Genetics of brain age suggest an overlap with common brain disorders

Tobias Kaufmann\(^1\), Dennis van der Meer\(^1\), Nhat Trung Doan\(^1\), Emanuel Schwarz\(^2\), Martina J. Lund\(^1\), Ingrid Agartz\(^3,4\), Dag Alnæs\(^1\), Deanna M. Barch\(^5,6,7\), Ramona Baur-Streube\(^8\), Alessandro Bertolino\(^9,10\), Francesco Bettella\(^1\), Mona K. Beyer\(^11,12\), Erlend Bøen\(^3\), Stefan Borgward\(^13,14\), Christine L. Brandt\(^1\), Jan Buitelaar\(^15,16\), Elisabeth G. Celius\(^17,18\), Simon Cervenka\(^a\), Annette Conzelmann\(^19\), Aldo Córdova-Palomera\(^1,20\), Anders M. Dale\(^21\), Dominique J.-F de Quervain\(^22,23\), Pasquale Di Carlo\(^10\), Srdjan Djurovic\(^24,25\), Erlend S. Dørum\(^1,26,27\), Sarah Eisenacher\(^2\), Torbjørn Elvsåshagen\(^1,11,18\), Thomas Espeseth\(^26\), Helena Fatouros-Bergman\(^4\), Lena Flyckt\(^4\), Barbara Franke\(^28\), Oleksandr Frei\(^1\), Beate Haatveit\(^1,26\), Asta K. Häberg\(^29,30\), Hanne F. Harbo\(^18,11\), Catharina A. Hartman\(^31\), Dirk Heslenfeld\(^32,33\), Pieter J. Hoekstra\(^34\), Einar A. Høgestøl\(^11,18\), Terry Jernigan\(^35,36,37\), Rune Jonassen\(^26\), Erik G. Jönsson\(^14\), Karolinska Schizophrenia Project (KaSP)\(^38\), Peter Kirsch\(^39,40\), Iwona Kłoszewska\(^41\), Knut-Kristian Kolskar\(^1,26,27\), Nils Inge Landrø\(^26,3\), Stephanie Le Hellard\(^25\), Klaus-Peter Lesch\(^42,43,44\), Simon Lovestone\(^45\), Arvid Lundervold\(^46,47\), Astri J. Lundervold\(^48,49\), Luigi A. Maglanoc\(^1,26\), Ulrik F. Malt\(^1,50\), Patrizia Mecocci\(^51\), Ingrid Melle\(^1\), Andreas Meyer-Lindenberg\(^2\), Torgeir Moberget\(^1\), Linn B. Norbom\(^1,26\), Jan Egil Nordvik\(^27\), Lars Nyberg\(^52\), Jaap Oosterlaan\(^32,53,54\), Marco Papalino\(^10\), Andreas Papassotiropoulos\(^55,23,56\), Paul Pauli\(^8\), Giulio Pergola\(^10\), Karin Persson\(^57,58\), Geneviève Richard\(^1,26,27\), Jaroslav Rokicki\(^1,26\), Anne-Marthe Sanders\(^1,26,27\), Geir Selbæk\(^58,59,17\), Alexey A. Shadrin\(^1\), Olav B. Smeland\(^1\), Hilkka Soininen\(^60,61\), Piotr Sowa\(^12\), Vidar M. Steen\(^25,62\), Magda Tsolaki\(^63\), Kristine M. Ulrichsen\(^1,26,27\), Bruno Vellas\(^64\), Lei Wang\(^65\), Eric Westman\(^66,14\), Georg C. Ziegler\(^42\), Mathias Zink\(^2,67\), [alphabetical order ends] Ole A. Andreassen\(^1\), Lars T. Westlye\(^1,26,9\)

[Author byline]: for the Alzheimer's Disease Neuroimaging Initiative\(^68\), for the Pediatric Imaging, Neurocognition and Genetics Study\(^69\), for the AddNeuroMed consortium\(^70\)

1 NORMENT, KG Jebsen Centre for Psychology Research, Division of Mental Health and Addiction, Oslo University Hospital & Institute of Clinical Medicine, University of Oslo, Oslo, Norway
2 Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany
3 Department of Psychiatry, Diakonhjemmet Hospital, Oslo, Norway
4 Department of Clinical Neuroscience, Center for Psychiatry Research, Karolinska Institutet and Stockholm County Council, Stockholm, Sweden
5 Department of Psychological and Brain Sciences, Washington University in St. Louis, St. Louis, USA
6 Department of Psychiatry, Washington University in St. Louis, St. Louis, USA
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

34 Department of Child and Adolescent Psychiatry, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
35 Center for Human Development, University of California, San Diego, USA
36 Department of Cognitive Science, University of California, San Diego, USA
37 Departments of Psychiatry and Radiology, University of California, San Diego, USA
38 Members of Karolinska Schizophrenia Project (KaSP) are listed before References
39 Department of Clinical Psychology, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany
40 Bernstein Center for Computational Neuroscience Heidelberg/Mannheim, Mannheim, Germany
41 Department of Old Age Psychiatry and Psychotic Disorders, Medical University of Lodz, Poland
42 Division of Molecular Psychiatry, Center of Mental Health, University of Würzburg, Würzburg, Germany
43 Laboratory of Psychiatric Neurobiology, Institute of Molecular Medicine, Sechenov First Moscow State Medical University, Moscow, Russia
44 Department of Neuroscience, School for Mental Health and Neuroscience (MHeNS), Maastricht University, Maastricht, The Netherlands
45 Department of Psychiatry, Warneford Hospital, University of Oxford, Oxford, UK
46 Department of Biomedicine, University of Bergen, Norway
47 Mohn Medical Imaging and Visualization Centre, Department of Radiology, Haukeland University Hospital, Bergen, Norway
48 Department of Biological and Medical Psychology, University of Bergen, Norway
49 K. G. Jøbsen Centre for Neuropsychiatric Disorders, University of Bergen, Norway
50 Department of Research and Education, Oslo University Hospital, Oslo, Norway
51 Institute of Gerontology and Geriatrics, University of Perugia, Perugia, Italy
52 Department of Radiation Science, Umeå University, Umeå, Sweden
53 Emma Children's Hospital Amsterdam Medical Center, Amsterdam, The Netherlands
54 VU University Medical Center, Department of Pediatrics, Amsterdam, The Netherlands
55 Division of Molecular Neuroscience, Department of Pediatrics, Amsterdam, The Netherlands
56 Life Sciences Training Facility, Department Biozentrum, University of Basel, Basel, Switzerland
57 Department of Geriatric Medicine, Oslo University Hospital, Oslo, Norway
58 Norwegian National Advisory Unit on Ageing and Health, Vestfold Hospital Trust, Tønsberg, Norway
59 Centre for Old Age Psychiatric Research, Innlandet Hospital Trust, Ottestad, Norway
60 Department of Neurology, Institute of Clinical Medicine, University of Eastern Finland, Kuopio, Finland
61 Neurocenter, Neurology, Kuopio University Hospital, Kuopio, Finland
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

62 Dr. E. Martens Research Group for Biological Psychiatry, Department of Medical Genetics, Haukeland University Hospital, Bergen, Norway.
63 1st Department of Neurology, Aristotle University of Thessaloniki, Makedonia, Greece.
64 INSERM U 1027, University of Toulouse, Toulouse, France
65 Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA
66 Department of Neurobiology, Care Sciences and Society, Karolinska Institute, Stockholm, Sweden
67 District hospital Ansbach, Germany
68 Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.
69 Data used in preparation of this article were obtained from the Pediatric Imaging, Neurocognition and Genetics Study (PING) database (http://ping.chd.ucsd.edu). As such, the investigators within PING contributed to the design and implementation of PING and/or provided data but did not participate in analysis or writing of this report. A complete listing of PING investigators can be found at https://ping-dataset.ucsd.edu/sharing/Authors10222012.pdf
70 AddNeuroMed consortium was led by Simon Lovestone, Bruno Vellas, Patrizia Mecocci, Magda Tsolaki, Iwona Kłoszewska, Hilkka Soininen

* Corresponding authors:

Tobias Kaufmann, Ph.D. & Lars T. Westlye, Ph.D.
Email: tobias.kaufmann@medisin.uio.no, l.t.westlye@psykologi.uio.no
Postal address: OUS, PoBox 4956 Nydalen, 0424 Oslo, Norway
Telephone: +47 23 02 73 50, Fax: +47 23 02 73 33

Counts:
Abstract: 144 words
Main text body: 3147 words
References: 41 in the main paper, 111 incl. supplemental
Figures: 5 (1x 1-column, 4x 2-column)

Key words: Brain age gap, brain disorders, genetic architecture, pleiotropy
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

Numerous genetic and environmental factors contribute to psychiatric disorders and other brain disorders. Common risk factors likely converge on biological pathways regulating the optimization of brain structure and function across the lifespan. Here, using structural magnetic resonance imaging and machine learning, we estimated the gap between brain age and chronological age in 36,891 individuals aged 3 to 96 years, including individuals with different brain disorders. We show that several disorders are associated with accentuated brain aging, with strongest effects in schizophrenia, multiple sclerosis and dementia, and document differential regional patterns of brain age gaps between disorders. In 16,269 healthy adult individuals, we show that brain age gap is heritable with a polygenic architecture overlapping those observed in common brain disorders. Our results identify brain age gap as a genetically modulated trait that offers a window into shared and distinct mechanisms in different brain disorders.

