POPULATION NEUROSCIENCE OF DEVELOPMENT AND AGING

EDITED BY: Tomáš Paus, Stephanie Debette and Sudha Seshadri
PUBLISHED IN: Frontiers in Systems Neuroscience, Frontiers in Genetics, Frontiers in Neuroscience, Frontiers in Neurology, Frontiers in Aging Neuroscience and Frontiers in Human Neuroscience

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88976-154-8
DOI 10.3389/978-2-88976-154-8

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

POPULATION NEUROSCIENCE OF DEVELOPMENT AND AGING

Topic Editors:
Tomáš Paus, Université de Montréal, Canada
Stephanie Debette, Université de Bordeaux, France
Sudha Seshadri, Boston University, United States

Citation: Paus, T., Debette, S., Seshadri, S., eds. (2022). Population Neuroscience of Development and Aging. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88976-154-8

Table of Contents

frontiers | Frontiers in Systems Neuroscience

EDITORIAL
published: 25 April 2022
doi: 10.3389/fnsys.2022.897943

Editorial: Population Neuroscience of Development and Aging

Tomáš Paus[1,2]*, *Stephanie Debette*[3,4] *and Sudha Seshadri*[5]

[1] Departments of Psychiatry and Neuroscience, Faculty of Medicine and Centre Hospitalier Universitaire Sainte-Justine, University of Montreal, Montreal, QC, Canada, [2] Departments of Psychology and Psychiatry, University of Toronto, Toronto, ON, Canada, [3] University of Bordeaux, Inserm, Bordeaux Population Health Research Center, Team VINTAGE, UMR 1219, Bordeaux, France, [4] CHU de Bordeaux, Department of Neurology, Bordeaux, France, [5] Department of Epidemiology and Biostatistics, Glenn Biggs Institute for Alzheimer's and Neurodegenerative Diseases, UT Health San Antonio, San Antonio, TX, United States

Keywords: brain development, brain aging, cohort studies, genetics, epidemiology, neuroimaging, psychiatric and neurologic disorders

Editorial on the Research Topic

Population Neuroscience of Development and Aging

Brain development and aging involve many inter-related biological processes that are shaped by genes and the environment from conception onwards. Population neuroscience endeavors to identify and model such processes and influences using a combination of epidemiology, "omics" sciences and neuroimaging, as applied in large cohorts and meta-analytical datasets. In these efforts, practitioners of population neuroscience are cognizant of three key challenges inherent in their pursuits: (1) An infinite combination of factors influencing the brain from *within* (genes and their regulation) and the *outside* (physical, built and social environment); (2) Presence of *developmental cascades* that carry such influences from one time point to the next, from one organ to another, and from one level of organization to a different one; and (3) Structural and functional *complexity* of the human brain (Paus, 2013, 2016).

This Research Topic collected contributions that speak to the current advancements in relevant methodological and conceptual issues, as well as reviews and original reports drawing on population-based studies of brain development and aging.

On the "development" side of the lifespan continuum, four papers report original findings obtained in four large cohorts of children, adolescents and young adults, namely the Generation R, ABCD Study, IMAGEN and iSHARE. Lopez-Vicente et al. analyzed data acquired during resting state with functional magnetic resonance imaging (MRI) in ∼ 3,000 children (8–15 years of age), and provided a glimpse of developmental changes in the dynamics of "functional connectivity" during this developmental period. Patel et al. studied the relationship between various structural properties of the cerebral cortex in ∼11,000 10-year old children, and reported a striking similarity between inter-regional profiles of the tangential growth of the cerebral cortex (i.e., its surface area) as it relates—in a distinct spatial pattern captured by the profiles—to both the general psychopathology and overall cognitive abilities. Penninck et al. tested, in over 1,500 adolescents, whether psychopathology at 18 years of age could be predicted by inflammation-related variations in brain structure at 14 years of age; they complemented this work in humans by experimental studies in mice. Finally, Tsuchida et al. showed, in a cohort of ∼1,700 university students (18–26 years of age),

Edited and reviewed by:
Hui-Jie Li,
Institute of Psychology (CAS), China

***Correspondence:**
Tomáš Paus
tpausresearch@gmail.com

Received: *16 March 2022*
Accepted: *07 April 2022*
Published: *25 April 2022*

Citation:
Paus T, Debette S and Seshadri S (2022) Editorial: Population Neuroscience of Development and Aging.
Front. Syst. Neurosci. 16:897943.
doi: 10.3389/fnsys.2022.897943

that the maturation of structural properties of white matter, characterized by multi-modal imaging, continue to change in the 3rd decade of life, with considerable variations across MRI-based parameters (e.g., fractional anisotropy, neurite density index) and fiber tracts.

On the "aging" side of the lifespan continuum, this Research Topic includes four papers that address a variety of questions about the aging brain. Lin et al. used a modest lifespan sample (417 individuals, 21–92 years of age), and identified striking age-related changes in cortical thickness and curvature in older (60+ years) adults that varied as a function of their location on the surface (gyri) vs. depth (sulci) of cortical folds. Rodriguez et al. studied 1,400 adults (40–80 years of age) to find out whether the type of employment (and associated mental demands) protect brain health, with an affirmative answer especially in individuals who continued to work in their older age. Sole-Padulles et al. examined, in a multi-center study (∼1,400 adults over 60 years of age), whether or not loneliness is associated with cognitive and brain health, and concluded that this is not the case if one excludes individuals who developed dementia at follow-up assessments. Finally, Xiang et al. analyzed a unique sample of 250 centenarians to assess the interplay between Apolipoprotein E (*APOE*) genotypes and education vis-à-vis cognitive decline, confirming the beneficial effect of education in carriers of the e3 allele but also revealing some additional, less expected, *APOE*-education interactions.

The above empirical reports are complemented by a series of methodological, conceptual and review papers. Thus, Balsor et al. provided a practical guide for studying the molecular development, across the lifespan, of the human visual cortex, addressing the key challenge of limited *post mortem* proteomic and transcriptomic data. Soumaré et al. addressed the important questions of incidental findings on MRI scans acquired in community-based samples of young adults, and provide a breakdown of the observed findings to those that required medical referral, active intervention, and clinical surveillance. Ness and colleagues asked to what extent developmental trajectories in brain structure and function, as obtained in large population-based samples, relate to symptom manifestations, and how to integrate other sources of data (e.g., social environment) to improve relevant predictions and risk calculations (Nees et al.). Abuga et al. provided a systematic review aimed at estimating the magnitude of premature mortality following childhood-onset neurological impairment (e.g., epilepsy), and identified several risk factors associated with premature death in these children in both high-income and low/middle-income countries. Finally, Royse et al. carried out a qualitative review to determine whether population ancestry (African American, Hispanic, non-Hispanic whites) plays a role in dementia-associated pathology and found that there are significant differences as a function of both ancestry and sex in the case of Tau (but not Aβ) load; given the limited literature, no conclusions could be drawn about the possible interaction between ancestry and sex.

The above contributions are but a few papers addressing important questions about the development, maturation and aging of the human brain using the population-neuroscience framework. This type of work has been enabled by several developments in the field, including (1) the pooling of existing datasets facilitated by collaborative efforts of several international consortia, such as CHARGE and ENIGMA, (2) the establishment of large community-based cohorts, such as the UK Biobank and the ABCD Study, and (3) open-science initiatives in the "omics" sciences, especially those in transcriptomics and proteomics (e.g., Allen Human Brain Atlas, the PsychENCODE Project). We and others have used these resources to answer a variety of questions about the development of the human cerebral cortex across the lifespan (Vidal-Pineiro et al., 2020; Bethlehem et al., 2022), its genetic architecture (Grasby et al., 2020; Hofer et al., 2020; Shin et al., 2020), cellular underpinning of inter-regional profiles of MRI-based phenotypic variations (Shin et al., 2018; Patel et al., 2021), and their prenatal origin (Patel et al., 2022). Now that this Research Topic is completed, we will endeavor to facilitate future inter-disciplinary cross-talks by publishing papers of this and similar Research Topics in the Section on Population Neuroimaging of Frontiers in Neuroimaging, a new Frontiers journal launched this year.

