Schizophrenia is characterized by age- and sex-specific effects on epigenetic aging

Anil P.S. Ori[1], Loes M. Olde Loohuis[1], Jerry Guintivano[2], Eilis Hannon[3], Emma Dempster[3],David St. Clair[4], Nick J Bass[5], Andrew McQuillin[5], Jonathan Mill[3,6], Patrick F Sullivan[2,7], Rene S. Kahn[8], Steve Horvath[9,10], Roel A. Ophoff[1,10,11].

1. *University of California Los Angeles, Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, Los Angeles, CA, USA*
2. *University of North Carolina, Department of Genetics, Chapel Hill, NC, US*
3. *University of Exeter, University of Exeter Medical School, Exeter, UK*
4. *University of Aberdeen, Institute of Medical Sciences, Aberdeen, Scotland, UK*
5. *University College London, Division of Psychiatry, UK*
6. *King's College London, London, UK*
7. *Karolinska Institutet, Department of Medical Epidemiology and Biostatistics, Stockholm, Sweden*
8. *Icahn School of Medicine at Mount Sinai, Department of Psychiatry, New York, NY, USA*
9. *University of California Los Angeles, Department of Biostatistics, Fielding School of Public Health, Los Angeles, CA, USA.*
10. *University of California Los Angeles, Department of Human Genetics, David Geffen School of Medicine, Los Angeles, CA, USA*
11. *Erasmus University Medical Center, Department of Psychiatry, Rotterdam, The Netherlands.*

Correspondence to: Anil Ori (anilori.contact@gmail.com) / Roel Ophoff (ophoff@ucla.edu)

Key words: DNA methylation, aging, schizophrenia, age acceleration, polygenic risk

Abstract

Schizophrenia (SCZ) is a severe mental illness that is associated with an increased prevalence of age-related disability and morbidity compared to the general population. An accelerated aging process has therefore been hypothesized as a component of the SCZ disease trajectory. Here, we investigated differential aging using three DNA methylation (DNAm) clocks (i.e. Hannum, Horvath, Levine) in a multi-cohort SCZ whole blood sample consisting of 1,100 SCZ cases and 1,200 controls. It is known that all three DNAm clocks are highly predictive of chronological age and capture different features of biological aging. We found that blood-based DNAm aging is significantly altered in SCZ with age- and sex-specific effects that differ between clocks and map to distinct chronological age windows. Most notably, the predicted phenotypic age (Levine clock) in female cases, starting at age 36 and beyond, is 3.21 years older compared to matching control subjects (95% CI: 1.92-4.50, P=1.3e-06) explaining 7.7% of the variance in disease status. Female cases with high SCZ polygenic risk scores present the highest age acceleration in this age group with +7.03 years (95% CI: 3.87-10.18, P=1.7E-05). Since increased phenotypic age is associated with increased risk of all-cause mortality, our findings suggests that specific and identifiable patient groups are at increased mortality risk as measured by the Levine clock. These results provide new biological insights into the aging landscape of SCZ with age- and sex-specific effects and warrant further investigations into the potential of DNAm clocks as clinical biomarkers that may help with disease management in schizophrenia.

Introduction

Schizophrenia (SCZ) is a severe psychiatric disorder with significant impact on the individual, their family, and society. A substantial proportion of people diagnosed with SCZ has chronic symptoms and disability across their lifespan[1–3]. SCZ is associated with an increased mortality[4–6] and a 15 year reduced life expectancy compared to the general population[7]. Despite elevated rates of suicide and other unnatural causes of death, most morbidity in SCZ is attributed to age-related diseases (e.g., cardiovascular and respiratory diseases, and diabetes mellitus[8,9,5]). The increased prevalence of age-related disabilities and morbidities suggests that biological aging may be accelerated in SCZ[10,11].

DNA methylation (DNAm) age predictors, or "epigenetic clocks", are biomarkers of ageing that generate a highly accurate estimate of chronological age, known as DNAm age[12–14]. The difference between (predicted) DNAm and chronological age (Δage) is associated with a wide-range of health and disease outcomes. For example, DNAm age acceleration in blood is associated with all-cause mortality[15–18], socioeconomic adversity and smoking[19], metabolic outcomes, such as body mass index and obesity[20,21], and brain-related phenotypes, such as Parkinson's disease, posttraumatic stress disorder, insomnia, major depressive disorder, and bipolar disorder[22–24,25,26].

While different aging biomarkers have been studied, there is no clear demonstration of altered aging in SCZ[11]. More specifically, DNAm age predictors have found limited to no evidence for altered epigenetic aging in either brain or blood [27–31]. These studies, however, (i) consisted of small sample sizes and thus limiting the ability to detect a biological signal, (ii) used a single epigenetic clock that may have not been most informative for aging studies of psychiatric disorders, and (iii) did not consider aging differences across the life-span of patients. As morbidities in the SCZ population differ between older and younger individuals, and females and males [5], analyses of both age- and sex-specific effects is warranted and could identify differential aging patterns nevertheless.

To investigate DNAm aging in SCZ, we used three independent DNAm age estimators; the Hannum[13], Horvath[12,32], and Levine clock[14]. Each clock is designed using

different training features and capture distinct characteristics of aging[12,32]: the Hannum age predictor was trained on whole blood adult samples, the Horvath predictor was trained across 30 tissues and cell types across developmental stages, and the Levine predictor combines DNAm from adult blood samples with clinical blood-based measures. As the Levine estimator is trained on chronological age and nine clinical markers, its output is referred to as DNAm PhenoAge or “phenotypic age”. The Hannum estimator captures measures of cell extrinsic aging in blood, whereas the Horvath clock measures cell intrinsic aging as it was trained across multiple tissues and therefore is less dependent on cell type composition. All three clocks, in different but complementary ways, capture the pace of biological aging that is associated with age-related conditions and diseases, including all-cause mortality[14,32].

Here, we used these clocks to investigate DNAm aging in SCZ using four European case-control cohorts. Analysis are performed across the full sample and stratified by age and sex. SCZ polygenic risk scores (PRS), age at onset, duration of illness, DNAm smoking scores, and blood cell type proportions were used to gain further insights into differential aging patterns. This study overall reports an in-depth investigation of the DNAm aging landscape in schizophrenia.

Material and Methods

Cohort and sample description

Details of samples included in this study can be found in the Supplementary Information. Briefly, unrelated individuals diagnosed with SCZ and ancestry-matched non-psychiatric controls from four cohorts of European ancestry were included; the Netherlands (N=1,116), Scotland (N=847), Sweden (N=96), and the United Kingdom (N=675). Cases were selected on the basis of a clinical diagnosis of SCZ using the Diagnostic and Statistical Manual for Mental Disorders (DSM-IV), Research Diagnostic Criteria (RDC), or the International Classification of Diseases 10 (ICD10). Controls were unaffected subjects without a history of any major psychiatric disorder. Whole blood DNAm data was available for a total of 2,707 samples (1,399 cases and 1,308 controls; (Table S1).

Genome-wide DNA methylation profiling and data processing

To quantify DNA methylation, DNA was extracted from whole blood and bisulfite converted for hybridization to the Illumina Infinium Human Methylation Beadchip. Samples were assayed with either the 27K or 450K beadchip, which contain 27,578 and 485,512 probes that interrogate CpG sites across the genome, respectively. For each platform, data processing pipelines were implemented, which includes background correction, color channel and probe type correction, and normalization of the data, to minimize the effect of technical variation on the final beta values. Samples with more than 5% of probes detected at $P > 0.05$ were excluded from further analyses (n=13). Full details are described in the supplementary methods.