Psychiatric disorders and other brain disorders are among the main contributors to morbidity and disability around the world\textsuperscript{1}, often placing debilitating disadvantages on the individual\textsuperscript{2}. The disease mechanisms are complex, spanning a wide range of genetic and environmental contributing factors\textsuperscript{3}. The inter-individual variability is large, but on a group-level, patients with common brain disorders perform worse on cognitive tests, are less likely to excel professionally, and engage in adverse health behaviours more frequently\textsuperscript{4}.

Dynamic processes influencing the rate of maturation and change throughout the lifespan play a critical role, as reflected in the wide range of disease onset times from early childhood to old age\textsuperscript{5}. This suggests that the age at which individual trajectories diverge from the norm is a key characteristic in the underlying pathophysiology. Whereas autism spectrum disorder (ASD) and attention-deficit/hyperactivity disorder (ADHD) have onset in childhood\textsuperscript{6}, schizophrenia
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

(SZ) and bipolar (BD) spectrum disorders likely evolve during adolescence, before the outbreak of severe symptoms, which is typically in early adulthood\textsuperscript{6,7}. Likewise, multiple sclerosis (MS) most often presents itself in early adulthood but the disease process likely starts much earlier\textsuperscript{8,9}. First episodes in major depressive disorder (MDD) can appear at any stage from adolescence to old age\textsuperscript{6,10}, whereas mild cognitive impairment (MCI) and dementia (DEM) most often evolve in old age\textsuperscript{11}. When attempting to decode the underlying brain dysfunction of these disorders, age-related deviations from the norm may also differ in terms of spatial location, direction, change rate and magnitude, all of which add complexity to the interpretation of observed effects.

Magnetic resonance imaging (MRI) is a powerful tool to unveil abnormal brain development\textsuperscript{12,13} and age-related degeneration\textsuperscript{14,15}. Machine learning techniques enable robust estimation of the biological age of the brain using MRI-derived features\textsuperscript{16}, and initial evidence suggests that a deviation between brain and chronological age – termed the brain age gap - is present in several brain disorders\textsuperscript{17}. These findings may render brain age gap a promising marker of brain health\textsuperscript{17}, but several critical issues remain to be addressed. First, while advantageous for narrowing the complexity, reducing a rich set of brain imaging features into a single estimate of brain age inevitably compromises spatial specificity, thereby potentially removing disorder-specific patterns. Second, most studies so far have been rather small-scale, performed within a limited age range and focusing on a single disorder, which left them unable to uncover clinical specificity and lifespan dynamics. Third, the genetic underpinnings of brain age gap are not understood and it is unknown to what degree they overlap with the genetic architecture of major clinical traits. To address these critical knowledge gaps, large imaging genetics samples covering a range of prevalent brain disorders are necessary.

The availability of unprecedented sample sizes of neuroimaging and genetics data through global data sharing and population-based efforts provide new opportunities for accurate
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

modelling of lifespan differences in brain anatomy and its application to brain disorders. Here, we have gathered raw structural MRI data from a large number of individuals. We employed a centralized and harmonized processing protocol including automated surface-based morphometry and subcortical segmentation using Freesurfer (Suppl. Fig. 1). The main analysis in this study is based on data from 36,891 individuals aged 3 to 96 years that passed quality control, representing the largest brain imaging study on brain age to date.

This sample included data from healthy controls (HC; n = 30,967; 3-95 years), as well as from 5924 individuals with diverse brain disorders with typical onset age distributed across the lifespan. We included data from individuals with ASD (n = 975; 5-64 years) and ADHD (n = 751; 7-62 years), individuals with prodromal SZ or at risk mental state (SZRISK; n = 98; 16-42 years), individuals with SZ (n = 1,145; 18-66 years), a heterogeneous group with mixed diagnoses in the psychosis spectrum (PSYMIX; n = 294; 18-69 years), individuals with BD (n = 445; 18-66 years), MS (n = 254; 19-68 years), MDD (n = 211; 18-71 years), MCI (n = 992; 38-91 years), and DEM (including Alzheimer’s disease (AD); n = 759; 53-96 years). Supplementary Tables 1 and 2 provide details on the sample’s characteristics.

Brain age prediction across brain disorders

We used machine learning to estimate individual biological brain age based on structural brain imaging features. First, we grouped all subjects into different samples. For each of the ten clinical groups, we identified a group of healthy individuals of equal size, matched on age, sex and scanning site using propensity score matching. All remaining individuals were joined into one sample comprising healthy individuals only. The latter constituted a training sample, used to train and tune the machine learning models for age prediction (n = 26,535; 14,182 females and 12,353 males).
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

males, aged 3-89 years), whereas the ten clinical samples were used as independent test samples. Figure 1a illustrates the respective age distributions per sex and diagnosis.

![Figure 1a](image)

Figure 1a illustrates the respective age distributions per sex and diagnosis. The large sample size and wide age-span of the training sample allowed us to model male and female brain age separately, thereby accounting for potential sexual dimorphisms in brain structural lifespan trajectories. For each sex, we built a machine learning model based on gradient tree boosting (xgboost) to predict the age of the brain from a set of thickness, area and volume features extracted using a multi-modal parcellation of the cerebral cortex as well as a set of cerebellar/subcortical volume features (1,118 features in total, Fig. 1b). Five-fold cross-validations confirmed the validity of the models, yielding high correlations between chronological age and predicted brain age (r=.94 and r=.95 for the female and male model,
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders respectively; **Suppl. Fig. 2**. Next, we applied the models to predict brain age for each individual in the ten independent test samples, and tested for effects of diagnosis on the brain age gap. We used mega-analysis (across-site analysis) as the main statistical framework as it may best exploit the benefits of the big data approach, while also providing results from a meta-analysis framework in the supplement. We controlled all associations and group differences reported in this paper for age, age², sex, and scanning site. Further, to rule out confounding effects of data quality on the results, we repeated the main analyses using a more stringent multivariate quality control and exclusion procedure.

**Figure 2** illustrates that the brain age gap was increased in several brain disorders. Strongest effects were observed in SZ (Cohen’s $d = 0.55$), MS ($d = 0.72$), MCI ($d = 0.45$) and DEM ($d = 1.06$). PSYMIX ($d = 0.22$) and BD ($d = 0.30$) showed small effects of increased brain age gap, whereas other groups showed negligible effects. The meta-analysis converged on the same findings (**Suppl. Fig. 3**) and the results replicated regardless of the quality control exclusion criterion applied (**Suppl. Fig. 4**). Compared to matched healthy controls, the average brain age gap was estimated to 1.1 years for ASD, 0.7 years for ADHD, 0.6 years for SZRISK, 3.9 years for SZ, 1.4 years for PSYMIX, 2.0 years for BD, 5.6 years for MS, 0.8 years for MDD, 3.0 years for...
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

...for MCI and 5.8 years for DEM. The brain age gap in all clinical groups was positive and there were no signs of a negative brain age gap (delay) in children with ASD or ADHD.

**Regional specificity of brain age gap**

We assessed the specificity of the spatial brain age gap patterns across clinical groups. We trained age prediction models similar to those for the full brain above, including only occipital, frontal, temporal, parietal, cingulate, insula, or cerebellar/subcortical features (Fig. 1b). Cross-validation confirmed the predictive performance of all regional models (Suppl. Fig. 2), so we used these to predict regional brain age in the ten independent test sets. Region-wise brain age gaps often corresponded to the ones observed on the full brain level, yet some notable differences in the spatial patterns of the disorders emerged (Fig. 3a). For example, increased cerebellar/subcortical age gap was most prominent in DEM ($d = 0.99$) and MS ($d = 0.89$) but was not present in SZ ($d = 0.08$). The largest effect in SZ was observed in the frontal lobe ($d = 0.71$).

In PSYMIX, brain age gaps in the insula ($d = 0.39$) and the temporal lobe ($d = 0.34$) were larger than the brain age gap observed on the full brain level ($d = 0.22$). A brain age gap in the temporal lobe was observed in MDD ($d = 0.26$), whereas there was no evidence for a brain age gap in ASD or ADHD in any of the regions.

