Data Supplement for Han et al., Epigenetic Aging in Major Depressive Disorder. Am J Psychiatry (doi: 10.1176/appi.ajp.2018.17060595)

CONTENTS

SUPPLEMENTARY METHODS

Assessed phenotypes

Body Mass Index (BMI) was calculated as weight (kilograms) divided by squared length (meters). Smoking behavior was represented by cotinine levels, an adequate marker for calculating recent tobacco exposure and discriminating smokers from non-smokers (1). Physical activity was assessed using the International Physical Activity Questionnaire and indicated by overall energy expenditure in Metabolic Equivalent Total-minutes per week. Health status was assessed as number of chronic diseases (heart disease, epilepsy, diabetes, osteoarthritis, cancer, stroke, intestinal disorders, ulcers, and lung-, liver-, and thyroid disease) for which participants received medical treatment. In line with earlier work, a cumulative childhood trauma index (CTI) was created that reported the sum of the categories that were scored from 0 to 2 (0: never happened, 1: sometimes, 2: happened regularly), resulting in an index score ranging from 0-8(2). Finally, frequent use of antidepressants was assessed through container inspection and categorized using World Health Organization Anatomical Therapeutic Chemical (ATC) classifications: antidepressants (tricyclic antidepressants (ATC code N06AA), selective serotonin reuptake inhibitors (ATC code N06AB), and other antidepressants (ATC codes N06AF, N06AG, N06AX). Childhood trauma was assessed using the NEMESIS childhood trauma interview with personal history questions including a structured inventory of trauma exposure during childhood (emotional neglect, psychological abuse, physical abuse, sexual abuse, and important life-events in early life).

Quality control of MBD-seq data in NESDA

Of the 1200 samples from participants eligible for methylome-wide sequencing, 34 samples were excluded because of failed methylation enrichment (n=16) or failed library construction (n=18) where we did not have sufficient DNA available for a repeated assay. Reads aligning to loci without CpGs

(non-CpGs) represent "noise" caused by, for example, imperfect enrichment leading to non-methylated fragments being sequenced. We used a threshold of 0.05 for "non-CpG/CpG coverage ratio" to remove samples with high "noise" levels (n=10), leaving an average ratio of 0.010 (SD=0.005) in the remaining samples. For 10 samples, sequence variants called from the methylation data did not match the genotype information, indicating that a sample swap or sample contamination may have occurred. As it is not possible to determine whether the sample handling errors occurred in the GWAS or in the MWAS data, we conservatively excluded all 10 samples from further analysis. We used the R function 'pcout' in the 'mvoutliers' package (with the upper boundary for outlier detection set to 15, the scaling constant set to 0.5, and the boundary for final outliers set to 0.2) to identify multidimensional outliers using principal components of the methylation data as input. Fourteen samples were multidimensional outliers and omitted. Finally, two samples were removed after DNAm age residual estimation because there were determined to be extreme outliers as they were both more than seven standard deviations away from the mean.

This left a sample of 1130 subjects. The mean number of reads for these samples was 59,954,723. The average alignment rate was 99.2%. We performed quality control (QC) for multi- and duplicate-reads. Although reads often map to multiple genomic locations, in most cases, a single alignment can be selected because it is clearly better than other alignments. In the case of multi-reads, multiple alignments are about equally good (in terms of alignment score). When Bowtie2 encounters a set of equally good alignments, it uses a pseudo-random number to select one primary alignment. Duplicate-reads are reads that start at the same nucleotide positions. When sequencing a whole genome, duplicate-reads often arise from artifacts in template preparation or amplification. However, in the context of sequencing an enriched genomic fraction such as methyl-CG binding domain sequencing (MBD-seq), duplicate-reads are increasingly likely to occur because reads originate from a smaller fraction of the genome. We therefore allow for three reads to occur at the same location but

for instances where more than 3 (duplicate) reads start at the same position, we reset the read count to 1 implicitly assuming excess reads are tagging a single clonal fragment. This left an average of 48,653,227 reads per sample (=81.9% of all reads).