AUTHOR CONTRIBUTIONS

TP drafted the manuscript. SD and SS reviewed it. All authors contributed to the article and approved the submitted version.

REFERENCES

Bethlehem, R. A. I., Seidlitz, J., White, S. R., Vogel, J. W., Anderson, K. M., Adamson, C., et al. (2022). Brain charts for the human lifespan. *Nature*. doi: 10.1038/s41586-022-04554-y. [Epub ahead of print].

Grasby, K. L., Jahanshad, N., Painter, J. N., Colodro-Conde, L., Bralten, J., Hibar, D. P., et al. (2020). The genetic architecture of the human cerebral cortex. *Science* 367:aay6690. doi: 10.1126/science.aay6690

Hofer, E., Roshchupkin, G. V., Adams, H. H. H., Knol, M. J., Lin, H., Li, S., et al. (2020). Genetic correlations and genome-wide associations of cortical structure in general population samples of 22,824 adults. *Nat. Commun.* 11:4796. doi: 10.1038/s41467-020-18367-y

Patel, Y., Parker, N., Shin, J., Howard, D., French, L., Thomopoulos, S. I., et al. (2021). Virtual histology of cortical thickness and shared neurobiology in 6 psychiatric disorders. *J. Am. Med. Assoc. Psychiatr.* 78, 47–63. doi: 10.1001/jamapsychiatry.2020.2694

Patel, Y., Shin, J., Abé, C., Agartz, I., Alloza, C., Alnæqs, D., et al. (2022). Virtual ontogeny of cortical growth preceding mental illness. *Biol. Psychiatry* 2:959. doi: 10.1016/j.biopsych.2022.02.959

Paus, T. (2013). *Population Neuroscience*. Berlin Heidelberg: Springer-Verlag.

Paus, T. (2016). Population neuroscience. *Handb. Clin. Neurol.* 138, 17–37. doi: 10.1016/B978-0-12-802973-2.00002-1

Shin, J., French, L., Xu, T., Leonard, G., Perron, M., Pike, G. B., et al. (2018). Cell-specific gene-expression profiles and cortical thickness in the human brain. *Cereb. Cortex* 28, 3267–3277. doi: 10.1093/cercor/bhx197

Shin, J., Ma, S., Hofer, E., Patel, Y., Vosberg, D. E., Tilley, S., et al. (2020). Global and regional development of the human cerebral cortex: molecular architecture and occupational aptitudes. *Cereb. Cortex* 30, 4121–4139. doi: 10.1093/cercor/bhaa035

Vidal-Pineiro, D., Parker, N., Shin, J., French, L., Grydeland, H., Jackowski, A. P., et al. (2020). I. Alzheimer's disease neuroimaging, B. the Australian imaging, and a. lifestyle flagship study of, cellular correlates of cortical thinning throughout the lifespan. *Sci. Rep.* 10:21803. doi: 10.1038/s41598-020-78471-3

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Genetics

ORIGINAL RESEARCH
published: 08 January 2021
doi: 10.3389/fgene.2020.606831

Studying the Interplay Between Apolipoprotein E and Education on Cognitive Decline in Centenarians Using Bayesian Beta Regression

Qingyan Xiang[1], Stacy Lynn Andersen[2], Thomas T. Perls[2] and Paola Sebastiani[3]*

[1] Department of Biostatistics, Boston University School of Public Health, Boston, MA, United States, [2] Section of Geriatrics, Department of Medicine, Boston University School of Medicine, Boston, MA, United States, [3] Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA, United States

Apolipoprotein E (*APOE*) is an important risk factor for cognitive decline and Alzheimer's disease in aging individuals. Among the 3 known alleles of this gene: e2, e3, and e4, the e4 allele is associated with faster cognitive decline and increased risk for Alzheimer's and dementia, while the e2 allele has a positive effect on longevity, and possibly on preservation of cognitive function. Education also has an important effect on cognition and longevity but the interplay between *APOE* and education is not well-characterized. Previous studies of the effect of *APOE* on cognitive decline often used linear regression with the normality assumption, which may not be appropriate for analyzing bounded and skewed neuropsychological test scores. In this paper, we applied Bayesian beta regression to assess the effect of *APOE* alleles on cognitive decline in a cohort of centenarians with longitudinal assessment of their cognitive function. The analysis confirmed the negative association between older age and cognition and the beneficial effect of education that persists even at the extreme of human lifespan in carriers of the e3 allele. In addition, the analysis showed an association between *APOE* and cognition that is modified by education. Surprisingly, an antagonistic interaction existed between higher education and *APOE* alleles, suggesting that education may reduce the positive effect of *APOE* e2 and increase the negative effect of *APOE* e4 at extreme old age.

Keywords: apolipoprotein E, beta regression, blessed information-memory-concentration test, cognitive function, centenarians

OPEN ACCESS

Edited by:
Tomáš Paus,
University of Toronto, Canada

Reviewed by:
Anbupalam Thalamuthu,
University of New South Wales, Australia
M. J. Mosher,
Western Washington University, United States

****Correspondence:***
Qingyan Xiang
qyxiang@bu.edu

Specialty section:
This article was submitted to Applied Genetic Epidemiology, a section of the journal Frontiers in Genetics

Received: *15 September 2020*
Accepted: *30 November 2020*
Published: *08 January 2021*

Citation:
Xiang Q, Andersen SL, Perls TT and Sebastiani P (2021) Studying the Interplay Between Apolipoprotein E and Education on Cognitive Decline in Centenarians Using Bayesian Beta Regression. Front. Genet. 11:606831. doi: 10.3389/fgene.2020.606831

INTRODUCTION

Declines of certain cognitive abilities are common complications of aging and identifying risk factors for cognitive decline is essential to search for therapeutic interventions. Known risk factors for cognitive decline include older age, lower education, and genes such as apolipoprotein E (*APOE*) that plays an important role in the risk for Alzheimer's disease (Fan et al., 2019). *APOE* is involved in the transport of cholesterol and other lipids between cellular structures (Mahley, 1988). The gene has three well-characterized alleles e2, e3, and e4 that are defined by the combinations of the genotypes of the single nucleotide polymorphisms rs7412 and rs429358. The three alleles differ in two amino acids and result in proteins produced by the e2 and e3 alleles that bind to high-density lipoprotein cholesterol (HDL) while e4 binds to very-low-density lipoprotein (VLDL)

(Belloy et al., 2019). Studies have shown that e4-carriers have increased risk for Alzheimer's disease and accelerated cognitive decline compared to non-carriers (Staehelin et al., 1999; Bondi et al., 2003; Wetter et al., 2005; Belloy et al., 2019), while e2-carriers appears to have a reduced risk for age-related neurodegenerative disease (Henderson et al., 1995; Raber et al., 2004; Kim et al., 2017). The review by O'donoghue et al. (2018) lists 40 studies of the association between *APOE* and cognition in longitudinal studies, but none of these studies examined this association among individuals at extreme ages (e.g., centenarians). Investigating the association of this gene with cognitive decline in centenarians is important for informing about the extent of the protective and deleterious effect of this gene on cognitive function at the extreme of human lifespan. In addition, examining the interaction between education and *APOE* alleles could help to better characterize the long term effect of education on cognition and the interplay between genetic and environmental risk factors. Therefore, in this work, we examined the association of e2 and e4 alleles of *APOE* with cognitive function in centenarians enrolled in the New England Centenarian Study (NECS) (Sebastiani and Perls, 2012), who are enriched for carriers of the e2 allele, have varying levels of education, and for whom we have longitudinally collected assessments of cognitive function.