DNAm-based estimation of biological age

To compute blood-based DNAm age estimates, processed beta values were used as input to the Hannum[13], Horvath[12], and Levine[14] DNAm clock. These DNAm age estimators use a set of CpGs that are selected via an optimization algorithm to collectively minimize the error associated with estimating chronological age (Supplementary Information). Horvath

DNAm age estimates were calculated using R scripts from the Horvath DNA Methylation Calculator (https://dnamage.genetics.ucla.edu)[12]. Hannum and Levine estimates were obtained by using the reported set of probes with corresponding regression weights. We define Δage by subtracting chronological age at the time of the blood draw from the predicted DNAm age.

Statistical analyses

To investigate epigenetic aging differences in SCZ, we first removed samples with discrepant phenotypic sex and predicted sex based on DNAm data (n=9), as well as samples with missing chronological age data (n=237), bipolar disorder diagnosis (n=26), and duplicate samples (n=126). For each epigenetic clock, we regressed Δage on technical principal components (PCs), using the first components that cumulatively explain >90% of variation in intensity values of control probes, and added the residuals to mean(Δage) to generate a measure in the same units as Δage that is adjusted for technical variation (Δage-adjusted). We used the adjusted value for subsequent analyses and refer to it as Δage.

To avoid skewing of chronological age distribution between groups[33], we removed any case older than the oldest control subject in each cohort (n = 5 for NLD, 16 for SCT, 4 for SWD, and 1 for UK). Chronological age was furthermore included as a covariate in all analyses. To minimize the effect of outlying samples, we excluded samples >3SD from mean Δage across cohorts (ranging from n=13 to 16 for the three clocks). These are samples for which DNAm age diverged substantially from chronological age, which are likely artifacts.

For each clock and each cohort, we implemented a multivariable regression model predicting Δage as a function of schizophrenia status, sex, and age. For the Dutch cohort, batch and array platform were also included as covariates, as this cohort consists of multiple datasets. For each clock, regression coefficients with corresponding standard errors for each of the four cohorts were then supplied to the rma() function of the metafor package[34] in R to fit a meta-analytic fixed-effect model with inverse-variance weights and obtain an overall

effect size and test statistic. To quantify the significance of age- and sex-specific effects, we determined the contribution of interaction effects on top of the main disease effect. We first combined all cohorts to maintain necessary sample sizes across age and sex groups. Age groups were defined by grouping samples by decades with ages 18 and 19 included in the first decade (18-30, 31-40, etc.). To quantify the gain in variance explained in Δage, models with the interaction term were compared to a baseline model without the interaction term. For each analysis, statistical significance was determined using Bonferroni correction, i.e. $P < 0.05$ / number of tests.

SCZ polygenic risk quantification

Polygenic risk scores (PRS) were obtained from analyses of the SCZ GWAS conducted by Psychiatric Genomics Consortium (PGC)[35]. Using a leave one out approach, weights were generated in a training dataset based on all samples minus the target cohort in which the PRS were calculated. For each individual, weighted single nucleotide polymorphisms (SNPs) were summed to a genetic risk score that represents a quantitative and normally distributed measure of SNP-based SCZ genetic risk. To reduce between cohort-variation and maximize statistical power, we used a previously developed analytical strategy that uses principal component analysis (PCA) to concentrate disease risk across PRSs of ten GWAS p-value thresholds into the first principal component (PRS1)[36] (Supplementary Information). PRS1 explains 70.7% of the variance in risk scores and 19.9% of the variance in SCZ status; the remaining PCs had no explanatory value in disease status (mean R^2 = 0.0%), indicating that PRS1 captures the majority of SNP-based SCZ polygenic risk. PRS1 was generated for 1,933 individuals, 853 cases and 1080 controls, and modeled as both a quantitative and categorical variable to predict Δage.

Defining age at onset and illness duration

Age at onset is defined as the earliest reported age of psychotic symptoms or by the Operational Criteria Checklist (OPCRIT), depending on the cohort. This data is available for

a subset of cases (N = 710) across the Dutch, Scottish, and UK cohorts. Illness duration is defined as the time between age at onset and blood collection. A more detailed description of each cohort's definition is available in the Supplementary Information.

DNA methylation-based smoking scores and blood cell type proportions

Smoking scores and blood cell type proportions were estimated from the data (see Supplementary Methods) and used as a proxy to further decompose differential aging effects.

Estimating the contribution of differential aging in schizophrenia

Using a multivariable logistic regression model for disease status, we fitted batch, cohort, DNAm smoking score, DNAm blood cell type proportions, and Δage as explanatory variables. We first performed a variable reduction step to select the most contributing variables to disease status by use of a regularized logistic regression using the glmnet() function in R ("glmnet" package, v2.13)[37]. Alpha was set to "1" (Lasso) and the lambda parameter estimated at the optimal value that minimizes the cross-validation prediction error rate using the cv.glmnet() function. For each selected variable, we then report the variance explained in SCZ status (glm, family = "binomial") for both the individual variable as well as adjusted for all other selected variables using the NagelkerkeR2() function in the "fmsb" package (v 0.6.3). The significance of each variable their contribution was computed by comparing the model with and without the variable of interest using the likelihood ratio test of the anova() function.

Availability of data and materials

The datasets used are available on the NCBI Gene Expression Omnibus (GEO) data repository, the European Genome-phenome Archive (EGA), or via the principal investigator of each cohort. See Table S2 and S3 for an overview and corresponding accession series numbers. See Table S4 for sample information, including DNAm age estimates.

Results

Figure 1 shows a schematic overview of the study design and analysis framework used to investigate DNAm aging in SCZ. After data preprocessing and quality control, 1,090 SCZ cases and 1,206 controls (2,296 subjects of 2,707 initial samples) were included in our analysis. The overall sample has a mean age of 40.3 years (SD=14.4) and consists of 34.5% women (Table S1 and Figure S1).

[FIGURE 1 ABOUT HERE - SCHEMATIC OVERVIEW OF STUDY DESIGN]

Across cohorts, all three clocks produce a high correlation with chronological age (Pearson's r = 0.92-0.94; Figure 2A and S2). Using duplicates in the Dutch cohort, we assessed consistency between pairs of technical replicates, i.e. samples for which blood was collected at the same time but DNA processed at different times and DNAm data obtained on different arrays. Comparing Δage estimates between these pairs, we find a significant correlation for each clock (Figure S3); Hannum (rho = 0.79, n = 10), Horvath (rho = 0.53, n=118), Levine (rho = 0.67, n=118). Δage directionality (i.e. age deceleration or acceleration) is concordant in 90%, 73%, and 86% of pairs for Hannum, Horvath, and Levine, respectively, highlighting that the obtained estimates of DNAm age are reproducible for all three clocks. Comparing Δage estimates between clocks using all samples, we find a moderate concordance (Pearson's r = 0.39-0.43; Figure S4), demonstrating that a significant proportion of the variation in Δage is clock-specific. This indicates that these estimators capture different features of biological aging and that investigating all three epigenetic clocks simultaneously may thus yield broader insights into differential aging.