To identify CpGs, we combined reference genome sequence (hg19/GRCh37) with common SNPs calculated on the European superpopulation from 1000 Genomes (Phase 3). To avoid including sites that are CpGs in only a very small proportion of subjects, we excluded CpGs created by SNPs with minor allele frequency <1%. This resulted in 27,916,990 CpGs. CpGs in loci prone to alignment errors, e.g., in repetitive regions, were eliminated prior to the analysis. To identify these CpGs, we used RaMWAS to perform the in-silico alignment experiment outlined elsewhere.(3) In this experiment, the vast majority of CpGs (89.3%) were located in regions that showed perfect alignment coverage and only 1.3% (365,223 CpGs) showed evidence of alignment problems defined as 15% or more reads from this locus not aligning properly. These CpGs with alignment problems were removed from further analyses. Finally, we eliminated 5,682,206 CpGs with average coverage less than 0.3 or having zero coverage in over 70% of the samples. This resembles filtering GWAS SNPs on minor allele frequency and avoids statistical problems associated with analyzing sparse data that is the result of sites that are not methylated in almost any subject in the study. This left a total of 21,869,561 CpGs.

Cell type measurements

To estimate cell type proportions, we used reference methylomes(4,5). Whole blood samples of six subjects were used to isolated cells with 5 clusters of differentiation (CD3, CD19, CD20, CD14, and CD15) that capture the most common cell types in blood (T-cells, B-cells, monocytes, and granulocytes). Cell populations were isolated by positive selection using EasySep™ kits (Stemcell technologies) that apply magnetic nanoparticles coated with antibodies against a particular surface antigen (CD molecules). All reference methylomes were generated using MBD-seq. In a previous

paper we showed that the estimated cell type proportions effectively controlled for cell type heterogeneity in methylome-wide association studies(6).

We further validated our cell type proportion estimates by correlating them with automated counts from 337 subjects obtained with the Abbott Sapphire system that uses optical scatter and impedance. We only had automated cell counts for these subjects 2-3 and 5-6 years after the methylation measurement and the cross-year correlation was 0.589. This correlation is attenuated by the fact that cell counts will have changed over the 2 to 3-year period but serves as a comparison. The correlation between MBD-seq estimates at baseline and automated counts 2-3 years later was 0.534 and only slightly lower. This suggests that the reliability of MBD-seq estimates of cell type proportions was comparable to that of automated counts as typically used in clinical settings.

Post-mortem brain samples for replication

Here, we summarize the post-mortem brain samples used in the replication analysis. Diagnosing disease in subjects providing post-mortem brain samples can be challenging(7). In most cases one or two psychiatrists determine the diagnosis by using information obtained from a family member who is well acquainted with the deceased. This technique has been validated for axis I and II diagnoses(8,9) and has shown to have high inter-rater agreement(10). We pooled the data of six brain collections from four different brain banks that are described in detail below.

The first subsample included post-mortem brain tissue from 30 MDD cases and 30 matched controls, obtained from the Victorian Brain Bank Network, Australia(11). For MDD cases, DSM-IV diagnoses were confirmed post-mortem by two psychiatrists, using clinical case histories and the Diagnostic Instrument for Brain Studies (DIBS)(12). The controls had no history of psychiatric symptoms or substance abuse (as determined by both information from relatives and medical

records) and were age/sex matched to the cases. The tissue samples were dissected from the cerebral cortex (BA25) for each subject.

The second subsample included 3 cases and 4 controls from the Harvard Brain Bank(13). Family members initially reported diagnoses at the time of death and next of kin were asked to complete a questionnaire/participate in a phone interview to provide further details. A staff psychiatrist then reviewed the clinical records and family questionnaires to confirm or correct the psychiatric diagnosis.

The third subsample included 9 cases and 9 controls from the Netherlands Brain Bank(14). Reports by family members of a lifetime diagnosis of MDD was confirmed post-mortem by a certified psychiatrist on the basis of the medical records following DSM-IV criteria. Controls never received any psychiatric diagnosis or long-term psychotropic medication.

The fourth and fifth independent replication samples included BA10 samples from the Stanley Medical Research Institute (SMRI)(15). The first consisted of 22 MDD cases (with or without psychosis) and 11 controls. The second collection comprised 10 (non-psychotic) cases and 13 controls. The SMRI uses DSM-IV diagnoses made by two senior psychiatrists on the basis of medical records and, when necessary, telephone interviews with family members. Diagnoses of unaffected controls are based on structured interviews by a senior psychiatrist with family member(s) to rule out Axis I diagnoses. In addition to the balanced MDD case-control collections, we also used methylome data from the same brain bank and brain region BA10 from an additional 78 individuals (25 schizophrenia cases, 18 bipolar cases and 25 controls) that were solely used to train the DNAmAge prediction model, but were excluded from all replication analyses.