Typically, longitudinal studies of the effect of *APOE* on cognitive decline use linear mixed models of data collected from a variety of neuropsychological tests (Blair et al., 2005; Caselli et al., 2009; Kim et al., 2017). These models rely on the assumption that errors follow a normal distribution, which is unbounded and symmetrically distributed. However, the outcome of any neuropsychological test can easily violate the normality assumption in two aspects: (1) the neuropsychological test scores are usually defined in a limited interval between 0 and a maximum test score, and (2) the distribution of neuropsychological test scores is often skewed. Violating the normality assumption when modeling cognitive test scores could lead to a poor estimation of the genetic effect on cognitive function, and predict scores that are either negative or exceed the maximum test value. To address this problem, we propose using a regression model based on the assumption that the response follows a beta distribution. Beta distribution is defined in a finite interval and can accommodate distribution with different shapes. Hence this distribution is well suited to model test scores and questionnaire outcomes that can take values in a finite range (Smithson and Verkuilen, 2006).

Beta regression was proposed by Ferrari and Cribari-Neto (2004) and was further developed by others (Smithson and Verkuilen, 2006; Simas et al., 2010; Figueroa-Zúniga et al., 2013). Studies in multiple fields have utilized beta regression to model different variables such as ischemic stroke lesion volume (Swearingen et al., 2011), genetic distance (Branscum et al., 2007), and understory vegetation communities (Eskelson et al., 2011). In this paper, we use a Bayesian hierarchical beta regression model to fit longitudinally collected neuropsychological test scores, and we compare the results of this analysis with the conventional linear mixed model. Using the Bayesian beta regression to analyze the association between *APOE* and cognitive decline among centenarians enrolled in the NECS is an important novelty of our analysis (Sebastiani and Perls, 2012).

MATERIALS AND METHODS

Participants

The NECS is an ongoing study that began in 1994 as a population-based study of centenarians living within eight towns in the Boston area and expanded enrollment to include centenarians, their siblings and offspring as well as controls to North America in 2000. Enrolled participants provide socio-demographic data, medical history, and physical function ability (Barthel Index, Mahoney and Barthel, 1965; Sinoff and Ore, 1997). Participants are also administered the Blessed Information-Memory-Concentration (BIMC) test (Blessed et al., 1968; Kawas et al., 1995), which is a brief test of global cognition that can be administered over the phone to the participant or with the help of a proxy. The BIMC has a maximum total score of 37 points. Scores of 33 or greater represent no impairment, 27–32 indicate mild impairment, 21–26 signify moderate impairment, and less than 20 are associated with severe impairment (Kawas et al., 1995). NECS participants are followed annually to collect new medical events, changes in medication and physical function ability, and are administered the BIMC annually. In this analysis, we use the data collected through November 2019. The data include information about sex, education, race, ages, *APOE* genotypes, and BIMC scores of participants who agreed to complete the test. Participants with Alzheimer's disease or dementia who could not complete the test were not included. Genotype data for *APOE* were inferred from the combinations of the SNPs s7412 and rs429358 as described in Sebastiani et al. (2019a).

Statistical Analysis

Beta Regression Specification

A variable y defined in the interval (0, 1) follows the beta distribution if the density function is proportional to

$$f\left(y \mid \mu,\ \phi\right) \propto y^{\mu\phi-1}\left(1-y\right)^{(1-\mu)\phi-1},$$

where μ represents the mean: $0 < \mu < 1$; the function $\mu(1-\mu)/(1+\phi)$ is the variance, and $\phi > 0$ is the precision parameter (Ferrari and Cribari-Neto, 2004). To parameterize the mean as a function of covariates, it is convenient to use the logit function

$$g\left(\mu\right) = log\left(\mu/(1-\mu)\right) = x_i^T\beta,$$

where β is a vector of coefficient and x_i^T is a vector of covariates (Smithson and Verkuilen, 2006; Zeileis et al., 2010). The beta distribution is defined in the open interval (0, 1). To fit this model to data defined in the range (a, b), we use the transformation $y' = (y-a)/(b-a)$. In our analysis, the BIMC scores range from 0 to 37. By adding/subtracting a small correct term from the minimum/maximum values of the BIMC scores, we rescaled the data to the interval [0.01, 0.99] to avoid zeros and ones.

Bayesian Beta Regression Modeling

We were interested in modeling the association of the following main covariates with BIMC scores: age, sex, education, and *APOE* alleles. We standardized the continuous variables *age* and *education* to generate parameters on the same scale and coded the dichotomous variable *sex* 0 for female and 1 for male. Since homozygote carriers of e2 or e4 are rare, we used the following allele groupings in the analyses:

1. "e2 group" comprising carriers of the *APOE* genotypes e2e2 or e2e3;
2. "e3 group" comprising carriers of the *APOE* genotype e3e3;
3. "e4 group" comprising carriers of the *APOE* genotypes e3e4 or e4e4.

We used dummy codings regarding these three groups, and we selected "e3 group" as the reference group in all of our analyses because it is the most frequent genotype in Whites. We performed two analyses to distinguish the beneficial effect of the e2 group allele relative to e3 group from lack of carrying the deleterious e4 allele. One analysis estimated the association between *APOE* and BIMC scores of the e2 group relative to the e3 group by including only the e2 groups and the e3 groups. The other analysis estimated the association between *APOE* and BIMC scores of the e4 group relative to the e3 group. For completeness, we also performed an additional analysis that included all three *APOE* groups.

We used backward selection for model fitting, and started with the hierarchical regression model:

$$y_{ij} \sim \text{Beta}(\mu_{ij}\phi, \phi(1-\mu_{ij}))$$

$$\text{logit}\left(\mu_{ij}\right) = \mu_{b_0} \cdot I_{single} + b_{i0} \cdot \left(1 - I_{single}\right) + \beta_{age} \cdot age_{ij} +$$

$$\beta_{age^2} \cdot age_{ij}^2 + \beta_{sex} \cdot sex_i + \beta_{edu} \cdot edu_i + \beta_{APOE} \cdot APOE_i +$$

$$\beta_{age.edu} \cdot age_i \cdot edu_i + \beta_{age.APOE} \cdot age_i \cdot APOE_i +$$

$$\beta_{edu.APOE} \cdot edu_i \cdot APOE_i,$$

where y_{ij} denotes the j^{th} cognitive test score of the i^{th} participant, the β coefficients are fixed effects, and $b_{i0}b_{i0}$ is the random intercept that we used to account for within participant correlation of the repeated measurements. Besides the covariates of the main effects of age, sex, education, and *APOE*, we also included a squared term of age and two-way interactions of these main effects in our model. We used a piecewise random intercept $\mu_{b0} \cdot I_{single} + b_{i0} \cdot \left(1 - I_{single}\right)$ to accommodate participants with different number of test administration, where the indicator variable I_{single} is 0 for participants with only one cognitive test score and $I_{single} = 1$ otherwise. The random intercepts b_{i0} were assumed to be independent and normally distributed, i.e., $b_{i0} \sim N(\mu_b, \sigma^2{}_b)$. We specified a normal prior for the mean parameter μ_b of the random intercept that $\mu_b \sim N(0,\ 1000)$, and a gamma distribution with the shape and scale parameters both equal to 1 for the precision parameter ϕ and the precision of random intercept $1/\sigma^2{}_b$. With this parameterization, the participants with only one test score would be assigned the fixed intercept μ_b in the regression, while participants with more than one test score have their own random intercept b_{i0}. We used normal priors for all the fixed effect parameters.

We implemented the backward model selection algorithm using the deviance information criterion (DIC) (Spiegelhalter et al., 2002, 2014), which is particularly useful for selection of hierarchical models. Since DIC has a tendency to overfit (Clarke and Clarke, 2018), we then refined the model selected by this search by retaining only interactions and main effects with a posterior credible interval (CI) that did not include 0. Once we selected the final model, we also conducted a sensitivity analysis with respect to the prior distributions. For regression coefficients and the mean parameter of the random intercept that use the normal priors, we reduced the variance from 1,000 to 10 and 100. For precision parameters of the beta distribution and the random intercept that use the gamma priors, we modified the variance of gamma priors from 1 to 100. We ran each case to assess if the parameter estimates altered after we changed the scale of the prior parameters. We estimated the explained variance of the model using the regression sum of squares divided by the total sum of squares. All analyses were conducted in R3.6 and all Bayesian models were analyzed using Markov Chain Monte Carlo (MCMC) implemented in the "rjags" package (Plummer, 2016). The posterior estimates of the parameters are derived from at least 8,000 burn-in adaptions and 4,000 iterations.