[FIGURE 2 ABOUT HERE]

DNA methylation age is altered in an age-dependent manner

Across the full sample, SCZ cases are on average 1.53 years older in phenotypic Δage (Levine clock) compared to controls (P_{meta}= 3.45E-08) (Figure 2B). The intrinsic cellular age (Horvath) predictor revealed an opposite pattern, with SCZ cases appearing 0.47 years younger compared to controls (P_{meta}= 0.06). No differences were observed between cases and controls when applying the blood-based Hannum DNAm age predictor. No evidence of heterogeneity between the four cohorts was observed for any of the DNAm age analyses (P_{het} > 0.05, Tables S5-S8).

Modeling the interaction effect between disease status and chronological age on Δage reveals a differential rate of aging between cases and controls (Figure 2C). That is, the slope of Δage across chronological age is 0.05 and 0.06 years steeper in cases compared to controls for the Horvath (P_{meta}=2.3E-03) and Levine clocks (P_{meta}=7.1E-03), respectively (Figure S5 and Table S6). As no significant effects were observed for the Hannum Δage, we decided to focus our downstream analysis on the phenotypic (Levine) age and intrinsic cellular (Horvath) age only. To further disentangle the relationship between Δage in SCZ conditional on chronological age, we estimated differential aging by 10 year intervals, with years 18 and 19 included in the first age group. We observe significant DNAm age deceleration in early adulthood (18-30 years) with cases estimated at -1.23 years younger (P_{meta}=3.9E-03) in intrinsic cellular age with no significant difference at later ages (Figure 3). In phenotypic age, SCZ cases displayed significant DNAm age acceleration from 30 years and older, with the most pronounced age acceleration between 50-60 years (2.29 years, P_{meta}=9.0E-03). We find no evidence of heterogeneity between cohorts (Figure S6 and Table S7-8).

[FIGURE 3 ABOUT HERE]

Age- and sex-specific effects contribute independently to DNAm aging

To quantify the contribution of age- and also sex-specific effects, we estimated the gain in variance explained of Δage by adding the interaction terms of age and sex with disease status to a baseline model and assesed the gain in model performance (Table 1 and S9). For intrinsic cellular aging, the baseline model explains 4.0% of the variance of Horvath Δage. The variance explained increases significantly to 5.5% and 5.9% when a two-way interaction between status and (categorical) age and a three-way interaction between status, age, and sex is added, respectively. For phenotypic age, the baseline model explains 3.2% of the variance of Levine Δage. Similarly, the variance-explained of the Levine Δage signifcantly increases to 3.9% and 4.7%, respectively, with the interaction terms added to the model. For both measures of aging, inclusion of interaction terms presented a significantly better fit, especially with the three-way interaction model (i.e. disease status, categorical age and biological sex). We do observe a larger gain in model fit for the three-way interaction for phenotypic aging (P=0.01) than for intrinsic cellular aging (P=0.24), suggesting that sex-specific effects may be less pronounced for Horvath Δage.

[TABLE 1 ABOUT HERE]

Estimating and mapping windows of differential aging in schizophrenia

In order to refine the age-dependent aging effects and decompose its effect to more specific age windows, we implemented a sliding window approach across chronological age, both in the full sample and within each sex separately. Using 5-year bins and sliding steps of 1 year, we tested cases versus age-matched controls and constructed a more precise picture of differential aging across chronological age in SCZ (Figure 4). We mapped changes in Δage to specific ages with different patterns between men and women. For intrinsic cellular age, we observe a deceleration effect during early adulthood from 29 years and younger across all samples, with the shift in differential aging occurring earlier in women (<25) (Figure 4B). While we previously did not observe differences in intrinsic cellular aging

across other categorical age groups (Figure 3 and S7), the higher resolution, sliding window analysis reveals additional disease effects. For both men and women, we observe age deceleration in mid-forties and for women we also find age acceleration between 50-56 years (Figure 4A-C).

For phenotypic age, we mapped the age acceleration effect to 27 years and older across the whole sample with male / female differences (Figure 4D-F). In women, we observe age acceleration between 25-29 years and from 36 years and older (Figure 4E). In men, we see age acceleration between 27-39 and 49-59 years (Figure 4F). More details on each age window and corresponding effect sizes are shown in Table S10. Thus far, our results show that DNAm aging, measured through the Horvath and Levine clock, is significantly different in SCZ and that this overall effect is characterized by age-specific effects with some distinctions between the sexes, particularly for Levine Δage.

[FIGURE 4 ABOUT HERE]

Age deceleration by multi-tissue Horvath clock is not present in brain

In order to examine whether the blood DNAm age findings are also observed in the brain, we investigated DNAm aging in frontal cortex postmortem brain samples of 221 SCZ cases and 278 controls. The multi-tissue Horvath clock accurately predicts DNAm age in brain as well (r=0.94, P < 2.2e-16). We however find no difference in DNAm aging between cases and controls (ß=-0.29, P=0.46) and no evidence of age-dependent aging either. More details are shown in the Supplementary Results (S2.1).

Phenotypic age acceleration is associated with SCZ polygenic risk in women

To further decipher the factors underlying the signal of differential aging in SCZ, we examined the possible role of SCZ polygenic risk (PRS1), age at onset, and illness duration (Figure S8). We first focus on the phenotypic age acceleration in female cases of age 36 years and older, the age range which we identified as showing the most consistent and

significant aging effect. We find stronger age acceleration in cases with both low and high SCZ genetic risk (Table 2). Cases in the highest PRS1 tertile are predicted to be 4.30 years older in phenotypic age compared to controls (P=1.3E-05), cases with median range PRS are 1.89 years older (P=4.5E-02), and cases in the lowest quartile are 2.89 years older (P=2.8E-03). By permutation of PRS1 groups, we find that the observed effect in female cases in the highest PRS1 group is unlikely to occur by chance (P=0.024). For the association between Levine Δage and PRS1 to be most pronounced in the low and high group, is even less likely to happen by chance (P=0.006). At maximum, this group of women carrying high SCZ genetic risk have on average 7.03 higher phenotypic Δage (95% CI: 3.87-10.18; P=1.7E-05) (Figure 5A). We do not observe such an association in women age < 36 years, men with age > 36 years, nor across the whole dataset (Figure S9). For age at onset and illness duration, we did not find significant association with Δage across partitioned bins (after permutation, P > 0.05) (Table 2). This is further confirmed when we integrated these two variables across PRS1 tertiles, demonstrating that the most pronounced differences in Δage are observed across PRS1 bins and not across the distribution of age at onset and illness duration (Figure 5B-C).

[TABLE 2 ABOUT HERE]

We conducted a similar investigation on the observed intrinsic cellular age deceleration in all SCZ cases age 29 years and younger but found no significant associations between Horvath Δage and PRS1, age at onset, or illness duration (Table S11 and Figure S10). While we did observe the strongest Horvath age deceleration in the high PRS1 tertile (β=-1.58, P=3.0E-03), this was not significant after permutation analysis (P>0.05). We did not analyze other identified age windows of differential aging as these either had too few individuals with genetic or phenotypic information available or more modest disease effects limiting any further stratification.