MBD-seq data from post-mortem brain samples

To study the CpG methylome of the post-mortem brain samples we again used the MBD-seq approach. All brain samples were assayed using the same MBD-seq protocol(16). In short, we used ultrasonicaton to shear genomic DNA into an average of 150 bp fragments. Next, we performed enrichment with MethylMiner™ (Invitrogen), following the same procedure as was described for the blood samples, to capture the methylated fraction of the genome. Barcoded sequencing libraries were manually created for each methylation capture, were pooled in equal molarities, sequenced on a SOLiD5500 wildfire instrument (Life technologies) with 50bp reads and were aligned with Cushaw3(17). As for the MBD-seq data from blood, the aligned MBD-seq data from brain was processed and analyzed using RaMWAS (RaMWAS: Fast Methylome-Wide Association Study Pipeline for Enrichment Platforms. https://bioconductor.org/packages/ramwas).

Quality control of MBD-seq data from post-mortem brain samples

Quality control for the post-mortem brain tissues largely followed the procedures used for the blood samples (see above). A summary of the samples available after QC is reported in **Table 2**. The mean number of reads after sample quality control was 57.5 million (SD = 18.6 million). The average alignment rate was 77.9%. After removing multi- and duplicate-reads, an average of 27.7 million reads (SD = 10.2 million reads) per sample remained. As previously described(3), to identify regions showing alignment problems, we conducted an *in silico* alignment experiment using the appropriate settings (50bp reads aligned with CUSHAW3(17)). With these settings 7.9% of the CpGs were removed from further analyses. After also excluding sporadically methylated CpGs (average coverage <0.3), 18.9 million CpGs remained for statistical analysis.

DNA methylation prediction model in post-mortem brain samples

To maximize the sample size of the data used for age prediction, we pooled all methylation data from post-mortem brain (regions BA10 and BA25) samples (N=211) available that had been generated using the same MBD-seq protocol and sequenced on the SOLiD system. Using the same 10-fold resampling approach as was described for the blood samples, chronological age could be predicted with a correlation of 0.69 (P<0.001) using 100,000 CpGs (Supplementary **Table S2**). The brain data included a subset of MDD patients (n=74) and matched controls (n=67) without psychiatric diagnosis used to examine case-control differences. We again regressed chronological age on DNAmAge, and used the unstandardized residuals as outcome measure.

Enrichment testing

To perform enrichment tests analysis of top MWAS findings in brain and blood, we used the R package shiftR. ShiftR cross-classifies CpGs as being in the top or bottom of the two MWAS. Using the resulting 2 by 2 tables as input, shiftR tests the null hypothesis that the enrichment odds ratio equals one. To perform these tests, it uses circular permutations(18) that destroy possible association signals while preserving the correlational structure between adjacent CpGs. Thus, it generates an empirical test statistic distribution under the null hypothesis that takes into account the dependency between CpGs. We used 1 million permutations for each enrichment test. Three thresholds of 0.5, 1 and 5% were specified to define the "top" MWAS findings. To account for this "multiple testing", the same thresholds were used in the permutations where the test statistic distribution under the null hypothesis is generated from most significant (combination of) thresholds. Of the overlapping sites in each comparison, only those below a combined q-value<0.1 were considered for pathway analysis.

SUPPLEMENTARY RESULTS

Accelerated aging and covariates

In addition, to further validate our main outcome measure, we examined the impact of all selected covariates on epigenetic aging (EA, unstandardized residuals of DNAmAge regressed on chronological age) with one multiple linear regression model (Supplementary **Table S2**). As expected based on earlier results (19–21) higher EA was associated with male sex (β=-0.11, P<0.001). In addition, higher EA was associated with increased body mass index (β=0.08, P=0.007) and low physical activity (β=-0.07, P=0.03). There was no relationship between EA and education (in years), cotinine levels, alcohol consumption or the number of chronic diseases under treatment (all Ps>0.05).

DNA methylation prediction model trained in controls-only

We repeated the between-group comparison of EA that was based on the prediction model trained in controls-only, rather than the full sample. This "controls-only model" included 10,000 CpGs, compared to the 80,000 CpG sites used in the "full sample model". *10*-fold cross-validation showed that chronological age could be predicted less precise, with a less slightly reliable correlation of 0.93 (P<0.001).