Interpretation of the Results

A limitation of beta regression is that the magnitude of the regression coefficients is not directly interpretable in terms of changes of the outcome. To better understand the association of *APOE* with BIMC scores and the interplay with education, we estimated the fitted means (marginal effect) using inverse transformation of the logit function, and then we rescaled the fitted means to the original scale of the BIMC scores (0, 37). We calculated fitted means for the e2, e3, and e4 groups for three different education levels: low education (25% quantile of study population's education; 8 years), median education (median of study population's education; 12 years), and high education (75% quantile of study population's education; 15 years). For each allele in each education level, we also calculated the corresponding age of onset of moderate cognitive impairment (BIMC score = 26) using the fitted trajectories of the cognitive test scores. These ages provide a quantitatively more interpretable metric of the genetic effects.

Comparison With the Conventional Method

For comparison with the conventional method, we performed two analyses that focus on the effect of e2 or e4 separately using linear mixed models. For a fair comparison, we used the same variables as selected by Bayesian beta regression, and we kept the random intercept term to adjust for the subject effect. The linear mixed model was conducted in R3.6 using the package "lme4." We visualized the results of these two methods by plotting the fitted BIMC scores (marginal effect) with their credible/confidence intervals against age. In the Bayesian beta regression, the credible interval of the fitted value was computed

using the 2.5th and 97.5 percentiles from the MCMC samples. In the linear mixed model, the confidence interval of the fitted values was derived using bootstrap. We also computed a residual sum of squares (RSS) for each method.

RESULTS

Applying Bayesian Beta Regression on the NECS

Out of 768 total participants in the NECS dataset with *APOE* genotypes, we excluded 167 participants with missing test scores, 111 participants with missing education information, and 4 participants with *APOE* genotype e2e4. **Table 1** summarizes the demographic characteristics and test scores at baseline of the 486 participants used in our analysis. The age range of the data used in our analysis was 91–113 years. The *APOE* e3 group was the most prevalent and used as the reference group in all subsequent analyses. Approximately 24% of centenarians in the study carried at least one e2 allele, while only 8% were carriers of an e4 allele. We did not find any e4 homozygous centenarians.

The histogram in **Figure 1** shows that the distribution of the baseline BIMC scores is bounded and highly skewed. Hence we analyzed the rescaled test scores using beta regression. The combination of backward selection using the DIC criterion and refinement of the model using the posterior credible intervals produced the models that are summarized in **Tables 2**, **3**. The diagnostic plots (trace plots, autocorrelation plots, and the Gelman plots of all MCMC chains) showed no indications of lack of convergence. The sensitivity analyses also showed that changing the prior distributions did not appreciably alter the parameter estimates. The diagnostic plots and sensitivity analyses are all included in **Supplementary Material**.

Table 2 summarizes the results of the analysis comparing *APOE* e2 with *APOE* e3, and **Table 3** summarizes the results of the analysis comparing *APOE* e4 with *APOE* e3. The significant coefficients are marked with bold fonts. In both analyses, age, sex, and education were all significantly associated with BIMC scores. Older age was associated with significantly worse performance on the test. More years of education and male sex were associated with significantly higher scores.

In the analysis comparing e2 with e3, the interaction between education and e2 was significant, thus suggesting an effect modification of education on the association of e2 with BIMC scores. It is noteworthy that the estimates of the main effect of education (0.063) and the interaction term (−0.062) were opposite, so that the negative interaction between e2 and education numerically canceled out the positive effect of education on BIMC scores in carriers of the e2 allele. This is shown in **Figure 2**, in which the fitted BIMC scores in e2 carriers remain nearly the same in different education levels. **Figure 2** also shows the positive effect of education on BIMC scores in the e3 group. Interestingly, only participants of the e3 group with low education had a worse BIMC score compared to carriers of one or more e2 alleles. The difference between the e2 and e3 groups became negligible in participants with median education level, and carriers of the e3e3 genotypes with a high level of education obtained a better score than carriers of the e2 alleles.

To better summarize the clinical implication of these results, we calculated the ages of onset of cognitive impairment predicted by the fitted model. From the fitted lines, we estimated that the ages of onset of moderate cognitive impairment were 99.5, 101.8, and 103.5 years in for e3 carriers with low, median, and high education groups, respectively. The age of onset of moderate cognitive impairment (BIMC score = 26) was 102.1 years in e2 carriers, independent of education. Therefore, e2

TABLE 1 | Demographic characteristics and the BIMC scores at baseline of the participants in the New England Centenarian Study.

	***APOE* e2 (e2e2, e2e3)**	***APOE* e3 (e3e3)**	***APOE* e4 (e3e4)**
N at enrollment	117	331	38
N with at least 2 follow-ups	60	170	20
Sex, male (%)	22 (18.8%)	85 (25.68%)	9 (23.68%)
Age at enrollment			
91–95 (%)	4 (3.42%)	10 (3.02%)	2 (5.26%)
96–100 (%)	22 (18.80%)	55 (16.61%)	9 (23.68%)
101–105 (%)	30 (25.64%)	94 (28.40%)	10 (26.32%)
106–110 (%)	50 (42.74%)	152 (45.92%)	16 (42.10%)
111-113 (%)	11 (9.40%)	20 (6.04%)	1 (2.63%)
Mean (SD)	103.54 ± 4.59	103.32 ± 4.43	102.42 ± 4.97
Years of Education, mean (SD)	11.63 ± 3.51	11.82 ± 3.85	12.29 ± 4.05
BIMC scores at baseline			
33–37 (%)	29 (24.79%)	98 (29.61%)	12 (31.58%)
27–32 (%)	33 (28.21%)	66 (19.94%)	7 (18.42%)
21–26 (%)	30 (25.64%)	91 (27.49%)	8 (21.05%)
0–20 (%)	25 (21.37%)	76 (22.96%)	11 (28.95%)
Mean (SD)	25.11 ± 8.23	25.00 ± 8.63	24.38 ± 10.20

The percentiles are with respected to each column stratum. The groups of the BIMC scores are used to classify the degree of cognitive impairment: normal, 33–77; mild, 27–32; moderate, 21–26; severe, 0–21.

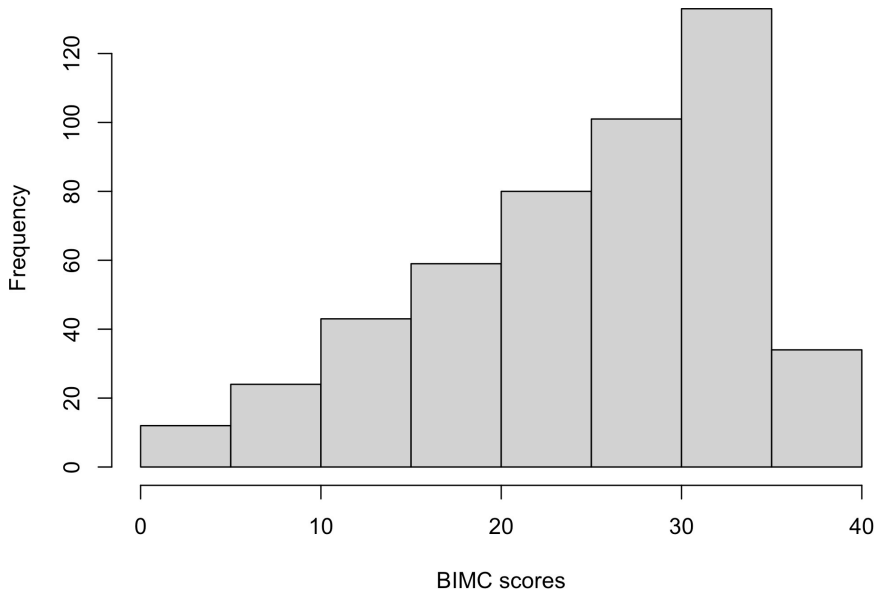

FIGURE 1 | Histogram of baseline Blessed Information-Memory-Concentration scores from the New England Centenarian Study.

