ, Gascon M, Perez-Leon D, Mudu P. 2019. Green spaces and mortality: a systematic review and meta-analysis of cohort studies. Lancet Planet Health 3(11):e469–e477, PMID: 31777338, https://doi.org/10.1016/S2542-5196(19)30215-3.
Sarkar C. 2017. Residential greenness and adiposity: findings from the UK Biobank. Environ Int 106:1–10, PMID: 28551493, https://doi.org/10.1016/j.envint.2017.05.016.
Sbihi H, Jones MJ, Sears MR, Subbarao P, Moraes TJ, Mandhane P, et al. 2018. Epigenetics, built environment and atopy. [Abstract.] In: *Proceedings of the ISEE Conference Abstracts*. Vol. 2018.
Stone J, Gurrin LC, Byrnes GB, Schroen CJ, Treloar SA, Padilla EJD, et al. 2007. Mammographic density and candidate gene variants: a twins and sisters study. Cancer Epidemiol Biomarkers Prev 16(7):1479–1484, PMID: 17627014, https://doi.org/10.1158/1055-9965.EPI-07-0107.
Stone J, Gurrin LC, Hayes VM, Southey MC, Hopper JL, Byrnes GB. 2010. Sibship analysis of associations between SNP haplotypes and a continuous trait with application to mammographic density. Genet Epidemiol 34(4):309–318, PMID: 19918759, https://doi.org/10.1002/gepi.20462.
Taylor L, Hochuli DF. 2017. Defining greenspace: multiple uses across multiple disciplines. Landsc Urban Plan 158:25–38, https://doi.org/10.1016/j.landurbplan.2016.09.024.
Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshirani R, et al. 2001. Missing value estimation methods for DNA microarrays. Bioinformatics 17(6):520–525, PMID: 11395428, https://doi.org/10.1093/bioinformatics/17.6.520.
Tsai VWW, Lin S, Brown DA, Salis A, Breit SN. 2016. Anorexia–cachexia and obesity treatment may be two sides of the same coin: role of the TGF-b superfamily cytokine MIC-1/GDF15. Int J Obes (Lond) 40(2):193–197, PMID: 26620888, https://doi.org/10.1038/ijo.2015.242.
United Nations. 2017. *World Population Ageing: Highlights*. New York, NY: United Nations. https://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2017_Highlights.pdf [accessed 19 August 2021].
van Buuren S, Groothuis-Oudshoorn K. 2011. mice: multivariate imputation by chained equations in R. J Stat Softw 45(3):1–67. https://www.jstatsoft.org/index.php/jss/article/view/v045i03/v45i03.pdf [accessed 19 August 2021].
Vienneau D, de Hoogh K, Faeh D, Kaufmann M, Wunderli JM, Röösli M, et al. 2017. More than clean air and tranquillity: residential green is independently associated with decreasing mortality. Environ Int 108:176–184, PMID: 28863390, https://doi.org/10.1016/j.envint.2017.08.012.
Wang C, Ni W, Yao Y, Just A, Heiss J, Wei Y, et al. 2021. DNA methylation-based biomarkers of age acceleration and all-cause death, myocardial infarction, stroke, and cancer in two cohorts: the NAS, and KORA F4. EBioMedicine 63:103151, PMID: 33279859, https://doi.org/10.1016/j.ebiom.2020.103151.
Ward-Caviness CK, Pu S, Martin CL, Galea S, Uddin M, Wildman DE, et al. 2020. Epigenetic predictors of all-cause mortality are associated with objective measures of neighborhood disadvantage in an urban population. Clin Epigenetics 12(1):44, PMID: 32160902, https://doi.org/10.1186/s13148-020-00830-8.
Wischhusen J, Melero I, Fridman WH. 2020. Growth/differentiation factor-15 (GDF-15): from biomarker to novel targetable immune checkpoint. Front Immunol 11:951, PMID: 32508832, https://doi.org/10.3389/fimmu.2020.00951.
Woo J, Tang N, Suen E, Leung J, Wong M. 2009. Green space, psychological restoration, and telomere length. Lancet 373(9660):299–300, PMID: 19167568, https://doi.org/10.1016/S0140-6736(09)60094-5.
Wootton RE, Richmond RC, Stuijfzand BG, Lawn RB, Sallis HM, Taylor GMJ, et al. 2020. Evidence for causal effects of lifetime smoking on risk for depression and schizophrenia: a Mendelian randomisation study. Psychol Med 50(14):2435–2443, PMID: 31689377, https://doi.org/10.1017/S0033291719002678.
WHO (World Health Organization). 2015. *World Report on Ageing and Health*. Geneva, Switzerland: WHO.
Xu R, Li S, Guo S, Zhao Q, Abramson MJ, Li S, et al. 2020. Environmental temperature and human epigenetic modifications: a systematic review. Environ Pollut 259:113840, PMID: 31884209, https://doi.org/10.1016/j.envpol.2019.113840.
Xu R, Zhao Q, Coelho MSZS, Saldiva PHN, Abramson MJ, Li S, et al. 2019. The association between heat exposure and hospitalization for undernutrition in Brazil during 2000–2015: a nationwide case-crossover study. PLoS Med 16(10):e1002950, PMID: 31661490, https://doi.org/10.1371/journal.pmed.1002950.
Zhao W, Ammous F, Ratliff S, Liu J, Yu M, Mosley TH, et al. 2019. Education and lifestyle factors are associated with DNA methylation clocks in older African Americans. Int J Environ Res Public Health 16(17):3141, PMID: 31466396, https://doi.org/10.3390/ijerph16173141.