**Figure 3b** illustrates a hierarchical clustering of clinical groups based on region-wise effect sizes. One cluster of groups with similar spatial brain age gap patterns comprised MS, MCI and DEM, whereas the other groups formed a second cluster. Notably, the spatial patterns of the groups in the first cluster were negatively associated with several disorders in the second cluster, pointing toward spatial specificity of these disorders.
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

Figure 3: Several disorders displayed regional specific aging patterns. a, Region-wise brain age gaps per disorder. Colours indicate Cohen’s d effect sizes for group comparison to healthy controls matched for age, sex and scanning site. Strongest patterns were observed in the cerebellar/subcortical region for DEM and MS, and in the frontal lobe for SZ. b, Correlation matrix based on the effect sizes from panel (a) indicates similarities (e.g. between MS, DEM and MCI) and dissimilarities (e.g. between SZ/PSYMIX and MS/DEM) in the spatial brain age patterns between the disorders. Sorting is based on hierarchical clustering. c, Effect sizes of region x group interaction effects from repeated measures ANOVAs run for each combination of regions and groups (1260 tests in total). Strongest interaction effects were observed between MS/MCI/DEM and the other disorders, confirming the dissimilarity pattern from panel (b) with an independent analysis.

To explore these differences further, we tested for group x region interactions on each pairwise combination of clinical groups and pairwise combination of region-wise brain age gaps (1260 tests). Figure 3c illustrates the effect sizes for all resulting group x region interactions. Confirming the results from Figure 3b, strongest interaction effects were observed between the groups from cluster 1 and those from cluster 2. For example, the three strongest interaction
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

effects indicated that brain age gaps for frontal and cerebellar/subcortical regions diverged mostly between MS and SZ (Partial Eta squared $\eta_p^2 = 0.16$), MS and PSYMIX ($\eta_p^2 = 0.15$) and between SZ and DEM ($\eta_p^2 = 0.15$; Suppl. Fig. 5 provides effect sizes for all tests). Together, these results demonstrate that several common disorders affecting the brain show anatomically differential patterns of increased brain age gap, indicating that the rate at which different regions age in relation to each other oftentimes showed opposite patterns between disorders typically considered neurodevelopmental and neurodegenerative disorders, respectively.

With converging findings suggesting largest brain age gaps in SZ, MS, MCI and DEM, we explored the functional relevance of the region-wise brain age gaps for these groups, testing for associations with clinical and cognitive data. Clinical data available in the SZ test sample included symptom ($n = 81$ HC, $n = 391$ SZ) and function ($n = 271$ SZ) scores of the Global Assessment of Functioning scale (GAF) as well as positive ($n = 57$ HC, $n = 653$ SZ) and negative ($n = 57$ HC, $n = 655$ SZ) scores of the Positive and Negative Syndrome Scale (PANSS). In the MS test sample, we assessed associations with scores from the Expanded Disability Status Scale (EDSS, $n = 188$ MS) and in the joint MCI and DEM test samples, we assessed associations with Mini Mental State Examination scores (MMSE, $n = 901$ HC, $n = 921$ MCI, $n = 707$ DEM). Figure 4a depicts association strengths for all tests and Figure 4b illustrates the strongest association for each test, except for the PANSS scores where only weak associations were found. In SZ, larger brain age gaps were associated with lower functioning, in particular for full brain, frontal, temporal and insula brain age gaps (GAF function all $z < -0.18$, all $P < 0.003$; GAF symptom all $z < -0.20$, all $P < 2 \times 10^{-5}$). In MS, larger brain age gap was associated with higher disability, in particular for the full brain age gap ($z = 0.23$, $P = .001$). Finally, lower cognitive functioning was associated with larger brain age gaps in the joint
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

MCI/DEM samples, with strongest effects for full brain ($z = -0.34$, $P = 4 \times 10^{-65}$) and cerebellar/subcortical ($z = -0.31$, $P = 2 \times 10^{-53}$) brain age gaps.

**Figure 4. Region-wise brain age gaps were associated with cognitive and clinical scores.** a, Fisher $z$ score for all tested associations. In the SZ test sample, we assessed associations with GAF (function and symptom) and PANSS (positive and negative). In MS, we tested for associations with EDSS and in the joint MCI and DEM test samples, we assessed associations with MMSE scores. b, Exemplary illustration of the strongest associations for GAF, EDSS and MMSE.

The genetic architecture of brain age gap

Given the known genetic contributions to brain disorders, our results pose the question to what degree brain age patterns are genetically constrained and if the implied genes overlap with the polygenic architectures of the disorders. In our cohorts, single nucleotide polymorphism (SNP) data were available for 16,269 adult healthy controls with European ancestry, after disregarding subsets with data from clinical groups, children and individuals with non-European ancestry, all of which were too small to warrant an analysis. We estimated full and region-wise brain age for these individuals using 5-fold cross-validation in a model trained on all healthy controls ($n = 30,967$) and regressed age, age$^2$, sex, and scanning site effects from the resulting brain age gaps.
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

First, we ran genome-wide complex trait analysis (GCTA) across all 16,269 individuals, including the first four population components from multidimensional scaling as covariates to further control for population stratification. The results revealed significant heritability (Fig. 5a), with common SNPs explaining 18.3% of the variance in brain age gap across all individuals (full brain, $h^2_{\text{SNP}} = 0.1828$, SE = 0.02, $P < 1 \times 10^{-16}$) and 11.1-18.4% of the variance in region-wise brain age gaps (all $P < 2 \times 10^{-9}$).

Next, we assessed the overlap between the genetic underpinnings of brain age gap and common brain disorders. Focusing on those disorders that showed a significant brain age gap in the main analysis, we gathered genome-wide association analysis (GWAS) summary statistics for SZ and BD from the Psychiatric Genomics Consortium, MS from the International Multiple Sclerosis Genetics Consortium, and AD from the International Genomics of Alzheimer's Project. In addition, we performed GWAS on the full brain and region-wise brain age gaps in the above-described set of 16,269 healthy controls. We used conditional Q-Q plots to assess polygenic overlap between two complex traits, conditioning GWAS summary statistics from each of the brain age gaps on GWAS summary statistics from each of the disorders. Notably, our results indicate genetic overlap between brain age gap and brain disorders. Figure 5b provides exemplary illustrations of conditional Q-Q plots for the frontal brain age gap stratified by SZ, the cingulate brain age gap stratified by BD and the full brain age gap stratified by MS. When selecting subsets of SNPs based on their associations with the disorders, the nominal -log10 transformed P-values of the brain age gaps deviated from the trajectories expected under the global null hypothesis, indicating that the brain age gaps are enriched for SNP associations with the relevant disorder. SZ and MS also showed patterns of enrichment with subcortical brain age gap and BD with frontal brain age gap, whereas no clear patterns were observed for AD (Suppl. Fig. 6).
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

Next, we combined GWAS summary statistics of brain age gaps and the disorders in conjunctional FDR analyses\textsuperscript{35,36}, to identify SNPs that are associated with both phenotypes. We found 15 independent, significant loci showing pleiotropy between brain age gaps and SZ (2 occipital, 4 frontal, 3 temporal, 1 parietal, 2 cingulate, 1 insula, 2 cerebellar/subcortical; 116 SNPs in total), 6 loci for BD (3 frontal, 2 cingulate, 1 insula; 40 SNPs in total), 7 loci for MS (2 full brain, 2 frontal, 1 temporal, 2 subcortical; 7 SNPs in total) and 1 locus for AD (temporal, 1 SNP), respectively (Suppl. Table 3). An intronic variant (rs940904) in protein coding gene PITPNM2 at chromosome 12q24.31 underlying the frontal brain age gap significantly overlapped both with SZ and MS.

Fig. 5: The brain age gaps are heritable and the genetic underpinnings overlap with those observed for several disorders. a, Heritability (H2) estimated using GCTA (all $P < 2 \times 10^{-9}$). b, Exemplary illustration of genetic enrichment between brain age gaps and SZ, BD and MS, assessed using conditional Q-Q plots. The dashed line is the expected trajectory under the global null hypothesis, whereas the coloured lines are the trajectories observed in the complete set, and in subsets of SNPs identified by their association with the disorder. Abbreviations: BAG, brain age gap. SNP, single nucleotide polymorphism.

Discussion

Taken together, our results provide strong evidence that several common brain disorders are associated with accentuated aging of the brain compared to chronological age, with effects observed in SZ, PSYMIX, BD, MS, MDD, MCI and DEM; but not in ASD, ADHD or SZRISK.
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

Importantly, we revealed a distinct neuroanatomical distribution of brain age gaps in several disorders. Associations with clinical and cognitive data underlined the functional relevance of the brain age gaps and genetic analyses in healthy controls provided evidence that the brain age gaps are heritable, with overlapping genes implicated in the genetic underpinnings of brain age gaps and common brain disorders.