TABLE 2 | Parameter estimates and 95% credible intervals from the analysis of the BIMC scores in carriers of *APOE* e2 and e3 alleles using Bayesian beta regression.

	Estimates	2.5% CI	97.5% CI	SD
Intercept (μ_b)	**0.592**	**0.489**	**0.695**	**0.053**
Age	**−0.110**	**−0.138**	**−0.099**	**0.009**
Sex, male	**0.262**	**0.076**	**0.447**	**0.096**
Education	**0.062**	**0.039**	**0.085**	**0.012**
*APOE*2	0.037	−0.162	0.186	0.088
*APOE*2*edu	**−0.063**	**−0.109**	**−0.019**	**0.023**

All covariates were standardized and the score was rescaled to the interval (0, 1) to fit the beta regression, so all parameters should be interpreted on the logit scale and the effects are relative to the rescaled scores. Bold fonts highlight significant parameters.

TABLE 3 | Parameter estimates and 95% credible intervals from the analysis of the BIMC scores in carriers of *APOE* e4 and e3 alleles using Bayesian beta regression.

	Estimates	2.5% CI	97.5% CI	SD
Intercept (μ_b)	**0.647**	**0.541**	**0.685**	**0.053**
Age	**−0.119**	**−0.141**	**−0.098**	**0.011**
Sex, male	**0.215**	**0.019**	**0.413**	**0.101**
Education	**0.065**	**0.041**	**0.089**	**0.012**
*APOE*4	**−0.382**	**−0.664**	**−0.103**	**0.139**
*APOE*4*edu	**−0.082**	**−0.159**	**−0.008**	**0.034**

All covariates were standardized and the score was rescaled to the interval (0, 1) to fit the Bayesian beta regression, so all parameters should be interpreted on the logit scale and the effects are relative to the rescaled scores. Bold fonts highlight significant parameters.

carriers were estimated to delay the onset of moderate cognitive impairment by approximately 2 years compared to e3 carriers with low education, but this advantage essentially disappeared with higher education.

In the analysis comparing e4 with e3, both the main effect term of e4 and the interaction with education were significantly negative, thus suggesting that higher education was not sufficient to remove the negative association of the e4 allele with BIMC scores. This is illustrated in **Figure 3** that shows the fitted BIMC scores in the e3 and e4 groups stratified by education. Among individuals with the e3e3 genotype, more years of education was associated with higher BIMC scores, but the positive effect modification of higher education was reduced in carriers of e4 compared to e3.

From the fitted lines, the ages of onset of moderate cognitive impairment in e4 carriers were 99.5, 98.9, and 98.57 years in participants with low, median, and high education, respectively. Hence, in centenarians carrying the e4 allele, more years of education was not associated with a delay of cognitive decline.

The results of the combined analysis including all three groups of *APOE* alleles are summarized in **Supplementary Table 1**. After using the same model selection strategy as the previous separated analyses, the selected model had exactly the same variables of the two separated analyses. The direction of the effects of *APOE* and the significance of the variables were consistent with the previous analyses. The major difference was that the magnitude of the effects of *APOE* decreased in the combined analysis. For example, in the separated analysis, the beta coefficients of e2 and e4 were 0.032 and −0.382, respectively, while in the combined analysis the beta coefficients of e2 and e4 were 0.026 and −0.326. Hence, the analysis that combined all *APOE* alleles together suggested slightly smaller but still significant effects of *APOE*.

Comparison With Conventional Linear Mixed Models

The results of the analyses using linear mixed models are summarized in **Tables 4**, **5**. In both analyses, only the effects of age

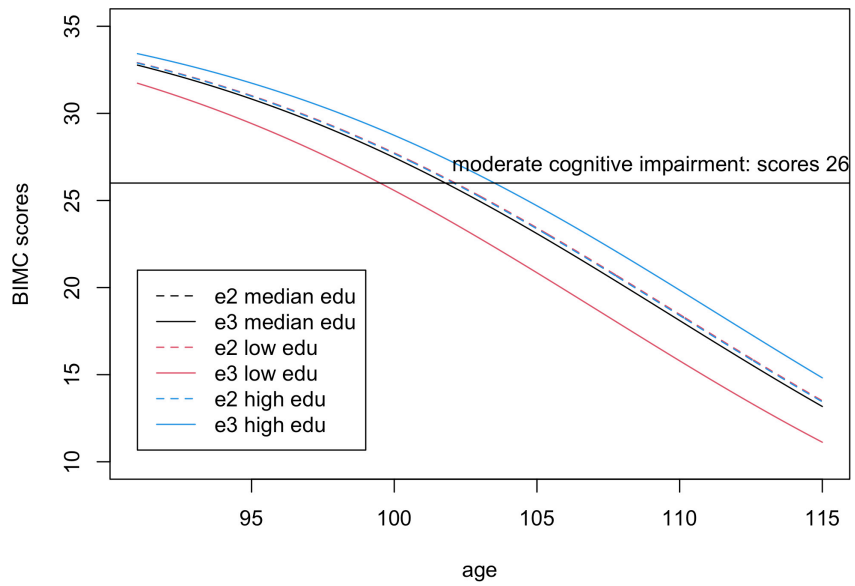

FIGURE 2 | Fitted Blessed Information-Memory-Concentration scores in carriers of one or more e2 alleles (dashed lines) compared to homozygotes e3e3 (continuous lines) for low education level (25% quantile, 8 years, red), median education level (12 years, black), and high education level (75% quantile, 15 years, blue). The scores were fitted using the estimates from Bayesian beta regression, assuming sex = female. Note visually there is only one dashed line, because the fitted lines of e2 carriers overlap in all three education groups.

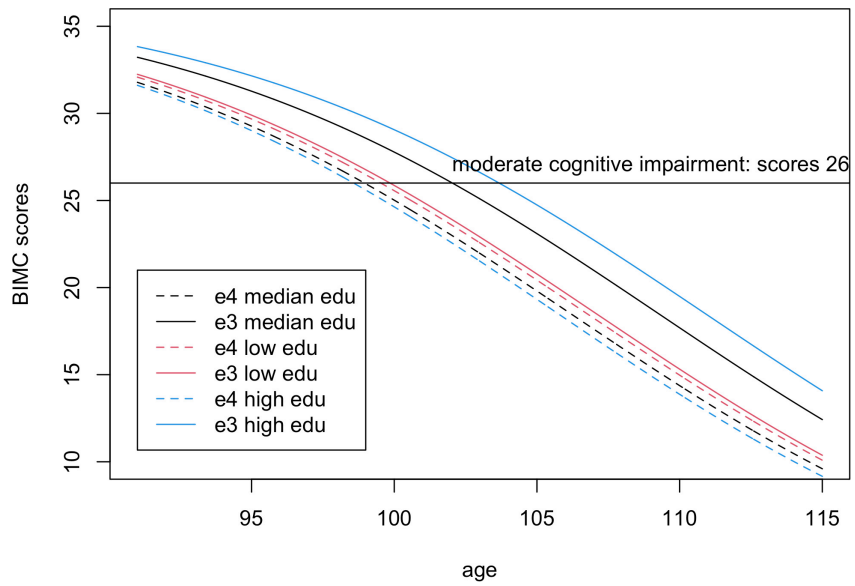

FIGURE 3 | Fitted Blessed Information-Memory-Concentration test scores in carriers of one or more e4 alleles (dashed lines) compared to homozygotes e3e3 (continuous lines) for low education level (25% quantile, 8 years, red), median education level (12 years, black), and high education level (75% quantile, 15 years, blue). The scores were fitted using the estimates from Bayesian beta regression, assuming sex = female.

and education reached statistical significance. The direction of the main effects of e2 and e4 were consistent with those estimated using the Bayesian beta regression. However, the effects did not reach statistical significance.