[FIGURE 5 ABOUT HERE]

DNAm aging affects SCZ above and beyond smoking and blood cell types

People suffering from SCZ smoke more than the general population[38] (Figure 5A) and blood cell type composition changes across the lifespan[39] (Figure 5B). To investigate the effect of these factors, we use DNAm-based smoking and cell type estimations (see Methods) as a proxy to evaluate their contribution to DNAm aging in SCZ. While DNAm clocks, by design, will encapsulate such effects, quantifying the contributions of each factor increases interpretability and helps understand the factors contributing to the differential aging findings. In women age 36 and older, the patient group in which we observed the most profound aging effects, a regularized logistic regression selected dataset/ethnicity, smoking, 5 cell types, and Levine Δage to explain a total of 23.6% of the variance in SCZ disease status (P=2.2E-08) (Table S12). Levine Δage explains 7.7% individually and 2.8% (P=3.3E-03) when adjusted for other selected variables (Figure 6C).

In individuals 29 years and younger, the group with the significant deceleration aging effects, the lasso regression selected batch/ethnicity, age, sex, smoking, 7 cell types, and Horvath Δage, which together explain 49.8% of the variance in disease status (Table S12). A large proportion of this effect is driven by smoking, which explains 28.8%. Horvath Δage explains 3.1% of the variance in SCZ individually and 0.6% adjusted for other select variables (P=0.14). A significant proportion of the Horvath Δage effect on disease status is reduced by adjusting for smoking. However, smoking has no association with Horvath Δage in controls (Pearson r=0.01, P=0.95) nor in cases (Pearson r=-0.08, P=0.28) (Figure S12). As smoking is confounded with SCZ disease status, it is difficult to distinguish these signals. In relation to SCZ genetic risk, smoking and blood cell types demonstrate limited effects on the observed pattern of differential aging across PRS1 (Figure S13).

[FIGURE 6 ABOUT HERE]

Discussion

We performed one of the largest aging and epigenetic studies in schizophrenia to date using multiple epigenetic clocks based on whole blood DNA methylation data. We observe significant patterns of sex-specific and age-dependent DNAm aging in SCZ, a finding consistent across four European cohorts. The most significant differential aging pattern that we observe is in females ages 36 years and older in which we detect advanced *phenotypic age acceleration*, as measured by the Levine clock, explaining 7.7% of the variance in disease status. We also observe *intrinsic cellular age deceleration* in SCZ during early adulthood, as measured by the Horvath clock, explaining 3.1% of the variance in disease status. The phenotypic age acceleration in females diagnosed with SCZ is associated with a high burden of SCZ polygenic risk. This high SCZ risk group displays accelerated aging of an average of 7 years compared to age-matched female controls. Our findings suggests that specific and identifiable patient groups are at increased mortality risk as measured by the Levine clock.

The Levine estimator was constructed by predicting a surrogate measure of phenotypic age, which is a weighted average of 10 clinical markers, including chronological age, albumin, creatinine, glucose and C-reactive protein levels, alkaline phosphatase and various blood cell related measures[14]. By design, the Levine estimator is a composite biomarker that strongly predicts mortality, in particular that of cardiovascular-related phenotypes. A 1-year increase in Levine DNAm age is associated with a 9% increased risk of all-cause mortality and a 10% and 20% increase of cardiovascular disease and diabetes mortality risk, respectively[14,40]. Our findings of multiple year increase in Levine DNAm age in SCZ could thus imply an increased mortality in these individuals that is linked to disease, a previously well-established epidemiological observation[4–6]. A recent study however found that DNAm age acceleration only predicts mortality in SCZ cases without pre-existing cancer using the Hannum clock[41]. They did not find such evidence using the Levine clock. The smaller sample size and predominantly male cohort may have reduced the predictive power of the study. Our findings warrant a more focused and larger study of DNAm aging in

females in later adulthood, preferably stratified by SCZ genetic risk. Our results align well with the observation that people diagnosed with SCZ, particularly women, are reported to be at high mortality risk due to cardiovascular disease and diabetes[5,42,43]. Assuming that cardiovascular risk is modifiable in SCZ[44], Levine DNAm age could serve as a potential biomarker to identify at-risk individuals and in this way help with disease management and improvement of life expectancy.

In contrast to *age acceleration* in Levine DNAm age, we observe *age deceleration* in intrinsic cellular age (i.e. the Horvath DNAm age), an effect that is most pronounced in cases age 29 and younger. Unlike the association in females, we did not observe clear patterns with genetic and phenotypic variables that could help deconstruct the signal. Horvath Δage furthermore showed strong age-specific effects but less clear sex-specific effects. We did not observe *age deceleration* in postmortem brain samples of the human cortex, indicating the the observe signal may be blood-specific. Horvath DNAm aging has been shown to associate with molecular processes of development and cell differentiation[12,32], including human (neuro)developmental phenotypes[45,46]. Our findings may indicate that individuals diagnosed with SCZ in this age group show evidence of delayed or deficient development and that this is detectable in blood through the multi-tissue Horvath clock. This however remains speculative and future work is needed to further dissect how blood-based *Horvath age deceleration* is associated with SCZ.

While we did observe aging effects with the Horvath and Levine clock, we did not with the Hannum clock. The Hannum clock is less predictive of age acceleration effects on mortality risk than the Levine clock[14], which could explain the lack of findings in our analyses. The Hannum estimator furthermore cannot be used on first generation 27K DNA methylation arrays which reduced the sample size of this study with 30% and may have impacted the statistical power of these specific analyses. This highlights the benefits of designing methods that are inclusive to all platforms, so all data, both old and new, can be leveraged.

Previous studies that investigated DNAm aging in SCZ did not observe significant differential aging [27–31], highlighting that sample size and/or explicit modeling of age- and sex-dependent effects are crucial factors when investigating aging in SCZ. Our findings also demonstrate that analyzing these clocks simultaneously can reveal new insights that may otherwise be missed. A systematic review of aging biomarkers found that less than a quarter of studies explored an interaction effect or statistically compared the regression slope between groups in SCZ[11]. Our findings support their recommendations to specifically examine age-specific effects in aging studies but also more general in epigenetic studies of SCZ, such as epigenome-wide association studies. Future work should also be extended to integrate nonlinear models to fully capture the complex relationship between DNAm aging and clinically relevant variables across the lifespan. These models will help validate and further refine the most relevant age intervals.

A limitation of the study is the cross-sectional design of the cohorts used. While we do find an association with SCZ polygenic risk, dissecting cause-and-effect relationships remains challenging. Large-scale longitudinal prospective cohorts with detailed clinical and genomic information are needed to further examine differential aging in SCZ and assess its clinical relevance above and beyond other known health risk factors and disease biomarkers, such as medication use. In addition, improvement of existing and/or development of new DNAm age biomarkers may help to better study differential aging in SCZ and related disorders with increased mortality. Combining blood-based DNAm age with that of other aging profiles, such as MRI-based brain age[47], may further advance our understanding of aging and SCZ disease progression, including the increased mortality[48].

Schizophrenia is a debilitating disorder that is associated with severe and oftentimes chronic disability. While health and life expectancy of the general population continues to improve, the mortality disparity between people suffering from schizophrenia and those unaffected continues to increase in the United States and the UK[8,9,42]. This highlights the importance of studies on improving the health and longevity of people with schizophrenia and other severe mental illnesses. Molecular biomarkers of aging, such as DNAm clocks,

are emerging as candidate tools for screening and intervention. Our findings provide evidence of its potential in clinical care of SCZ. This study strengthens the need for future large-scale epidemiological studies of DNA methylation aging in SCZ, a population vulnerable to age-related diseases and excess mortality.