EA (from "controls only model") showed a normal distribution with, by design, a mean of zero (s.d.=3.92 years), ranging from -13.21 to 15.17. Depressed patients had significantly higher EA compared to controls (mean EA ± s.e. (CI) MDD: 0.21 ± 0.14 (-0.07, 0.48), controls: -0.52 ± 0.22 (-0.96, -0.08); $F_{(1,1121)}$=7.46, P=0.006, effect size (Cohen's d)=0.18) after full adjustment for covariates. Consistent with a dose-response association, a fully-adjusted linear regression showed that higher EA was significantly associated with higher IDS-score in the overall sample (β=0.10, P=0.002). Thus, although the "controls only model" resulted in a less accurate prediction of chronological age, likely due to the fact that it is trained in only 1/3 of the sample, results remained unchanged.

Cell type proportion correction

To examine whether the between-group difference in EA was not confounded by differences in blood cell composition, we performed additional analyses including cell type proportions (CD03, CD14, CD15) as covariates. Given that these cell type proportion estimates were highly dependent on some lab technical covariates, we also included those variables as covariates in the model. The analysis of covariance model showed that the between-group difference in EA remained significant ($F_{(1,1102)}$=3.92, P=0.048, effect size (Cohen's d)=0.14). Overall, these results indicate that MDD disease status explained the difference in EA independent from blood cell composition.

Accelerated aging and childhood trauma

To gain further insight into the association of EA with childhood trauma, we conducted a one-way analysis of variance of fully-adjusted EA against groups of controls and MDD patients with and without childhood trauma (CT, yes≥1 and no=0) Bonferroni-corrected for multiple testing. This revealed that MDD patients with CT showed the highest EA compared to controls without CT (P=0.001, Cohen's d=0.29), highlighting that this MDD+CT subgroup is associated with the highest EA (**Supplementary Figure S1**).

Post-hoc analyses with telomere length

To investigate the relationship of our main outcome measure with another biological aging marker available in our dataset (for detailed methods and telomere assays, see(22)), we performed post-hoc age-corrected partial and non-adjusted bivariate correlations between telomere length (TL) and EA. The results showed non-significant relationships of r=-0.03, P=0.34 and r=-0.03, P=0.37, respectively. Thus, EA and TL were not significantly correlated in our study.

Additional post-hoc analyses showed that TL did not change any of the study's findings when added as a covariate to the analysis of covariance model. Significantly higher EA was still uniquely observed in the MDD group (mean EA ± s.e. MDD: 0.18 ± 0.13, controls: -0.46 ± 0.20; $F_{(1,1120)}$=7.00, P=0.008). There were also significant main effects of sex, BMI, and physical activity (all Ps<0.05). However, there was no effect of TL on EA ($F_{(1,1120)}$=0.07, P=0.79).

To further investigate whether TL was associated with sociodemographic and lifestyle parameters, we performed a similar multiple linear regression model as performed with EA. As was expected from a previous study(22), shorter TL was significantly associated to male sex (β=0.08, P=0.009), BMI (β=-0.11, P=0.001), and increased alcohol intake (β=-0.08, P=0.01). TL was not associated to education, cotinine levels, physical activity or number of chronic diseases under treatment (all Ps>0.05).

SUPPLEMENTARY REFERENCES

1. Benowitz NL, Bernert JT, Caraballo RS, Holiday DB, Wang J. Optimal serum cotinine levels for distinguishing cigarette smokers and nonsmokers within different racial/ethnic groups in the United States between 1999 and 2004. Am J Epidemiol. 2009;169(2):236–48.
2. Hovens JGFM, Wiersma JE, Giltay EJ, Van Oppen P, Spinhoven P, Penninx BWJH, et al. Childhood life events and childhood trauma in adult patients with depressive, anxiety and comorbid disorders vs. controls. Acta Psychiatr Scand. 2010;122(1):66–74.
3. Aberg KA, Mcclay JL, Nerella S, Xie LY, Clark SL, Hudson AD, et al. MBD-seq as a cost-effective approach for methylome-wide assocciation studies: demonstration in 1500 case-control samples. Epigenomics. 2012;4(6):605–21.
4. Houseman E, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics [Internet]. 2012;13(1):86. Available from: http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-13-86
5. Koestler DC, Christensen BC, Karagas MR, Marsit CJ, Langevin SM, Kelsey KT, et al.

Blood-based profiles of DNA methylation predict the underlying distribution of cell types: A validation analysis. Epigenetics. 2013;8(8):816–26.