To visualize the differences of the results from Bayesian beta regression and the linear mixed model, we plotted the fitted BIMC scores with their 95% credible/confidence intervals against age in **Figure 4**. The plot shows that the credible intervals of the Bayesian beta regression were slightly narrower compared with the confidence intervals of the linear mixed model. The linear mixed model did not fit the data of the youngest age group well, and the fitted values as well as their confidence

TABLE 4 | Parameter estimates from the analysis of the BIMC scores in carriers of *APOE* e2 and e3 alleles using the linear mixed model.

	Estimates	SD	*T*-value	*P*-value
Intercept	**23.935**	**0.474**	**50.544**	**< 0.001**
Age	**−1.074**	**0.064**	**−16.844**	**< 0.001**
Sex, male	1.251	0.859	1.457	0.146
Education	**0.438**	**0.109**	**4.032**	**< 0.001**
APOE2	0.075	0.82 0	0.092	0.927
*APOE2**edu	−0.430	0.229	−1.876	0.061

Bold fonts highlight significant parameters.

TABLE 5 | Parameter estimates from the analysis of the BIMC scores in carriers of *APOE* e4 and e3 alleles using the linear mixed model.

	Estimates	SD	*T*-value	*P*-value
Intercept	**24.386**	**0.495**	**49.233**	**< 0.001**
Age	**−1.082**	**0.071**	**−15.174**	**< 0.001**
Sex, male	1.073	0.940	1.142	0.254
Education	**0.437**	**0.111**	**3.934**	**< 0.001**
APOE4	−2.286	1.330	−1.719	0.087
*APOE4**edu	−0.485	0.333	−1.458	0.146

Bold fonts highlight significant parameters.

interval were even greater than the maximum BIMC score (max score = 37). This result is consistent with our hypothesis that using a linear mixed model with normality assumptions is not suitable when the outcomes are bounded neuropsychological scores. In addition, the Bayesian beta regression appeared to better capture the gradually increasing rate of decline, but the linear mixed model did not capture this feature. The RSS of the Bayesian beta regression was 51032.5 for the e2 analysis and 43151.54 for the e4 analysis, while the RSS of the linear mixed models was slightly larger (52443.72 for the e2 analysis and 43652.30 for the e4 analysis), thus confirming that the Bayesian beta regression fit the data better.

DISCUSSION

In this paper, we investigated the relationship between *APOE* alleles and change of cognitive function in a large cohort of centenarians enriched for carriers of the e2 allele. Instead of using the conventional linear mixed model, we used a Bayesian hierarchical beta regression model to characterize the effect of *APOE* on cognitive function assessed through the Blessed Information-Memory-Concentration test. Our analyses confirmed the decline of cognitive function and the positive effect of education on preservation of cognitive function at extreme old age. The analysis also showed that the *APOE* e4 allele has a negative association with cognitive function even at extreme old age, and the *APOE* e2 allele appears to be protective only in centenarians with low education.

Our findings are consistent with previous studies that showed a decline of cognitive function with older age in multiple domains (Harada et al., 2013), and the negative effect of *APOE* e4 on cognitive decline in centenarians (Arosio et al., 2017; Du et al., 2020). There is some evidence that, among centenarians, the e2 allele confers protection from cognitive decline (Kim et al., 2017; Sebastiani et al., 2020), in addition to increasing the chance for longevity (Sebastiani et al., 2019a) and conferring protection from aging-related diseases (Wolters et al., 2019; Kuo et al., 2020). Our studies also found that carriers of e2 with low education can delay their onset of moderate cognitive impairment. These results suggest that targeting e2 gene products could lead to important therapeutics for the preservation of cognitive function (Sebastiani et al., 2019b). Interestingly, the persistence of the e4 allele in the population has been linked to a survival advantage at a young age and treatments that target the effect of *APOE* alleles at an older age will need to consider the pleiotropic effect of this gene (Belloy et al., 2019).

Studies have shown the importance of early education on better cognitive function at an older age (Alley et al., 2007; Wilson et al., 2009). However, our study showed an antagonistic interaction between higher education and *APOE* alleles that suggests higher education may reduce the positive effect of e2 and increase the negative effect of e4 at extreme old age. A study of Japanese centenarians detected an education by *APOE* e4 interaction on cognition that differed by sex (Ishioka et al., 2016). The study showed that highly educated centenarians who carried at least one e4 allele had worse performance in the Mini-Mental State Examination, compared to poorly educated centenarians. Our findings are consistent with this counter-intuitive result but go one step further and disentangle the negative association of the e4 allele from the positive association of the e2 allele.

Our study focused on the association between *APOE* alleles and cognitive decline in extreme old individuals. Our analysis included well-known risk factors of cognitive decline such as older age and education. However, many more factors may affect the onset and rate of cognitive decline together or independently of *APOE*. Nutrition and dietary habits, for example, may be important risk factors to be considered in future analyses, given the role of *APOE* in lipid metabolism (Belloy et al., 2019). Studies showed that place of living changes the association between *APOE* alleles and extreme human longevity, after controlling for the overall genetic background (Gurinovich et al., 2019), suggesting that lifestyle may modify the genetic effect of *APOE*. However, we do not have data about dietary patterns in NECS centenarians to perform these analyses. The explained variance of our model was 0.29 for the e2 analysis and 0.30 for the e4 analysis. These results suggest that age, sex, *APOE*, and education can only explain less than 1/3 of the variability in BIMC scores and there are many factors, yet to be discovered, that should account for the unexplained variance. In future work, it will be interesting to collect data about additional physiological, medical, nutritional, or social factors that may contribute to maintain good cognitive function as people age and investigate their interplay with *APOE* alleles.

The Bayesian beta regression we used in this analysis has several strengths. First, this method allowed us to fit the test scores in the range of admissible values, and to model non-linear relations of the score with age and education. We illustrated the advantages of Bayesian beta regression by comparing it