Our approach of estimating brain age at the level of brain regions was useful to reveal differential spatial patterns between disorders. Whereas the implicated regions in the spatial brain age profiles of the disorders matched previously reported structural and functional abnormalities (e.g. frontal in SZ\cite{37-39}, or the widespread volume loss in AD with large effects in subcortical structures\cite{40}), our region-wise brain age approach preserved the well-established benefit of down-sampling a large number of brain imaging features into a highly condensed and interpretable score without a total loss of spatial sensitivity. As such, the analysis revealed substantial differences in spatial aging profiles between disorders typically regarded neurodegenerative disorders (MS, MCI, DEM) and disorders with established neurodevelopmental sources, especially SZ and PSYMIX. Whereas these disorders were all associated with an increased brain age gap on the full brain level, the region-wise analysis uncovered an interaction between the frontal brain age patterns observed in SZ and PSYMIX and the cerebellar/subcortical patterns observed in MS and DEM. Moreover, brain age gaps covered functional relevance beyond the group differences. We identified significant associations with clinical and cognitive data, in particular with scores of the Global Assessment of Functioning scale\cite{25} in SZ, with the Expanded Disability Status Scale\cite{27} in MS and with Mini Mental State Examination scores\cite{28} in the dementia spectrum. These results may warrant further research in which the link between the rate of changes in brain age gaps and the clinical and cognitive outcome can be studied in a longitudinal setting. Depending on the sensitivity of such associations to dynamic changes in clinical and
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

cognitive function, such studies may also explore the biological mechanisms underlying these associations as a potential anchor point for treatment options.

Genetic analysis offers one way of exploring the constraining factors underlying phenotypic variation. Here, we provided evidence that full and region-wise brain age gaps represent genetically influenced traits, and illustrated that the genetic variants associated with brain age gaps are also associated with SZ, BD, MS and AD. In line with the accumulating evidence that common disorders of the brain are highly polygenic and partly overlapping\textsuperscript{30-34,36,41}, these results suggest shared molecular genetic mechanisms between brain age gaps and brain disorders. Statistical associations do not necessarily signify causation, and functional interpretations of the identified genes should be made with caution. Larger imaging genetics samples, in particular those including individuals with common brain disorders, may in the future allow the investigation of specificity of the implicated genes. Considering the observed interaction effects in spatial brain age profiles between some disorders, we speculate that such analyses may offer novel insight into specific molecular mechanisms and will allow us to delineate the processes that affect the pace and profile of global and regional brain aging for each of these disorders.

In conclusion, in this largest brain age study to date, we established that the brain age gap is genetically constrained, increased in several common brain disorders, and linked to clinical and cognitive phenotypes. Our results establish the potential of advanced lifespan modelling in the clinical neurosciences, highlighting the benefit of big data resources that cover a wide span of ages and disease conditions. Delineating dynamic lifespan trajectories within and across individuals will be essential to disentangle the pathophysiological complexity of brain disorders.
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

Acknowledgements

The authors were funded by the Research Council of Norway (213837, 223273, 204966/F20, 229129, 249795/F20, 276082), the South-Eastern Norway Regional Health Authority (2013-123, 2014-097, 2015-073), Stiftelsen Kristian Gerhard Jebsen, and the European Commission 7th Framework Programme (602450, IMAGEMEND). The data used in this study were gathered from various sources. A detailed overview of the included cohorts and acknowledgement of their respective funding sources and cohort-specific details is provided in Supplementary Table 1.

Author contributions

T.K. and L.T.W. conceived the study; T.K., N.T.D. and L.T.W. pre-processed all data in Freesurfer; N.T.D., M.J.L., C.L.B, L.B.N., L.T.W. and T.K. performed quality control of the data; T.K. performed the analysis with contributions from L.T.W. and D.v.d.M.; T.K., L.T.W., N.T.D., D.v.d.M. and O.A.A. contributed to interpretation of the results. All remaining authors were involved in data collection at various sites as well as cohort-specific tasks. T.K. and L.T.W. wrote the first draft of the paper and all authors contributed to and approved the final manuscript.

Competing financial interests

Some authors received educational speaker’s honorarium from Lundbeck (O.A. Andreassen, A. Bertolino, T. Elvsåshagen, M. Zink, N. I. Landrø), Sunovion (O.A. Andreassen), Shire (B. Franke), Medice (B. Franke), Otsuka (A. Bertolino, M. Zink) and Janssen (A. Bertolino), Roche (M. Zink), Ferrer (M. Zink), Trommsdorff (M. Zink), Servier (M. Zink), all of these unrelated to this work. A. Bertolino is a stockholder of Hoffmann-La Roche Ltd and has received consultant fees from Biogen Idec. E. G. Celius and H. F. Harbo have received travel support, honoraria for advice and lecturing from Almirall (Celius), Biogen Idec (both), Genzyme (both), Merck (both), Novartis(both), Roche (both), Sanofi-Aventis (both) and Teva (both). They have received
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

unrestricted research grants from Novartis (Celsius, Harbo), Biogen Idec (Celsius) and Genzyme (Celsius). G. Pergola has been the academic supervisor of a Roche collaboration grant (years 2015-16) that funds his salary. None of the mentioned external parties had any role in the analysis, writing or decision to publish this work. Other authors declare no competing financial interests.

Materials & Correspondence

The data incorporated in this work were gathered from various resources (see acknowledgements). Material requests will need to be placed with individual PIs. Corresponding authors Tobias Kaufmann (tobias.kaufmann@medisin.uio.no) and Lars T. Westlye (l.t.westlye@psykologi.uio.no) will provide additional detail upon correspondence.

Collaborators


1Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet, & Stockholm County Council, Stockholm, Sweden; 2Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden; 3Neuroimmunology Unit, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden; 4NORMENT, KG Jebsen Centre for Psychosis Research, Division of Mental Health and Addiction, University of Oslo, Oslo, Norway; 5Department of Psychiatry Research, Diakonhjemmet Hospital, Oslo, Norway.
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

References

Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

470 **Online methods**

471 **Samples**

472 We have included data collected through collaborations, data sharing platforms, consortia as well as available in-house cohorts. **Supplementary Table 1** and **2** provide detailed information on the individual cohorts. All included cohorts have been published on, and we refer to a list of publications that can be consulted for a more detailed overview of cohort characteristics. Data collection in each cohort was performed with participants’ written informed consent and with approval by the respective local Institutional Review Boards.

478 **Image pre-processing and quality control**

479 Raw T1 data for all study participants were stored and analysed locally at University of Oslo, following a harmonized analysis protocol applied to each individual subject data (**Suppl. Fig. 1**). We performed automated surface-based morphometry and subcortical segmentation using Freesurfer 5.3\(^\text{10}\). We deployed an automated quality control protocol executed within each of the contributing cohorts (**Suppl. Tables 1-2**), that excluded potential outliers based on global data quality measures. In brief, we regressed age, age\(^2\), sex and (in case of multiple scanners) scanning site from mean cortical thickness, cortex volume, subcortical grey matter volume and from estimated total intracranial volume. Next, we z-standardized the resulting absolute of the residuals and excluded those subjects that exceeded a pre-defined standard deviation (SD) threshold of 4 SD. On top of this, a random set of data was carefully screened by trained research personnel to identify segmentation errors, assess the quality of each subject’s brain images manually, edit segmentation where possible and to exclude data of insufficient quality (\(n = 3957\) manually controlled, \(n = 166\) excluded, \(n = 219\) edited). Taken together, the main analysis excluded cases identified by manual QC as well as cases exceeding a threshold of 4 SD in the
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

automated quality control on either of the four brain imaging measures, yielding 36,891 subjects.

In addition, we performed supplementary analyses using a subset of data, where a more stringent quality control and exclusion procedure was applied. Using multivariate outlier detection based on robust methods as implemented in the R package mvoutlier\textsuperscript{24}, we identified an additional 5513 data sets with potentially less sufficient data quality. Thus, supplemental analysis provides a sanity check with those subjects excluded (sample size: \( n = 31,378 \)).

**Brain age prediction**

We utilized the most recent cortical parcellation scheme\textsuperscript{21} to extract cortical thickness, area and volume for 180 regions of interest (ROI) per hemisphere. In addition, we extracted the classic set of cerebellar/subcortical and cortical summary statistics\textsuperscript{19}. This yielded a total set of 1118 structural brain imaging features (360/360/360/38 for cortical thickness/area/volume as well as cerebellar/subcortical and cortical summary statistics, respectively).