Declarations

Funding

This work was supported by the US NIH under award number R01 DA028526, R01 MH078075, R21 MH098035, R01 MH115676, RF1AG058484 granted to RAO. LMOL was supported by the NIH under award number K99 MH116115. PFS was supported by the Swedish Research Council (Vetenskapsrådet, award D0886501), the Horizon 2020 Program of the European Union (COSYN, RIA grant agreement n° 610307), and US NIMH (U01 MH109528 and R01 MH077139).

Authors' contribution

APSO and RAO planned the study. RAO supervised the study. APSO was responsible for data analysis and presentation of results, LMOL, SH, and RAO provided feedback. APSO, LMOL, RAO interpreted the main findings and wrote the manuscript, with input from co-authors. RAO, RSK, EH, ED, DSC, NJB, AM, JM, JG and PFS provided access to data of cohorts included in the study. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

References

1. Robinson, D. G., Woerner, M. G., McMeniman, M., Mendelowitz, A. & Bilder, R. M. Symptomatic and functional recovery from a first episode of schizophrenia or schizoaffective disorder. *Am. J. Psychiatry* **161**, 473–479 (2004).
2. Harrison, G. *et al.* Recovery from psychotic illness: a 15- and 25-year international follow-up study. *Br. J. Psychiatry* **178**, 506–517 (2001).
3. Owen, M. J., Sawa, A. & Mortensen, P. B. Schizophrenia. *Lancet* **388**, 86–97 (2016).
4. McGrath, J., Saha, S., Chant, D. & Welham, J. Schizophrenia: a concise overview of incidence, prevalence, and mortality. *Epidemiol. Rev.* **30**, 67–76 (2008).
5. Olfson, M., Gerhard, T., Huang, C., Crystal, S. & Stroup, T. S. Premature Mortality Among Adults With Schizophrenia in the United States. *JAMA Psychiatry* **72**, 1172–1181 (2015).
6. Allebeck, P. Schizophrenia: A Life-shortening Disease. *Schizophrenia Bulletin* **15**, 81–89 (1989).
7. Hjorthøj, C., Stürup, A. E., McGrath, J. J. & Nordentoft, M. Years of potential life lost and life expectancy in schizophrenia: a systematic review and meta-analysis. *Lancet Psychiatry* **4**, 295–301 (2017).
8. Saha, S., Chant, D. & McGrath, J. A systematic review of mortality in schizophrenia: is the differential mortality gap worsening over time? *Arch. Gen. Psychiatry* **64**, 1123–1131 (2007).
9. Hayes, J. F., Marston, L., Walters, K., King, M. B. & Osborn, D. P. J. Mortality gap for people with bipolar disorder and schizophrenia: UK-based cohort study 2000-2014. *Br. J. Psychiatry* **211**, 175–181 (2017).
10. Kirkpatrick, B., Messias, E., Harvey, P. D., Fernandez-Egea, E. & Bowie, C. R. Is schizophrenia a syndrome of accelerated aging? *Schizophr. Bull.* **34**, 1024–1032 (2008).
11. Nguyen, T. T., Eyler, L. T. & Jeste, D. V. Systemic Biomarkers of Accelerated Aging in

Schizophrenia: A Critical Review and Future Directions. *Schizophr. Bull.* **44**, 398–408 (2018).

12. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol.* **14**, R115 (2013).
13. Hannum, G. *et al.* Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol. Cell* **49**, 359–367 (2013).
14. Levine, M. E. *et al.* An epigenetic biomarker of aging for lifespan and healthspan. *Aging* **10**, 573–591 (2018).
15. Marioni, R. E. *et al.* DNA methylation age of blood predicts all-cause mortality in later life. *Genome Biol.* **16**, 25 (2015).
16. Chen, B. H. *et al.* DNA methylation-based measures of biological age: meta-analysis predicting time to death. *Aging* **8**, 1844–1865 (2016).
17. Perna, L. *et al.* Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort. *Clin. Epigenetics* **8**, 64 (2016).
18. Levine, M. E. *et al.* DNA methylation age of blood predicts future onset of lung cancer in the women's health initiative. *Aging* **7**, 690–700 (2015).
19. Fiorito, G. *et al.* Social adversity and epigenetic aging: a multi-cohort study on socioeconomic differences in peripheral blood DNA methylation. *Sci. Rep.* **7**, 16266 (2017).
20. Quach, A. *et al.* Epigenetic clock analysis of diet, exercise, education, and lifestyle factors. *Aging* **9**, 419–446 (2017).
21. Horvath, S. *et al.* Obesity accelerates epigenetic aging of human liver. *Proceedings of the National Academy of Sciences* 201412759 (2014).
22. Horvath, S. & Ritz, B. R. Increased epigenetic age and granulocyte counts in the blood of Parkinson's disease patients. *Aging* **7**, 1130–1142 (2015).
23. Boks, M. P. *et al.* Longitudinal changes of telomere length and epigenetic age related to traumatic stress and post-traumatic stress disorder. *Psychoneuroendocrinology* **51**, 506–512 (2015).

24. Carroll, J. E. *et al.* Epigenetic Aging and Immune Senescence in Women With Insomnia Symptoms: Findings From the Women's Health Initiative Study. *Biol. Psychiatry* **81**, 136–144 (2017).

25. Han, L. K. M. *et al.* Epigenetic Aging in Major Depressive Disorder. *Am. J. Psychiatry* appiajp201817060595 (2018).

26. Fries, G. R. *et al.* Accelerated epigenetic aging and mitochondrial DNA copy number in bipolar disorder. *Transl. Psychiatry* **7**, 1283 (2017).

27. Voisey, J. *et al.* Epigenetic analysis confirms no accelerated brain aging in schizophrenia. *NPJ Schizophr* **3**, 26 (2017).

28. McKinney, B. C., Lin, H., Ding, Y., Lewis, D. A. & Sweet, R. A. DNA methylation evidence against the accelerated aging hypothesis of schizophrenia. *NPJ Schizophr* **3**, 13 (2017).

29. McKinney, B. C., Lin, H., Ding, Y., Lewis, D. A. & Sweet, R. A. DNA methylation age is not accelerated in brain or blood of subjects with schizophrenia. *Schizophr. Res.* (2017). doi:10.1016/j.schres.2017.09.025

30. Okazaki, S. *et al.* Epigenetic clock analysis of blood samples from Japanese schizophrenia patients. *NPJ Schizophr* **5**, 4 (2019).

31. Viana, J. *et al.* Schizophrenia-associated methylomic variation: molecular signatures of disease and polygenic risk burden across multiple brain regions. *Hum. Mol. Genet.* **26**, 210–225 (2017).

32. Horvath, S. & Raj, K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat. Rev. Genet.* **19**, 371–384 (2018).

33. Khoury, L. E. *et al.* Properties of the epigenetic clock and age acceleration. doi:10.1101/363143

34. Viechtbauer, W. Conducting Meta-Analyses inRwith themetaforPackage. *J. Stat. Softw.* **36**, (2010).

35. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).