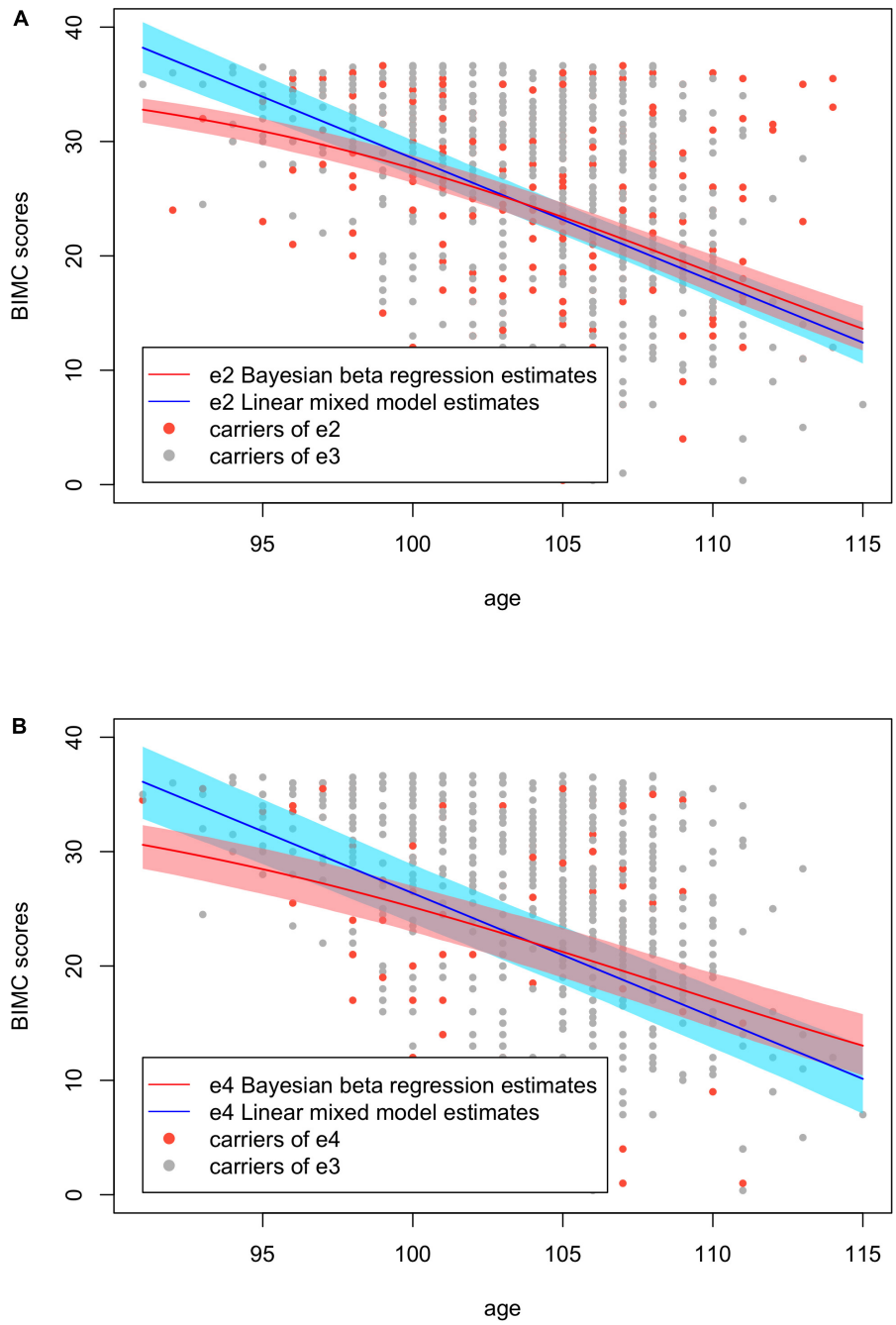

FIGURE 4 | Fitted Blessed Information-Memory-Concentration test scores using Bayesian beta regression and the linear mixed models with their credible/confidence bands. **(A)** Analysis in carriers of *APOE* e2 and e3. **(B)** Analysis in carriers of *APOE* e4 and e3.

with the conventional linear mixed model using the NECS data. Second, to increase statistical power, we included all participants with at least one cognitive function assessment in our model. A piecewise random intercept was used to include subjects with either repeated measurements or just one measurement. Finally, we carried out a model selection using the DIC criterion and then refined our model by credible intervals. From the diagnostic plots of the MCMC, the final models converged well. In addition, in previous studies (Caselli et al., 2009; Kulminski et al., 2015; Barral et al., 2017) *APOE* genotype was characterized as carriers of e4 (i.e., e2e4, e3e4, and e4e4), and non-carriers of e4 (i.e., e2e2, e2e3, and e3e3). This comparison could yield biased results since the non-carriers of the e4 group might not represent the most prevalent population. There are also studies (Kim et al., 2017) included e2 (e2e2 and e2e3), e3 (e3e3), and e4 (e3e4 and e4e4) in a combined analysis with dummy codings. In fact, we

performed a combined analysis on the NECS data and found that it produced smaller effects of e2 and e4. To eliminate these potential bias, we divided the participants into subgroups, where we only compared e2 carriers to e3 carriers and e4 carriers to e3 carriers separately.

Notable limitations of this study include loss to follow-up and mortality. Due to the extreme age of the study population, the mortality rate is considerably high: 49% of participants included in our analysis only completed the baseline evaluation. The absence of longitudinal data may induce bias, and thus we used a piecewise random intercept to compensate for those only with one record.

In conclusion, we examined the association between *APOE* alleles and cognitive decline in a cohort of centenarians. We confirmed the negative correlation between the e4 allele and cognition even at the extreme of human lifespan, and newly found the carrying the e2 alleles appears to be beneficial only in centenarians with poor education. The interaction between *APOE* e4 and education produced a counter-intuitive result that is, however, consistent with other results in centenarians. The antagonistic relation between higher education and carrying the e4 allele may be confounded by other risk factors and warrants more in-depth studies.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: The use of the data is restricted and needs approval. Requests to access these datasets should be directed to PS, psebastiani@tuftsmedicalcenter.org.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Boston University IRB. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

QX proposed the method used in the manuscript, conducted the statistical analysis, and drafted the manuscript. SA and TP designed the study, collected the data from the study cohort, and helped editing the manuscript. PS designed the analytic strategy, supervised the work, and critically revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Institute on Aging (NIA cooperative agreements U19-AG023122, R01 AG061844, and K01 AG057798).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.606831/full#supplementary-material

REFERENCES

Alley, D., Suthers, K., and Crimmins, E. (2007). Education and cognitive decline in older americans. *Res. Aging* 29, 73–94. doi: 10.1177/0164027506294245

Arosio, B., Ostan, R., Mari, D., Damanti, S., Ronchetti, F., Arcudi, S., et al. (2017). Cognitive status in the oldest old and centenarians: a condition crucial for quality of life methodologically difficult to assess. *Mech. Ageing Dev.* 165, 185–194. doi: 10.1016/j.mad.2017.02.010

Barral, S., Singh, J., Fagan, E., Cosentino, S., Andersen-Toomey, S. L., Wojczynski, M. K., et al. (2017). Age-related biomarkers in LLFS families with exceptional cognitive abilities. *J. Gerontol.Ser. A* 72, 1683–1688. doi: 10.1093/gerona/glx034

Belloy, M. E., Napolioni, V., and Greicius, M. D. (2019). A quarter century of APOE and Alzheimer's Disease: progress to date and the path forward. *Neuron* 101, 820–838. doi: 10.1016/j.neuron.2019.01.056

Blair, C., Folsom, A., Knopman, D. S., Bray, M., Mosley, T., and Boerwinkle, E. (2005). APOE genotype and cognitive decline in a middle-aged cohort. *Neurology* 64, 268–276.

Blessed, G., Tomlinson, B. E., and Roth, M. (1968). The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *Br. J. Psychiatry* 114, 797–811.

Bondi, M. W., Houston, W. S., Salmon, D. P., Corey-Bloom, J., Katzman, R., Thal, L. J., et al. (2003). Neuropsychological deficits associated with Alzheimer's disease in the very-old: discrepancies in raw vs. standardized scores. *J. Int. Neuropsychol. Soc.* 9, 783–795. doi: 10.1017/s1355617703950119

Branscum, A. J., Johnson, W. O., and Thurmond, M. C. (2007). Bayesian beta regression: applications to household expenditure data and genetic distance between foot-and-mouth disease viruses. *Austr. N. Z. J. Stat.* 49, 287–301.

Caselli, R. J., Dueck, A. C., Osborne, D., Sabbagh, M. N., Connor, D. J., and Ahern, G. L., et al. (2009). Longitudinal modeling of age-related memory decline and theAPOEε4 effect. *N. Engl. J. Med.* 361, 255–263. doi: 10.1056/nejmoa0809437

Clarke, B. S., and Clarke, J. L. (2018). *Predictive Statistics: Analysis and Inference Beyond Models*. Cambridge: Cambridge University Press.

Du, M., Andersen, S. L., Schupf, N., Feitosa, M. F., Barker, M. S., Perls, T. T., et al. (2020). Accepted. association between APOE alleles and change of neuropsychological tests in the long life family study. *J. Alzheimer's Dis.* (in press). doi: 10.3233/JAD-201113

Eskelson, B. N., Madsen, L., Hagar, J. C., and Temesgen, H. (2011). Estimating riparian understory vegetation cover with beta regression and copula models. *For. Sci.* 57, 212–221.

Fan, J., Tao, W., Li, X., Li, H., Zhang, J., and Wei, D., et al. (2019). The contribution of genetic factors to cognitive impairment and dementia: apolipoprotein E Gene, gene interactions, and polygenic risk. *Int. J. Mol. Sci.* 20:1177. doi: 10.3390/ijms20051177

Ferrari, S., and Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *J. Appl. Stat.* 31, 799–815.