We used machine learning on this feature set to predict the age of each individual’s brain. First, we split the available data into a training sample and ten independent test samples, as described in the main text (Fig. 1a). Next, for each sex, we trained machine learning models utilizing the \textit{xgboost} package in R\textsuperscript{42}, chosen due to its resource efficiency and demonstrated superior performance in previous machine learning competitions, to predict the age of the brain using data available in the training set. First, model parameters were tuned using a 5-fold cross-validation of the training data. This step identified the optimal number of model training iterations by assessing the prediction error for 1500 rounds and implementing an early stopping if the performance did not improve for 20 rounds. Based on previous experience, the learning rate was pre-set to \textit{eta}=0.01 and all other parameters were set to default\textsuperscript{42} for linear \textit{xgboost} tree models. After determining the optimal number of training iterations, the full set of training data
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

was used to train the final models with the adjusted nrounds parameter. In addition, to assess overall model performance, prediction models were cross-validated within the training set using a 5-fold cross validation, each fold implementing the above described training procedure and testing on the hold-out part of the training set. Brain age predictions on the level of individual brain regions followed the same procedures as those described for the full brain level, except that the feature set was reduced to cover only those features that overlapped more than 50% with a given lobe. Regions were defined following the FreeSurfer lobesStrict segmentation as occipital, frontal, temporal, parietal, cingulate and insula. In addition, given the limited number of cerebellar features available in the FreeSurfer summary statistics, cerebellar and subcortical features were grouped into a cerebellar/subcortical region (Fig. 1b).

For the genetic analyses, a different approach had to be taken to ensure maximum exploitation of the limited availability of genetic data. Rather than splitting into training and test sets, we selected all healthy subjects and estimated their brain age using a 5-fold cross-validation approach like the one performed when validating performance of the training set. The resulting unbiased estimates of brain age gaps for all individuals with genetic data available went into the genome-wide complex trait analysis and conjunctional FDR.

Main statistical analysis framework

We performed both mega- (across cohorts) and meta- (within cohort) analyses. To estimate group effects on a given measure in a mega-analysis framework, we computed the effect of diagnosis in relation to the healthy controls for each of the ten test samples in a linear model accounting for age, age², sex and scanning site. Cohen’s d effect sizes were estimated based on contrast t-statistics\(^43\) following Formula 1:

\[
d = \frac{i(n_1 + n_2)}{\sqrt{n_1 n_2 \sqrt{df}}}
\] (1)
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

For the meta-analysis, similar models were computed within cohorts. In addition to estimating Cohen’s $d$ (Formula 1), we estimated the variance of $d$ following Formula 2.

$$v = \frac{n_1 + n_2}{n_1 n_2} \cdot \frac{\left( \frac{n_1 + n_2}{2} \right)}{n_1 + n_2 - 2}$$  

Cumulative effects across cohorts were then estimated using a variance-weighted random-effects model as implemented in the `metafor` package in R.

Assessment of regional specificity

The clustering in Figure 3b was performed using heatmap.2 from the `gplots` package in R. A correlation matrix was computed based on the case-control effect sizes obtained from each test sample and region and hierarchical clustering was performed using the default settings. To further explore regional specificity, we performed another analysis that involved only the clinical groups. We regressed age, age$^2$, sex and scanning site from the brain age gaps in each test sample. Next, we joined data from each pair of clinical groups and each pair of regions for repeated measures analysis of variance and estimated the effect sizes of region x group interactions (1260 ANOVAs in total). The interaction effects were visualized in Figure 3c using the `circlize` package in R.

Genetic analyses

We restricted all genetic analyses to individuals with European ancestry, as determined through multidimensional scaling (MDS), and included the first four population components as covariates to further control for population stratification. Single Nucleotide Polymorphism (SNP) data were available for 16,269 adult healthy individuals with European ancestry. We used genome-wide complex trait analysis (GCTA) to estimate the proportion of variance in brain age explained by SNPs. Before the analysis, we removed high LD regions from the genetic data and pruned it, using a sliding window approach with a window size of 50 base pair (bp), a step size of 5 bp and
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

an $r^2$ of 0.2, leaving 133,147 SNPs. All GCTA analyses accounted for age, age$^2$, sex, scanning site and genetic batch.

Furthermore, we used conditional Q-Q plots$^{34}$ and conjunctional FDR analyses$^{35,36}$ to assess polygenic overlap between two complex traits. We gathered genome-wide association analysis (GWAS) summary statistics for SZ and BD from the Psychiatric Genomics Consortium$^{30,31}$, MS from the International Multiple Sclerosis Genetics Consortium$^{32}$, and AD from the International Genomics of Alzheimer's Project$^{33}$; and performed GWAS on the full brain and region-wise brain age gaps in the above-described sample of 16,269 healthy adults. The MHC region was excluded from the analysis. The SNPs were pruned using a pairwise correlation coefficient approximation to LD ($r^2$), where SNPs were disregarded at $r^2<0.2$ and pruning performed with 20 iterations, as described elsewhere$^{34}$. Conjunctional FDR was run for each pair of full brain / region-wise brain age gap and group, using conjunctional FDR threshold of 0.05. SNPs were annotated using the Ensembl Variant Effect Predictor$^{47}$.

Cognitive and clinical associations

Cognitive and clinical associations were tested in subsets based on data availability as described in the main text. First, we regressed age, age$^2$, sex and scanning site from the brain age gaps. Next, we correlated the resulting residuals with scores of the Global Assessment of Functioning scale$^{25}$ (GAF), the Positive and Negative Syndrome Scale$^{26}$ (PANSS), the Expanded Disability Status Scale$^{27}$ (EDSS) and Mini Mental State Examination scores$^{28}$ (MMSE). We transformed the resulting correlations using Fisher’s z transform. Therefore, the reported associations essentially reflect a partial correlation between full brain / region-wise brain age gaps and clinical/cognitive scores, controlling for confounding effects of age, sex and site.

Code availability.
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

The main analysis was performed using R statistics\textsuperscript{48}. The code needed to reproduce the results is available from the authors upon request.

References of the Supplement

42. Tianqi, C., Tong, H., Benesty, M., Khotilovich, V. & Tang, Y. in Xgboost: extreme gradient boosting. R package v0.4-2 (2015).
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

679 Supplementary Figures

1. Data sharing platforms → Structural magnetic resonance images (raw data)

2. Unified pre-processing pipeline (run at University of Oslo)
   - Full brain reconstruction (Freesurfer)
   - Extraction of cortical features from Human Connectome Project atlas and cerebellar/subcortical features
   - Quality control

3. Split data into a training and ten test sets
   
   Training set
   26,535 individuals, 3-89 years

   Test sets
   One test set per clinical group incl. a healthy control group, matched for age, sex and scanning site

4. Train and cross validate age prediction models
   - Female model
   - Male model

5. Predict brain age in each test set
   
   Predicted brain age
   Chronic age
   gap

6. Statistical analysis, assessing group differences in brain age gap
   - Meta analysis (across-) and meta analysis (within-samples)
   - Effect of diagnosis, accounting for age, age\(^2\), sex and scanning site

Suppl. Figure 1: Outline of the main analysis pipeline.
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders.

Suppl. Figure 2: Cross-validation of the prediction models confirmed validity of the models. The correlation between chronological age and predicted brain age estimated using 5-fold cross-validation within the training set is shown for each feature set.
Suppl. Figure 3: Meta-analysis confirmed mega-analysis results. All Cohen’s d effect sizes for the effect of group accounted for age, age² and sex. Further, Cohen’s d for all cohorts that were collected at multiple sites also accounted for scanning site.
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

Suppl. Figure 4: Replication of results in a subset of 31,378 individuals following more stringent, multivariate exclusion criteria. a, Replication of group effects (for comparison, see Fig. 2) b, Replication of spatial brain age gap patterns (for comparison, see Fig. 3a) c, Replication of interaction effect pattern (for comparison, see Fig. 3c). d, Replication of associations with clinical and cognitive scores (for comparison, see Fig. 4).
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

Suppl. Figure 5: Results from 1260 repeated measures ANOVAs confirm group x region interaction effects in brain age patterns. Strongest effects were observed between MS and SZ, MS and PSYMIX as well as DEM and SZ, suggesting divergent aging patterns in these disorders.
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

Suppl. Figure 6: Conditional Q-Q plots for all brain age gaps conditioned for SZ, BD, MS or AD suggests enrichment for several brain age gaps.
Supplementary Tables

Suppl. Table 1: Summary of the included cohorts. Several of the cohorts are under ongoing data collection, thus the subject numbers provided in the reference publications may not match those in this article.