36. Bergen, S. E. *et al.* Joint Contributions of Rare Copy Number Variants and Common SNPs to Risk for Schizophrenia. *Am. J. Psychiatry* **176**, 29–35 (2019).

37. Friedman, J., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* **33**, 1–22 (2010).

38. Kelly, C. & McCreadie, R. Cigarette smoking and schizophrenia. *Advances in Psychiatric Treatment* **6**, 327–331 (2000).

39. Jaffe, A. E. & Irizarry, R. A. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biol.* **15**, R31 (2014).

40. Liu, Z. *et al.* A new aging measure captures morbidity and mortality risk across diverse subpopulations from NHANES IV: A cohort study. *PLoS Med.* **15**, e1002718 (2018).

41. Kowalec, K. *et al.* Methylation age acceleration does not predict mortality in schizophrenia. *Transl. Psychiatry* **9**, 157 (2019).

42. Osby, U., Correia, N., Brandt, L., Ekbom, A. & Sparén, P. Time trends in schizophrenia mortality in Stockholm county, Sweden: cohort study. *BMJ* **321**, 483–484 (2000).

43. Galletly, C. A. *et al.* Cardiometabolic risk factors in people with psychotic disorders: the second Australian national survey of psychosis. *Aust. N. Z. J. Psychiatry* **46**, 753–761 (2012).

44. Kugathasan, P. *et al.* Association of Secondary Preventive Cardiovascular Treatment After Myocardial Infarction With Mortality Among Patients With Schizophrenia. *JAMA Psychiatry* **75**, 1234 (2018).

45. Jeffries, A. R. *et al.* Growth disrupting mutations in epigenetic regulatory molecules are associated with abnormalities of epigenetic aging. *Genome Res.* **29**, 1057–1066 (2019).

46. Hoshino, A., Horvath, S., Sridhar, A., Chitsazan, A. & Reh, T. A. Synchrony and asynchrony between an epigenetic clock and developmental timing. *Sci. Rep.* **9**, 3770 (2019).

47. Schnack, H. G. *et al.* Accelerated Brain Aging in Schizophrenia: A Longitudinal Pattern Recognition Study. *Am. J. Psychiatry* **173**, 607–616 (2016).

48. Cole, J. H., Marioni, R. E., Harris, S. E. & Deary, I. J. Brain age and other bodily ‘ages’: implications for neuropsychiatry. *Mol. Psychiatry* (2018).

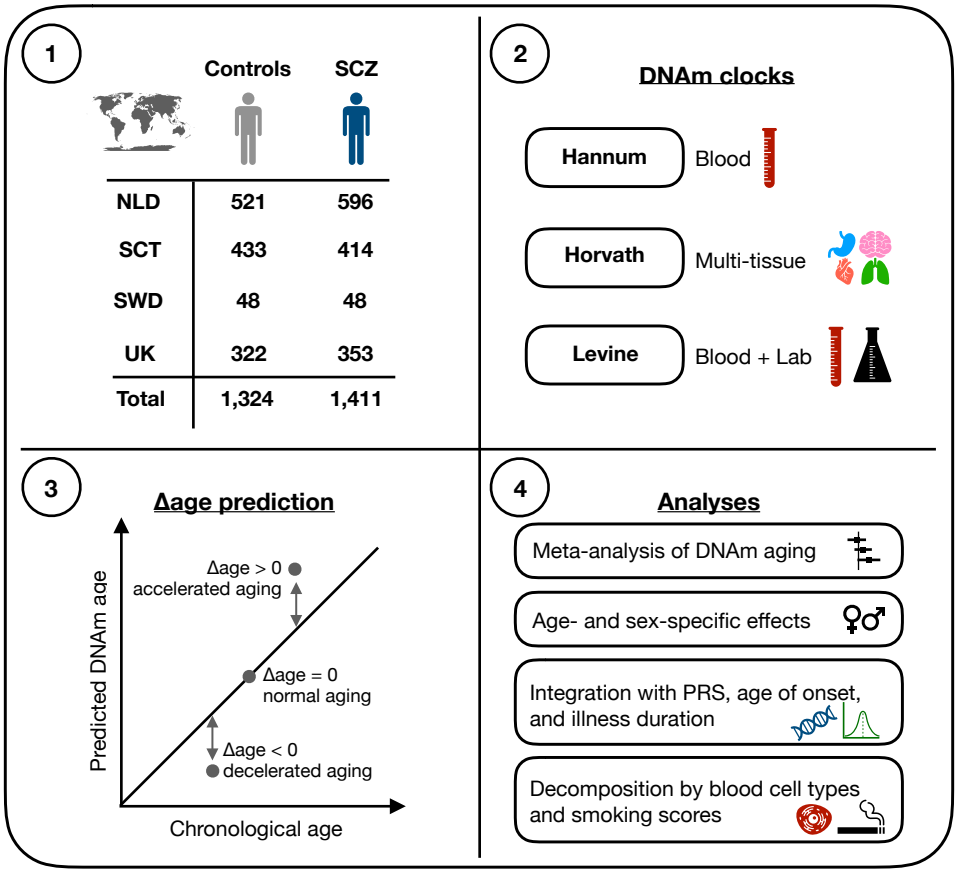

Figure 1. Overview of study design and analysis framework. DNA methylation (DNAm) data was available for a total of 2,735 samples across four European cohorts. See Table S2 for more details on samples. DNAm age estimates were generated using three DNAm clocks, each designed to capture different features of aging (box 2). To investigate differences in aging between cases and controls, Δage was computed (box 3) and analyzed according to the step-wise framework shown in box 4. SCZ = schizophrenia, NLD=Netherlands, SCT=Scotland, SWD=Sweden, UK=United Kingdom, PRS=polygenic risk scores.