6. Hattab MW, Shabalin AA, Clark SL, Zhao M, Kumar G, Chan RF, et al. Correcting for cell-type effects in DNA methylation studies: reference-based method outperforms latent variable approaches in empirical studies. Genome Biol [Internet]. Genome Biology; 2017;18(24). Available from: http://genomebiology.biomedcentral.com/articles/10.1186/s13059-017-1149-7
7. Deep-Soboslay A, Iglesias B, Hyde TM, Bigelow LB, Imamovic V, Herman MM, et al. Evaluation of tissue collection for postmortem studies of bipolar disorder. Bipolar Disord. 2008;10(7):822–8.
8. Conner KR, Conwell Y, Duberstein PR. The validity of proxy-based data in suicide research: a study of patients 50 years of age and older who attempted suicide. II. Life events, social support and suicidal behavior. Acta Psychiatr Scand [Internet]. 2001;104(6):452–7. Available from: http://doi.wiley.com/10.1034/j.1600-0447.2001.00091.x
9. Kelly TM, Mann JJ. Validity of DSM-III-R diagnosis by psychological autopsy: a comparison with clinician ante-mortem diagnosis. Acta Psychiatr Scand [Internet]. 1996;94(5):337–43. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9124080
10. Dumais A, Lesage AD, Alda M, Rouleau G, Dumont M, Chawky N, et al. Risk factors for suicide completion in major depression: A case-control study of impulsive and aggressive behaviors in men. Am J Psychiatry. 2005;162(11):2116–24.
11. Gibbons AS, Brooks L, Scarr E, Dean B. AMPA receptor expression is increased post-mortem samples of the anterior cingulate from subjects with major depressive disorder. J Affect Disord [Internet]. Elsevier B.V.; 2012;136(3):1232–7. Available from: http://dx.doi.org/10.1016/j.jad.2011.10.001
12. Keks N et al. Diagnostic instrument for brain studies (DIBS). 1997;24(34).
13. Deep-Soboslay A, Benes FM, Haroutunian V, Ellis JK, Kleinman JE, Hyde TM. Psychiatric brain banking: Three perspectives on current trends and future directions. Biol Psychiatry [Internet]. Elsevier Inc.; 2011;69(2):104–12. Available from: http://dx.doi.org/10.1016/j.biopsych.2010.05.025
14. Alt SR, Turner JD, Klok MD, Meijer OC, Lakke EAJF, DeRijk RH, et al. Differential expression of glucocorticoid receptor transcripts in major depressive disorder is not epigenetically programmed. Psychoneuroendocrinology [Internet]. Elsevier Ltd;

2010;35(4):544–56. Available from: http://dx.doi.org/10.1016/j.psyneuen.2009.09.001

15. Torrey EF, Webster M, Knable M, Johnston N, Yolken RH. The Stanley Foundation brain collection and Neuropathology Consortium. Schizophr Res. 2000;44(2):151–5.
16. Van Den Oord EJCG, Clark SL, Xie LY, Shabalin AA, Dozmorov MG, Kumar G, et al. A whole methylome CpG-SNP association study of psychosis in blood and brain tissue. Schizophr Bull. 2016;42(4):1018–26.
17. Liu Y, Popp B, Schmidt B. CUSHAW3: Sensitive and accurate base-space and color-space short-read alignment with hybrid seeding. PLoS One. 2014;9(1).
18. Cabrera CP, Navarro P, Huffman JE, Wright AF, Hayward C, Campbell H, et al. Uncovering Networks from Genome-Wide Association Studies via Circular Genomic Permutation. G3: Genes|Genomes|Genetics [Internet]. 2012;2(9):1067–75. Available from: http://g3journal.org/lookup/doi/10.1534/g3.112.002618
19. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, et al. Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates. Mol Cell [Internet]. Elsevier Inc.; 2013;49(2):359–67. Available from: http://dx.doi.org/10.1016/j.molcel.2012.10.016
20. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol [Internet]. 2013;14(10):R115. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4015143&tool=pmcentrez&rendertype=abstract
21. Marioni RE, Shah S, McRae AF, Chen BH, Colicino E, Harris SE, et al. DNA methylation age of blood predicts all-cause mortality in later life. Genome Biol [Internet]. 2015;16(1):25. Available from: http://genomebiology.com/2015/16/1/25
22. Verhoeven JE, Révész D, Epel ES, Lin J, Wolkowitz OM, Penninx BWJH. Major depressive disorder and accelerated cellular aging: results from a large psychiatric cohort study. MolPsychiatry. 2014;19(8):895–901.

FIGURE S1. Epigenetic aging (EA) in major depressive disorder (MDD) and childhood trauma (CT). Mean EA by control and MDD group, with and without presence of CT. The means are adjusted for sex, education, body mass index, smoking, drinking, physical activity, and somatic diseases. The error bars represent 95% confidence intervals.