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Source</th>
<th>Comment</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABIDE1</td>
<td><a href="http://fcon_1000.projects.nitrc.org/">http://fcon_1000.projects.nitrc.org/</a></td>
<td>Primary support for the work by Adriana Di Martino was provided by the NIMH (K23MH087770) and the Leon Levy Foundation. Primary support for the work by Michael P. Milham and the INDI team was provided by gifts from Joseph P. Healy and the Starrs Niarchos Foundation to the Child Mind Institute, as well as by an NIMH award to MPM (R03MH096321).</td>
<td>49</td>
</tr>
<tr>
<td>ABIDE2</td>
<td><a href="http://fcon_1000.projects.nitrc.org/">http://fcon_1000.projects.nitrc.org/</a></td>
<td>Primary support for the work by Adriana Di Martino and her team was provided by the National Institute of Mental Health (NIMH 5R21MH107045). Primary support for the work by Michael P. Milham and his team provided by the National Institute of Mental Health (NIMH 5R21MH107045); Nathan S. Kline Institute of Psychiatric Research. Additional Support was provided by gifts from Joseph P. Healy, Phyllis Green and Randolph Cowen to the Child Mind Institute.</td>
<td>50</td>
</tr>
<tr>
<td>ABM</td>
<td>Authors</td>
<td>ABM was supported by the Research Council of Norway (grant number 229133) and Health South East Research Funding Agency (grant number 2015052).</td>
<td>51,52</td>
</tr>
<tr>
<td>ADDNEUROMED</td>
<td>Authors</td>
<td>AddNeuroMed consortium was led by Simon Lovestone, Bruno Vellas, Patrizia Meccoci, Magda Tsoiaki, Iwona Kłoszewska, Hilka Soiminen. Their work was supported by InnoMed (Innovative Medicines in Europe), an integrated project funded by the European Union of the Sixth Framework program priority (FP6-2004-LIFESCIHEALTH-5)</td>
<td>53,54</td>
</tr>
<tr>
<td>ADHD200</td>
<td><a href="http://fcon_1000.projects.nitrc.org/">http://fcon_1000.projects.nitrc.org/</a></td>
<td>F. Xavier Castellanos, David Kennedy, Michael Milham, and Stewart Mostofsky are responsible for the initial conception of the ADHD-200 Consortium. Consortium steering committee includes Jan Buitera, F. Xavier Castellanos, Dan Dickstein, Damien Fair, David Kennedy, Beatriz Luna, Michael Milham (Project Coordinator), Stewart Mostofsky, and Julie Schweitzer. Data aggregation and organization was coordinated by the INDI team, which included Saroja Bangaru, David Gutman, Maarten Mennes, and Michael Milham. Web infrastructure and data storage were coordinated by Robert Bucigrossi, Albert Crowley, Christian Hasselgrove, David Kennedy, Kimberly Pohland, and Nina Preuss. The ADHD-200 Global Competition Coordinators were Damien Fair (Chair of Selection Committee, Editor in Chief for Global Competition Special issue) and Michael Milham.</td>
<td>55,56</td>
</tr>
<tr>
<td>ADHWUE</td>
<td>Authors</td>
<td>Primary support for the study was provided by the German Research Foundation, grant number DFG KFO 125 1/2 and Pa5607-3. KPL and his team are supported by the Deutsche Forschungsgemeinschaft (DFG: CRU 125, CRC TTR 58 A1/A5), European Community (EC: AGGRESSIONTYPE FP7/No. 602805; Fritz Thyssen Foundation (No. 10.13.1185), ERA-Net NEURON/RESPOND, No. 01EW1602B, and S-100 Russian Academy of Science Academic Excellence Project.</td>
<td>57,58</td>
</tr>
<tr>
<td>ADNI1</td>
<td><a href="http://adni.loni.usc.edu/">http://adni.loni.usc.edu/</a></td>
<td>Data used in the preparation of this article were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer’s disease (AD). For up-to-date information, see <a href="http://www.adni-info.org">www.adni-info.org</a>.</td>
<td>49</td>
</tr>
</tbody>
</table>
| ADNI2    | http://adni.loni.usc.edu/                  | Data collection and sharing for this project was funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer’s Association; Alzheimer’s Drug Discovery Foundation; Aanlon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpit Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; }
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

NCNG

Authors

The sample collection was supported by grants from the Bergen Research Foundation and the University of Bergen, the Dr Einar Martens Fund, the K.G. Jebsen Foundation, the Research Council of Norway, to SLH, VMS, AJL, and TE. The authors thank Dr. Eike Wehling for recruiting participants in Bergen, and Professor Jonn-Terje Geitgung and Haraldplass Deaconess Hospital for access to the MRI facility. Additional support by RCN grants 177458/V50 and 231286/F20.

NIMAGE

Authors

This project was supported by grants from National Institutes of Health (grant R01MH062873 to SV Fararone) for initial sample recruitment, and from NWO Large Investment (grant 1750102007010 to JK Buitelaar), NWO Brain & Cognition (grant 433-09-242 to JK Buitelaar), ZonMW Grant 60-60600-97-193, and grants from Radboud University Medical Center, University Medical Center Groningen, Accare, and VU University Amsterdam for subsequent assessment waves. NeuroIMAGE also receives funding from the European Community’s Seventh Framework Programme (FP7/2007 – 2013) under grant agreements n° 602805 (Agressotpe), n° 278948 (TACTICS), and n° 602450 (EMAGEMEND), and from the European Community’s Horizon 2020 Programme (H2020/2014 – 2020) under grant agreements n° 643051 (MiND) and n° 667302 (CoCA).

NORCOG

Authors

The Norwegian register of persons assessed for cognitive symptoms (NorCog) includes clinical, imaging and biological data from memory clinics in Norway (https://www.aldringogelse.no/norkog/). The register is owned by Oslo University Hospital and administered by Norwegian National Advisory Unit on Ageing and Health.

OASIS

http://www.oasis-brains.org/

The study was supported by grants P50 AG05681, P01 AG03991, R01 AG021910, P50 MH071616, U24 RR021382, R01 MH56584.

PING

http://pingstudy.ucsd.edu/

Data used in the preparation of this article were obtained from the Pediatric Imaging, Neurocognition and Genetics (PING) Study database (http://ping.chd.ucsd.edu/). PING was launched in 2009 by the National Institute on Drug Abuse (NIDA) and the Eunice Kennedy Shriver National Institute Of Child Health & Human Development (NICHD) as a 2-year project of the American Recovery and Reinvestment Act. The primary goal of PING has been to create a data resource of highly standardized and carefully curated magnetic resonance imaging (MRI) data, comprehensive genotyping data, and developmental and neuropsychological assessments for a large cohort of developing children aged 3 to 20 years. The scientific aim of the project is, by openly sharing these data, to amplify the power and productivity of investigations of healthy and disordered development in children, and to increase understanding of the origins of variation in neurobehavioral phenotypes. For up-to-date information, see http://ping.chd.ucsd.edu/.

Data collection and sharing for this project was funded by the Pediatric Imaging, Neurocognition and Genetics Study (PING) (National Institutes of Health Grant RC2DA029475). PING is funded by the National Institute on Drug Abuse and the Eunice Kennedy Shriver National Institute of Child Health & Human Development. PING data are disseminated by the PING Coordinating Center at the Center for Human Development, University of California, San Diego.

PNC

https://www.med.upenn.edu

Support for the collection of the data sets was provided by grant RC2MH089983 awarded to R. Gur and RC2MH089924 awarded to H. Hakonarson.

RSI-MN

Authors

Data collection in this MS cohort was supported by the South-Eastern Norway Regional Health Authority project 39569, Research Council of Norway grant 240102 and 240102.

SALD

http://icon_1000.projects.nitrc.org/

Data collection was supported by the South-Eastern Norway Regional Health Authority project 39569, Research Council of Norway grant 240102 and 240102, Oslo MS Society, Older Woman’s Society for MS research. Healthy controls were sampled from the TOP study (same scanner).

SCHIZCONNECT1

http://schizconnect.org/

Data used in preparation of this article were obtained from the SchizConnect (http://schizconnect.org) database. As such, the investigators within SchizConnect contributed to the design and implementation of SchizConnect and/or provided data but did not participate in analysis or writing of this report. Data collection and sharing for this project was funded by NIMH cooperative agreement 1U01 MH097435.

SCHIZCONNECT2

http://schizconnect.org/

SCHIZCONNECT1 comprised BrainGluSch, COBRE and MCIC samples (COINS). SCHIZCONNECT2 comprised NUSDAST and NUNDA samples. Duplicate subjects in different sources were excluded.