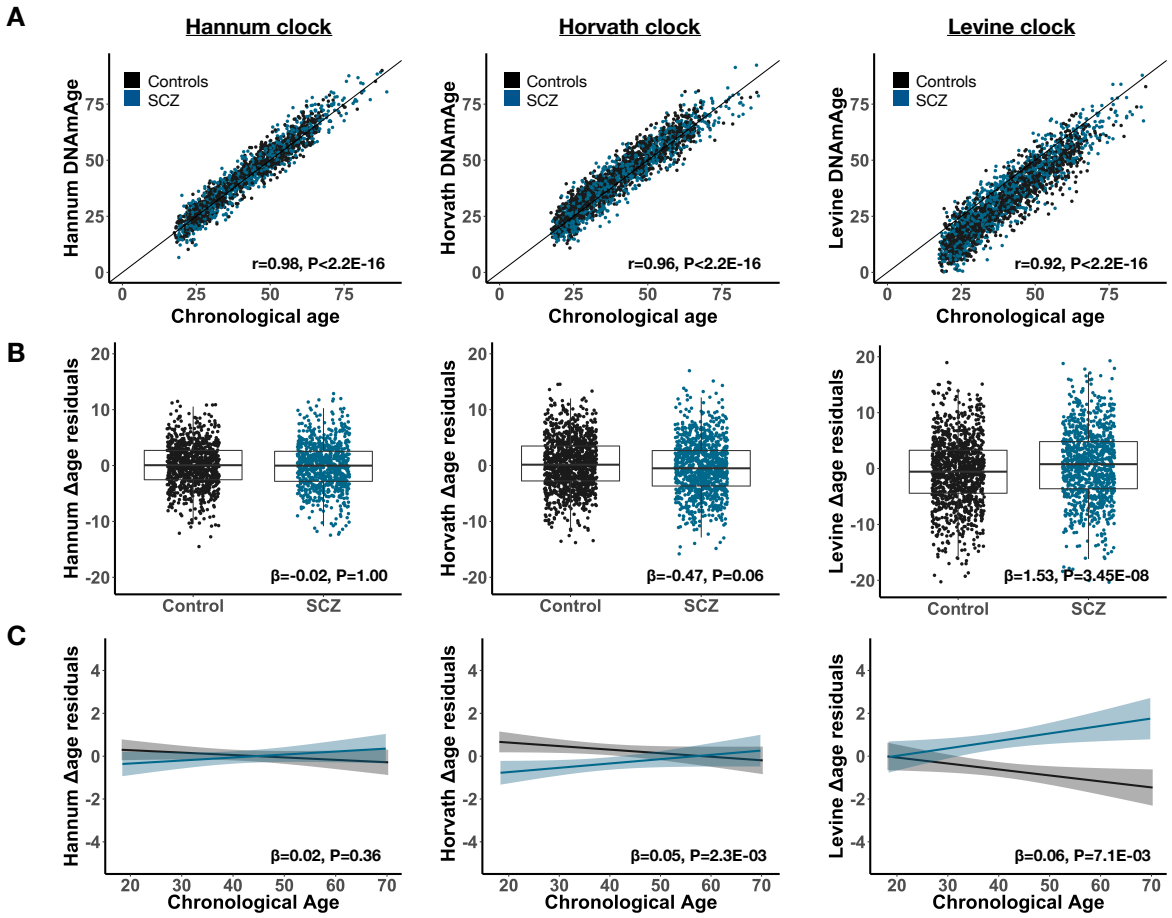

Figure 2. DNA methylation aging is altered in schizophrenia and conditional on chronological age. Presented are results visualizing DNAm aging in SCZ for each clock; Hannum (left), Horvath (middle), Levine (right). Cases are shown in blue and controls in black. (A) The correlation between DNAm age and chronological age. The Pearson's correlation estimate and corresponding p-value are shown in the bottom corner. (B) Boxplots of Δage between cases and controls with the meta-analytic effect size and p-value across cohorts shown. β represents the mean change in Δage in cases compared to controls. (C) Δage is visualized across chronological age with a regression line fitted separately for cases and controls and the meta-analytic interaction effect and p-value shown. β represents the change in Δage in cases per year of chronological age compared to controls. P-values are adjusted for multiple testing across clocks (n=3).

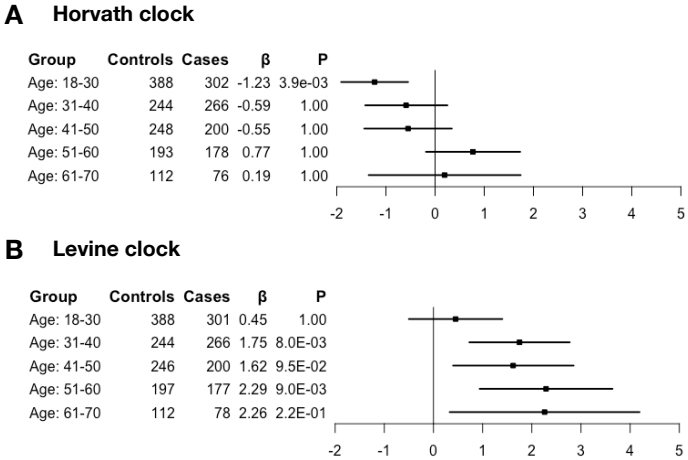

Figure 3. Differential DNAm aging rates across age groups. Shown are Δage differences between cases and controls across age groups for the Horvath (top) and Levine clock (bottom). For each age group, number of cases and controls, and meta-analytic effect size (β) and p-value (P) are presented. P-values are corrected for multiple testing (2 clocks x 5 groups = 10 tests). See Table S5 for more details on results and corresponding statistics.

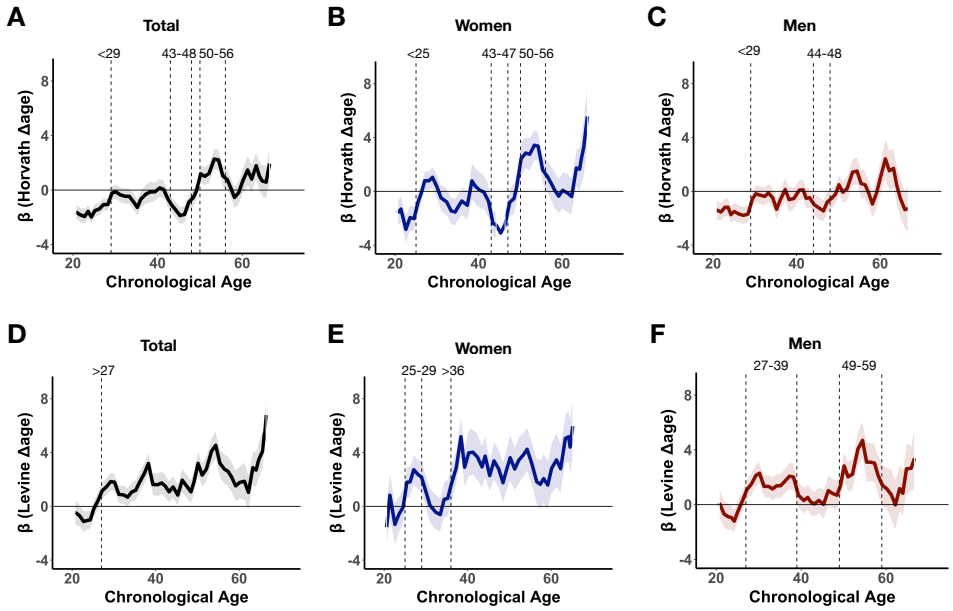

Figure 4. Differential aging in schizophrenia maps to specific age windows. Sliding age-windows, using 5-year bins with steps of 1-year, were used to estimate differential aging (β) at finer resolution across the range of chronological age. Significant shifts in Δage between cases and controls, defined by the standard error of β deviating from zero for at least 3 steps, are highlighted by the dotted vertical lines at their respective ages. Identified age intervals for the Horvath and Levine clock are shown in A-C and D-F, respectively. Results for women (middle) and men (right) are presented in blue and red, respectively. The effects in the total sample are displayed in black (left).