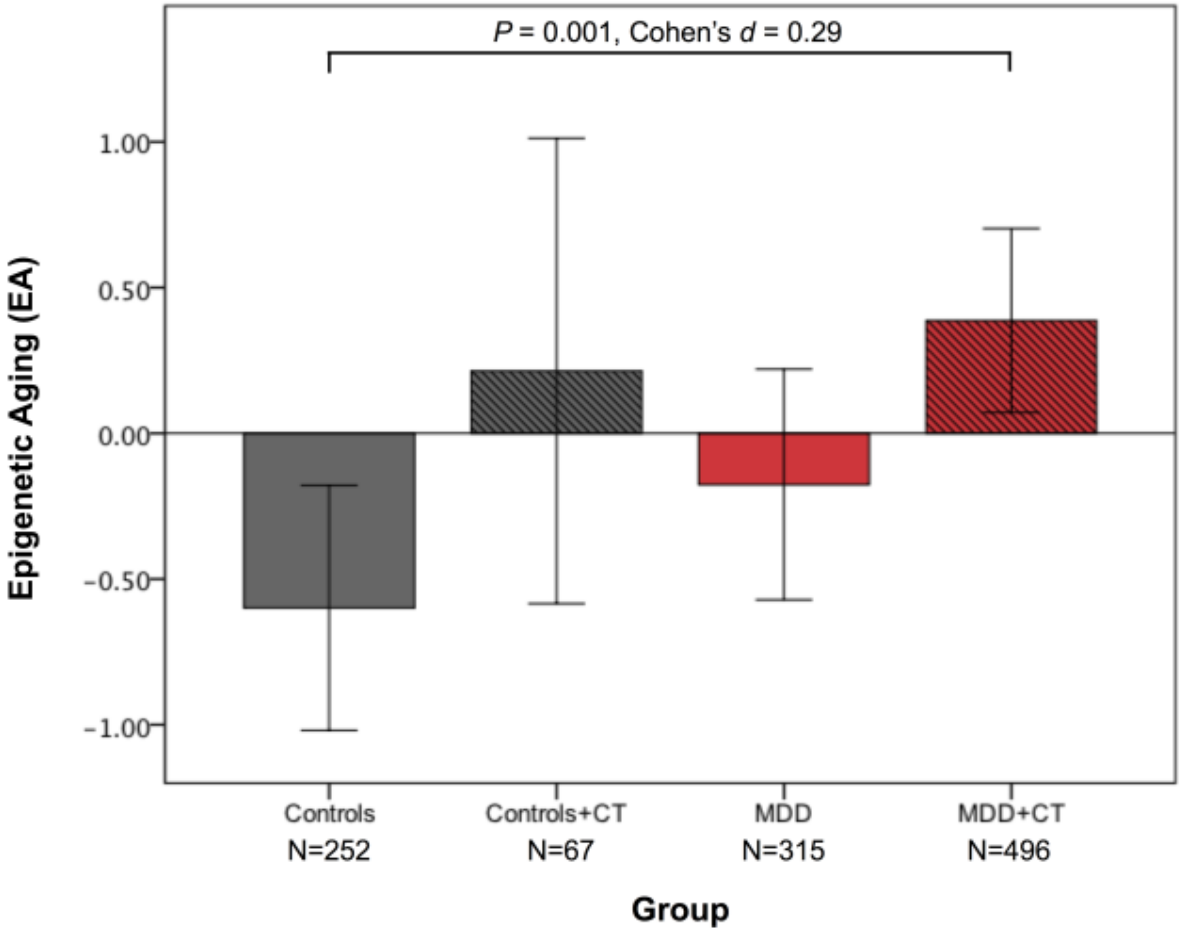

TABLE S1. ***See accompanying Excel files***

TABLE S2. ***See accompanying Excel files***

TABLE S3. The relationship between all selected covariates and epigenetic aging

	β	*P*-value
Sex (M/F)	-0.11	<0.001
Education (in years)	0.01	0.72
BMI	0.08	0.007
Cotinine levels (ng/ml)	0.04	0.15
Alcohol consumption (mean number of drinks/week)	-0.03	0.30
Physical activity (MET-minutes/week)	-0.07	0.03
Number of chronic diseases	0.001	0.98

Abbreviations: epigenetic aging, unstandardized residuals of DNAmAge regressed on chronological age; M, Male; F, female; BMI, Body Mass Index; MET, Metabolic Equivalent Total (MET level * minutes of activity * events per week).

TABLE S4. ***See accompanying Excel files***