The respective samples were supported by the following grants:

BrainGluSch: NIMH R01MH084898-01A1. COBRE: 5P20RR021938 /P20GM103472 from the NIH to Dr. Vince
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

SCORE Authors
This work was supported by the Swiss National Science Foundation (grant No. 119382)

SLIM Authors
Support was provided by grant numbers 31271087; 31470981; 31571137, 31509885, SWU1509383, SWU1509451, csti2015jc10106, 151023, 2015M572423, 2015MS07067, Xm2015037, 14JD88009

STROKE/MOT Authors
Supported by the Research Council of Norway (213837, 223271, 249795/F20, 248778), the South-Eastern Norway Regional Health Authority (2013-123, 2014-097, #2017-112) and Stiftelsen Kristian Gerhard Jebsen.

TOP Authors
The work was funded by the Research Council of Norway (213837, 223271, 2015073, 2016083), and the Norwegian ExtraFoundation for Health and Rehabilitation (2015/FO5146).

UBA Authors
European Community’s Seventh Framework Programme (FP7/2007–2013) grant agreement #602450 (IMAGEMEND)

UKBB Authors
This research has been conducted using the UK Biobank Resource (access code 27412). All subjects with a primary or secondary ICD-10 diagnosis with a mental or neurological disorder were excluded prior to analysis and the remaining subjects included as healthy controls.

UNIBA Authors
This work was supported by a “Capitale Umano ad Alta Qualificazione” grant by Fondazione Con Il Sud awarded to Alessandro Bertolino and by a Hoffmann-La Roche Collaboration Grant awarded to Giulio Pergola. This project has received funding from the European Union Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 602450 (IMAGEMEND). This paper reflects only the author’s views and the European Union is not liable for any use that may be made of the information contained therein.

Suppl. Table 2: Summary of group size, age and sex for each cohort.

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Number of subjects per group</th>
<th>Age in years: mean ± sd</th>
<th>Sex: f/m</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABIDE1</td>
<td>561 (HC) 517 (ASD)</td>
<td>17.1±7.8</td>
<td>6.5-56.2 (HC)</td>
</tr>
<tr>
<td>ABIDE2</td>
<td>498 (HC) 440 (ASD)</td>
<td>15.8±9.9</td>
<td>5.9-64.0 (HC)</td>
</tr>
<tr>
<td>ABM</td>
<td>73 (HC) 192 (MDD)</td>
<td>42.2±13.3</td>
<td>19.0-71.0 (MDD)</td>
</tr>
<tr>
<td>AddNeuroMed</td>
<td>126 (HC) 141 (MCI) 148 (DEM)</td>
<td>73.3±6.5</td>
<td>53.0-87.8 (HC)</td>
</tr>
<tr>
<td>ADHD200</td>
<td>542 (HC) 320 (ADHD)</td>
<td>11.3±3.0</td>
<td>7.0-21.0 (HC)</td>
</tr>
<tr>
<td>ADHDWUE</td>
<td>55 (HC) 59 (ADHD)</td>
<td>10.6±2.5</td>
<td>7.0-17.0 (ADHD)</td>
</tr>
<tr>
<td>ADNI1</td>
<td>218 (HC) 386 (MCI) 187 (DEM)</td>
<td>75.9±5.1</td>
<td>59.9-89.6 (HC)</td>
</tr>
<tr>
<td>ADNI2</td>
<td>268 (HC) 323 (MCI) 139 (DEM)</td>
<td>72.9±6.0</td>
<td>56.2-90.1 (MCI)</td>
</tr>
<tr>
<td>BETULA</td>
<td>347 (HC)</td>
<td>63.0±13.2</td>
<td>25.0-81.3 (HC)</td>
</tr>
<tr>
<td>CAMCAN</td>
<td>648 (HC)</td>
<td>54.2±18.6</td>
<td>18.0-88.0 (HC)</td>
</tr>
<tr>
<td>CIMH</td>
<td>43 (HC) 53 (SZ)</td>
<td>30.6±11.5</td>
<td>18.0-60.0 (HC)</td>
</tr>
<tr>
<td>CNP</td>
<td>125 (HC) 41 (ADHD) 50 (SZ) 49 (BD)</td>
<td>31.5±8.8</td>
<td>21.0-50.0 (HC)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>32.3±10.4</td>
<td>21.0-50.0 (ADHD)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>36.5±8.9</td>
<td>22.0-49.0 (SZ)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>35.4±9.0</td>
<td>21.0-50.0 (BD)</td>
</tr>
</tbody>
</table>
Kaufmann et al., Genetics of brain age suggest an overlap with common brain disorders

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Samples</th>
<th>Mean Age</th>
<th>Standard Deviation</th>
<th>Reference</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>CoRR</td>
<td>1360</td>
<td>26.1±15.9</td>
<td>6.0-88.0</td>
<td>HC</td>
<td>694/666</td>
</tr>
<tr>
<td>DLBS</td>
<td>311</td>
<td>54.3±20.0</td>
<td>20.6-89.1</td>
<td>HC</td>
<td>194/117</td>
</tr>
<tr>
<td>ds000119</td>
<td>73</td>
<td>16.2±4.7</td>
<td>8.1-26.8</td>
<td>HC</td>
<td>43/30</td>
</tr>
<tr>
<td>ds000171</td>
<td>20</td>
<td>29.4±11.3</td>
<td>18.0-59.0</td>
<td>MDD</td>
<td>22/17</td>
</tr>
<tr>
<td>ds000202</td>
<td>94</td>
<td>22.1±2.7</td>
<td>18.0-30.0</td>
<td>HC</td>
<td>94/0</td>
</tr>
<tr>
<td>ds000222</td>
<td>79</td>
<td>44.4±20.1</td>
<td>21.0-73.0</td>
<td>HC</td>
<td>41/38</td>
</tr>
<tr>
<td>HCP</td>
<td>1113</td>
<td>28.8±3.7</td>
<td>22.0-37.0</td>
<td>HC</td>
<td>606/507</td>
</tr>
<tr>
<td>HUBIN</td>
<td>102</td>
<td>42.0±8.8</td>
<td>19.4-56.2</td>
<td>HC</td>
<td>57/139</td>
</tr>
<tr>
<td>HUNT</td>
<td>909</td>
<td>58.8±4.2</td>
<td>50.5-66.8</td>
<td>HC</td>
<td>477/432</td>
</tr>
<tr>
<td>IXI</td>
<td>562</td>
<td>48.6±16.5</td>
<td>20.0-86.3</td>
<td>HC</td>
<td>312/250</td>
</tr>
<tr>
<td>KASP</td>
<td>45</td>
<td>25.7±5.2</td>
<td>19.0-43.0</td>
<td>HC</td>
<td>43/66</td>
</tr>
<tr>
<td>MALTOSLO</td>
<td>44</td>
<td>31.2±9.1</td>
<td>20.0-50.0</td>
<td>BD</td>
<td>58/30</td>
</tr>
<tr>
<td>MOT</td>
<td>57</td>
<td>25.3±3.9</td>
<td>21.0-36.0</td>
<td>HC</td>
<td>33/24</td>
</tr>
<tr>
<td>NCNG</td>
<td>465</td>
<td>53.3±16.5</td>
<td>19.4-82.3</td>
<td>HC</td>
<td>310/155</td>
</tr>
<tr>
<td>NIMAGE</td>
<td>372</td>
<td>16.9±3.8</td>
<td>7.7-28.6</td>
<td>HC</td>
<td>310/393</td>
</tr>
<tr>
<td>NORCOG</td>
<td>312</td>
<td>63.3±10.9</td>
<td>40.0-95.0</td>
<td>HC</td>
<td>351/289</td>
</tr>
<tr>
<td>OASIS</td>
<td>315</td>
<td>65.0±10.2</td>
<td>38.1-85.2</td>
<td>MCI</td>
<td>254/160</td>
</tr>
<tr>
<td>PING</td>
<td>1162</td>
<td>12.0±5.1</td>
<td>3.0-21.0</td>
<td>HC</td>
<td>556/606</td>
</tr>
<tr>
<td>PNC</td>
<td>1466</td>
<td>15.1±3.6</td>
<td>8.2-23.2</td>
<td>HC</td>
<td>775/691</td>
</tr>
<tr>
<td>RSI-MS</td>
<td>254</td>
<td>41.0±10.0</td>
<td>19.0-68.3</td>
<td>MS</td>
<td>196/58</td>
</tr>
<tr>
<td>SALD</td>
<td>493</td>
<td>45.2±17.4</td>
<td>19.0-80.0</td>
<td>HC</td>
<td>307/186</td>
</tr>
<tr>
<td>SCHIZCONNECT1</td>
<td>230</td>
<td>35.2±12.2</td>
<td>18.0-65.0</td>
<td>SZ</td>
<td>117/338</td>
</tr>
<tr>
<td>SCHIZCONNECT2</td>
<td>145</td>
<td>36.3±13.0</td>
<td>18.0-66.0</td>
<td>SZ</td>
<td>120/183</td>
</tr>
<tr>
<td>SCORE</td>
<td>44</td>
<td>32.3±13.1</td>
<td>14.0-68.1</td>
<td>SZ</td>
<td>120/183</td>
</tr>
<tr>
<td>SLIM</td>
<td>556</td>
<td>25.5±4.3</td>
<td>19.0-39.0</td>
<td>SZ</td>
<td>69/133</td>
</tr>
<tr>
<td>STROKEMRI</td>
<td>61</td>
<td>24.3±5.2</td>
<td>18.0-42.0</td>
<td>SZ</td>
<td>24.3±5.2</td>
</tr>
<tr>
<td>TOP</td>
<td>958</td>
<td>27.7±9.3</td>
<td>18.6-48.2</td>
<td>ASD</td>
<td>913/1066</td>
</tr>
<tr>
<td>UBA</td>
<td>1099</td>
<td>22.5±3.4</td>
<td>18.0-35.0</td>
<td>HC</td>
<td>714/385</td>
</tr>
<tr>
<td>UKBB</td>
<td>14581</td>
<td>55.7±7.5</td>
<td>40.2-70.2</td>
<td>HC</td>
<td>7573/7008</td>
</tr>
<tr>
<td>UNIBA</td>
<td>435</td>
<td>26.7±7.7</td>
<td>18.0-63.0</td>
<td>SZ</td>
<td>246/273</td>
</tr>
</tbody>
</table>