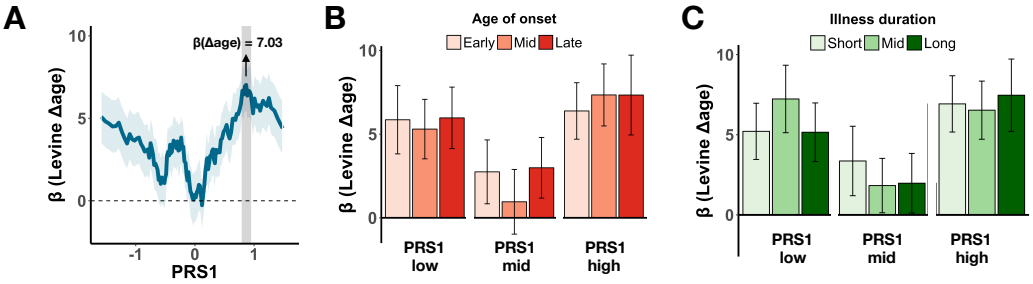

Figure 5. Integration of DNAm aging with PRS, age of presentation, and illness duration across identified age intervals in women >36. (A) Using a sliding-window approach, Levine Δage difference between cases and controls are shown across bins of ranked PRS1. Each bin contains 20 cases and slides from low to high PRS1 per shifts of one sample. The estimated Δage difference compared to all female controls >36 years is shown for each sliding bin in blue with the standard error in shaded blue. The most significant bin is highlighted by the grey vertical bar. (B) DNAm aging effects stratified by PRS1 and age of onset. (C) DNAm aging effects stratified by PRS1 and illness duration.

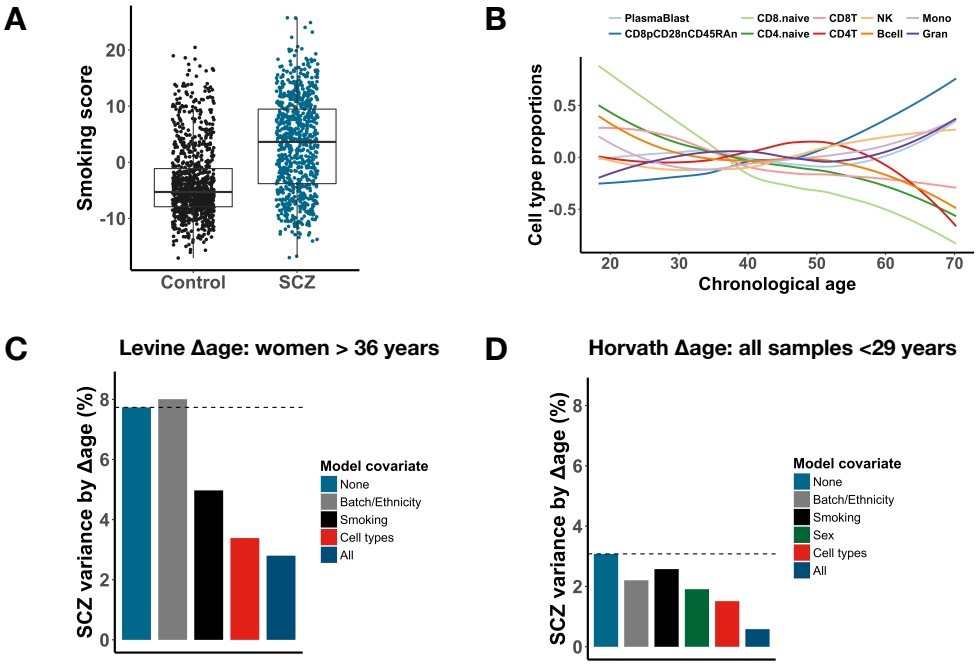

Figure 6. Smoking and blood cell type composition contribute in part to DNAm aging. Presented are the results of a sensitivity analysis of DNAm-based estimated smoking score and blood cell type proportions in the 450K subsample of the cohort. (A) A box plot of the estimated smoking scores between cases and controls. (B) The mean of standardized blood cell type proportions plotted across chronological age in controls only (n=906). (C and D) The proportion of schizophrenia variance explained by Δage after adjustment of various variables that are significantly associated with disease status. The "All" model presents the variance explained by Δage independent from all other variables. The baseline model represents the effect of Δage adjusted for batch, ethnicity and chronological age. Results are shown for Levine Δage women > 36 years, 99 cases and 181 controls (C) and Horvath Δage all samples < 29 years, 141 cases and 238 controls (D) separately.

		Horvath Δage		Levine Δage	
Model variables	**Model comparison**	**Δage R^2**	**P-value**	**Δage R^2**	**P-value**
Model 0: baseline	-	3.6%	-	2.1%	-
Model 1: + status	Model 0 vs 1	4.0%	6.6E-03	3.2%	4.4E-06
Model 2: + status*age.continuous	Model 1 vs 2	4.3%	0.08	3.7%	3.5E-03
Model 3: + status*age.groups	Model 1 vs 3	5.5%	1.4E-05	3.9%	0.02
Model 4: + status*age.groups*sex	Model 3 vs 4	5.9%	0.24	4.7%	0.01

Table 1. Age- and sex-specific effects significantly contribute to DNAm aging in schizophrenia. Shown are the contributions of interaction effects between disease status and age and sex on Δage. The baseline model corresponds to Δage ~ dataset + ethnicity + platform + age.continuous + sex. For other models, the variable(s) in addition to the baseline variables are shown with the corresponding variance explained (R^2) in Δage. Interaction terms with chronological age are modeled as a continuous variable (age.continuous) or a categorical variable (age.groups). The latter uses previously defined decades. Model comparison is performed to assess if the contribution of an interaction term is significant compared to a model without that term. The chi-square test is used to test two models with corresponding p-value presented. The results of these analysis are shown for both the Horvath and Levine clock. P-values are corrected for the number of tests performed (2 clocks x 4 comparisons = 8).

Women: >36	Controls	Cases	Mean value in cases	β (Levine Δage)	95% CI	P
Polygenic risk						
All - no stratification	227	149	0.35	3.02	1.76 — 4.27	3.1E-06
PRS1 - continuous	-	149	0.35	0.42	-0.37 — 1.21	3E-01
PRS1 - low	227	50	-0.68	2.89	1.00 — 4.77	2.8E-03
PRS1 - mid	227	50	0.34	1.89	0.05 — 3.73	4.5E-02
PRS1 - high	227	49	1.40	4.30	2.40— 6.20	**1.3E-05***
Age of onset						
All - no stratification	227	111	26.50	3.73	2.34 — 5.12	2.3E-07
AOO - continuous	-	111	26.5	-0.08	-0.21 — 0.05	2.2E-01
AOO - early	227	37	17.43	3.26	1.11 — 5.41	3.1E-03
AOO - mid	227	37	25.43	3.70	1.58 — 5.81	6.7E-04
AOO - late	227	37	36.62	4.24	2.09 — 6.40	1.3E-04
Illness duration						
All - no stratification	227	111	23.37	3.73	2.34 — 5.12	2.3E-07
DUR - continuous	-	111	23.37	0.03	-0.07 — 0.13	6.1E-01
DUR - short	227	37	10.76	3.90	1.76 — 6.03	3.9E-04
DUR - mid	227	37	23.33	2.91	0.78 — 5.05	7.7E-03
DUR - long	227	37	36.01	4.39	2.25 — 6.53	7.3E-05

Table 2. Integration of Levine Δage with PRS, age of onset, and illness duration in women in later adulthod. Analyses were performed using women >36 years of age. Only cases with available information were included in the analyses. Each phenotype was analyzed as both a continuous variable and as a categorical variable using equal tertiles from low to high bins. Mean values in cases for each phenotype are presented along with the association with Δage (β) and corresponding 95% confidence intervals and p-values. PRS1 = polygenic risk score PC1 (see Supplementary Information) scaled to mean zero with standard deviation of 1, AOO = age of onset, DUR = illness duration. Asterisk* indicates that significance ($P < 0.05$) by permutation analyses.