*Certified by peer review*
Suppl. Table 3: Significant loci from conjunctional FDR analysis reflecting overlap between brain age gaps and the respective disorders. The gene column reflects the gene closest to the significant SNP as identified via the Ensembl Variable Effect Predictor\textsuperscript{47}, unless the closest gene is more than 5000 bp away in which case no annotation is provided.

<table>
<thead>
<tr>
<th>Brain region, group</th>
<th>Lead SNP</th>
<th>Chr. region</th>
<th>Chr. position</th>
<th>Con. FDR</th>
<th>Consequence</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>occipital, SZ</td>
<td>rs354232</td>
<td>2p16.2</td>
<td>54916910</td>
<td>0.042</td>
<td>Intergenic variant</td>
<td>-</td>
</tr>
<tr>
<td>occipital, SZ</td>
<td>rs994261</td>
<td>2q33.1</td>
<td>199908378</td>
<td>0.033</td>
<td>Intron, non-coding transcript variant</td>
<td>AC018717.1</td>
</tr>
<tr>
<td>frontal, SZ</td>
<td>rs10803691</td>
<td>2p25.2</td>
<td>5449582</td>
<td>0.047</td>
<td>Downstream gene variant</td>
<td>AC073143.1</td>
</tr>
<tr>
<td>frontal, SZ</td>
<td>rs12820906</td>
<td>12q24.31</td>
<td>123493123</td>
<td>0.014</td>
<td>Intron variant</td>
<td>PITPNM2</td>
</tr>
<tr>
<td>frontal, SZ</td>
<td>rs4981695</td>
<td>14q12</td>
<td>29700996</td>
<td>0.028</td>
<td>Intergenic variant</td>
<td>-</td>
</tr>
<tr>
<td>frontal, SZ</td>
<td>rs221923</td>
<td>14q24.2</td>
<td>71584473</td>
<td>0.035</td>
<td>Downstream gene variant</td>
<td>PCNX</td>
</tr>
<tr>
<td>temporal, SZ</td>
<td>rs11892879</td>
<td>2q24.1</td>
<td>156114425</td>
<td>0.048</td>
<td>Intergenic variant</td>
<td>-</td>
</tr>
<tr>
<td>temporal, SZ</td>
<td>rs4689287</td>
<td>16p12</td>
<td>59901110</td>
<td>0.049</td>
<td>Missense variant</td>
<td>C4orf50</td>
</tr>
<tr>
<td>temporal, SZ</td>
<td>rs2764264</td>
<td>6q21</td>
<td>108934461</td>
<td>0.0065</td>
<td>Intron variant</td>
<td>FOXO3</td>
</tr>
<tr>
<td>parietal, SZ</td>
<td>rs9379851</td>
<td>6p22.2</td>
<td>26354780</td>
<td>0.0061</td>
<td>Upstream gene variant</td>
<td>RNU6-1259P</td>
</tr>
<tr>
<td>cingulate, SZ</td>
<td>rs6704641</td>
<td>2q33.1</td>
<td>200164252</td>
<td>0.031</td>
<td>Intron variant</td>
<td>SATB2</td>
</tr>
<tr>
<td>cingulate, SZ</td>
<td>rs7729320</td>
<td>5p13.1</td>
<td>38725149</td>
<td>0.04</td>
<td>Upstream gene variant</td>
<td>RP11-122C.51</td>
</tr>
<tr>
<td>insula, SZ</td>
<td>rs10460398</td>
<td>2q33.1</td>
<td>200214742</td>
<td>0.0061</td>
<td>Intron variant</td>
<td>SATB2</td>
</tr>
<tr>
<td>subcortical, SZ</td>
<td>rs673253</td>
<td>1p34.2</td>
<td>44062154</td>
<td>0.045</td>
<td>Intron variant</td>
<td>PTPRF</td>
</tr>
<tr>
<td>subcortical, SZ</td>
<td>rs13107325</td>
<td>4q24</td>
<td>103188709</td>
<td>0.00026</td>
<td>Missense variant</td>
<td>SLC39A8</td>
</tr>
<tr>
<td>frontal, BD</td>
<td>rs9489193</td>
<td>6q22.1</td>
<td>117769728</td>
<td>0.039</td>
<td>Downstream gene variant</td>
<td>RP1-92C.8.3</td>
</tr>
<tr>
<td>frontal, BD</td>
<td>rs10160070</td>
<td>10q21.1</td>
<td>54201572</td>
<td>0.046</td>
<td>Intergenic variant</td>
<td>-</td>
</tr>
<tr>
<td>frontal, BD</td>
<td>rs1884431</td>
<td>20q11.22</td>
<td>33338585</td>
<td>0.05</td>
<td>Intron variant</td>
<td>NCOA6</td>
</tr>
<tr>
<td>cingulate, BD</td>
<td>rs174583</td>
<td>11q12.2</td>
<td>61609750</td>
<td>0.016</td>
<td>Intron variant</td>
<td>FADS2</td>
</tr>
<tr>
<td>cingulate, BD</td>
<td>rs1034589</td>
<td>22q12.2</td>
<td>31579233</td>
<td>0.013</td>
<td>Intron variant</td>
<td>RNF185</td>
</tr>
<tr>
<td>insula, BD</td>
<td>rs2273684</td>
<td>20q11.22</td>
<td>33529766</td>
<td>0.035</td>
<td>Intron variant</td>
<td>GSS</td>
</tr>
<tr>
<td>full brain, MS</td>
<td>rs4465231</td>
<td>1p36.22</td>
<td>9347278</td>
<td>0.018</td>
<td>Intergenic variant</td>
<td>-</td>
</tr>
<tr>
<td>full brain, MS</td>
<td>rs12146713</td>
<td>12q23.3</td>
<td>106476805</td>
<td>0.02</td>
<td>Intron variant</td>
<td>NUAK1</td>
</tr>
<tr>
<td>frontal, MS</td>
<td>rs11164855</td>
<td>1p22.1</td>
<td>93469690</td>
<td>0.032</td>
<td>Intergenic variant</td>
<td>-</td>
</tr>
<tr>
<td>frontal, MS</td>
<td>rs940904</td>
<td>12q24.31</td>
<td>123491572</td>
<td>0.032</td>
<td>Intron variant</td>
<td>PITPNM2</td>
</tr>
<tr>
<td>temporal, MS</td>
<td>rs10484566</td>
<td>6p21.32</td>
<td>32835258</td>
<td>0.045</td>
<td>Regulatory region variant</td>
<td>-</td>
</tr>
<tr>
<td>subcortical, MS</td>
<td>rs6482190</td>
<td>10p12.31</td>
<td>22037809</td>
<td>0.049</td>
<td>Intergenic variant</td>
<td>-</td>
</tr>
<tr>
<td>subcortical, MS</td>
<td>rs12146713</td>
<td>12q23.3</td>
<td>106476805</td>
<td>0.026</td>
<td>Intron variant</td>
<td>NUAK1</td>
</tr>
<tr>
<td>temporal, AD</td>
<td>rs6485758</td>
<td>11p11.2</td>
<td>47530024</td>
<td>0.046</td>
<td>Intron variant</td>
<td>CELFL1</td>
</tr>
</tbody>
</table>