Figueroa-Zúniga, J. I., Arellano-Valle, R. B., and Ferrari, S. L. P. (2013). Mixed beta regression: a bayesian perspective. *Comput. Stat. Data Anal.* 61, 137–147. doi: 10.1016/j.csda.2012.12.002

Gurinovich, A., Andersen, S. L., Puca, A., Atzmon, G., Barzilai, N., and Sebastiani, P. (2019). Varying effects of APOE alleles on extreme longevity in European Ethnicities. *J. Gerontol. Ser. A* 74(Suppl._1), S45–S51. doi: 10.1093/gerona/glz179

Harada, C. N., Natelson Love, M. C., and Triebel, K. L. (2013). Normal cognitive aging. *Clin. Geriatr. Med.* 29, 737–752. doi: 10.1016/j.cger.2013.07.002

Henderson, A. S., Jorm, A. F., Korten, A. E., Christensen, H., Jacomb, P. A., Easteal, S., et al. (1995). Apolipoprotein E allele ∈4, dementia, and cognitive decline

in a population sample. *Lancet* 346, 1387–1390. doi: 10.1016/s0140-6736(95)92405-1

Ishioka, Y. L., Gondo, Y., Fuku, N., Inagaki, H., Masui, Y., Takayama, M., et al. (2016). Effects of the APOE ε4 allele and education on cognitive function in Japanese centenarians. *AGE* 38, 495–503. doi: 10.1007/s11357-016-9944-8

Kawas, C., Karagiozis, H., Resau, L., Corrada, M., and Brookmeyer, R. (1995). Reliability of the blessed telephone information-memory-concentration test. *J. Geriatr. Psychiatry Neurol.* 8, 238–242.

Kim, Y. J., Seo, S. W., Park, S. B., Yang, J. J., Lee, J. S., Lee, J., et al. (2017). Protective effects of APOE e2 against disease progression in subcortical vascular mild cognitive impairment patients: a three-year longitudinal study. *Sci. Rep.* 7:1910. doi: 10.1038/s41598-017-02046-y

Kulminski, A. M., Arbeev, K. G., Culminskaya, I., Ukraintseva, S. V., Stallard, E., Province, M. A., et al. (2015). Trade-offs in the effects of the apolipoprotein E Polymorphism on risks of diseases of the heart, cancer, and neurodegenerative disorders: insights on mechanisms from the long life family study. *Rejuvenation Res.* 18, 128–135. doi: 10.1089/rej.2014.1616

Kuo, C.-L., Pilling, L. C., Atkins, J. L., Kuchel, G. A., and Melzer, D. (2020). ApoE e2 and aging-related outcomes in 379,000 UK Biobank participants. *Aging* 12, 12222–12233. doi: 10.18632/aging.103405

Mahley, R. (1988). Apolipoprotein E: cholesterol transport protein with expanding role in cell biology. *Science* 240, 622–630. doi: 10.1126/science.3283935

Mahoney, F. I., and Barthel, D. W. (1965). Functional evaluation: the Barthel Index: a simple index of independence useful in scoring improvement in the rehabilitation of the chronically ill. *Md State Med. J.* 14, 61–65.

O'donoghue, M. C., Murphy, S. E., Zamboni, G., Nobre, A. C., and Mackay, C. E. (2018). APOE genotype and cognition in healthy individuals at risk of Alzheimer's disease: a review. *Cortex* 104, 103–123. doi: 10.1016/j.cortex.2018.03.025

Plummer, M. (2016). *rjags: Bayesian Graphical Models Using MCMC. R Package Version 4 (6).*

Raber, J., Huang, Y., and Ashford, J. W. (2004). ApoE genotype accounts for the vast majority of AD risk and AD pathology. *Neurobiol. Aging* 25, 641–650. doi: 10.1016/j.neurobiolaging.2003.12.023

Sebastiani, P., Andersen, S. L., Sweigart, B., Du, M., Cosentino, S., Thyagarajan, B., et al. (2020). Patterns of multi-domain cognitive aging in participants of the long life family study. *GeroScience* 42, 1335–1350. doi: 10.1007/s11357-020-00202-3

Sebastiani, P., Gurinovich, A., Nygaard, M., Sasaki, T., Sweigart, B., Bae, H., et al. (2019a). APOE alleles and extreme human longevity. *J. Gerontol. Ser. A* 74, 44–51. doi: 10.1093/gerona/gly174

Sebastiani, P., Monti, S., Morris, M., Gurinovich, A., Toshiko, T., Andersen, S. L., et al. (2019b). A serum protein signature of APOE genotypes in centenarians. *Aging Cell* 18:e13023. doi: 10.1111/acel.13023

Sebastiani, P., and Perls, T. T. (2012). The genetics of extreme longevity: lessons from the New England centenarian study. *Front. Genet.* 3:277. doi: 10.3389/fgene.2012.00277

Simas, A. B., Barreto-Souza, W., and Rocha, A. V. (2010). Improved estimators for a general class of beta regression models. *Comput. Stat. Data Anal.* 54, 348–366. doi: 10.1016/j.csda.2009.08.017

Sinoff, G., and Ore, L. (1997). The barthel activities of daily living index: self-reporting versus actual performance in the Old-Old (= 75 years). *J. Am. Geriatr. Soc.* 45, 832–836. doi: 10.1111/j.1532-5415.1997.tb01510.x

Smithson, M., and Verkuilen, J. (2006). A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychol. Methods* 11, 54–71. doi: 10.1037/1082-989x.11.1.54

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *J. R. Stat. Soc. Ser. B* 64, 583–639. doi: 10.1111/1467-9868.00353

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2014). The deviance information criterion: 12 years on. *J. R. Stat. Soc. Ser. B* 76, 485–493. doi: 10.1111/rssb.12062

Staehelin, H. B., Perrig-Chiello, P., Mitrache, C., Miserez, A. R., and Perrig, W. J. (1999). Apolipoprotein E genotypes and cognitive functions in healthy elderly persons. *Acta Neurol. Scand.* 100, 53–60. doi: 10.1111/j.1600-0404.1999.tb00723.x

Swearingen, C. J., Tilley, B. C., Adams, R. J., Rumboldt, Z., Nicholas, J. S., Bandyopadhyay, D., et al. (2011). Application of beta regression to analyze ischemic stroke volume in NINDS rt-PA clinical trials. *Neuroepidemiology* 37, 73–82. doi: 10.1159/000330375

Wetter, S. R., Delis, D. C., Houston, W. S., Jacobson, M. W., Lansing, A., Cobell, K., et al. (2005). Deficits in inhibition and flexibility are associated with the APOE-E4 allele in nondemented older adults. *J. Clin. Exp. Neuropsychol.* 27, 943–952. doi: 10.1080/13803390490919001

Wilson, R. S., Hebert, L. E., Scherr, P. A., Barnes, L. L., Mendes De Leon, C. F., and Evans, D. A. (2009). Educational attainment and cognitive decline in old age. *Neurology* 72, 460–465. doi: 10.1212/01.wnl.0000341782.71418.6c

Wolters, F. J., Yang, Q., Biggs, M. L., Jakobsdottir, J., Li, S., Evans, D. S., et al. (2019). The impact of APOE genotype on survival: results of 38,537 participants from six population-based cohorts (E2-CHARGE). *PLoS One* 14:e0219668. doi: 10.1371/journal.pone.0219668

Zeileis, A., Cribari-Neto, F., Grün, B., and Kos-Midis, I. (2010). Beta regression in R. *J. Stat, Softw.* 34, 1–24.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 13 January 2021
doi: 10.3389/fnagi.2020.622321

Do High Mental Demands at Work Protect Cognitive Health in Old Age *via* Hippocampal Volume? Results From a Community Sample

Francisca S. Rodriguez[1,2,3,4], Sebastian Huhn[5,6,7], William A. Vega[8], Maria P. Aranda[8], Matthias L. Schroeter[5,6], Christoph Engel[4,9], Ronny Baber[4,10], Ralph Burkhardt[10], Markus Löffler[4,9], Joachim Thiery[10], Arno Villringer[5,6], Tobias Luck[11], Steffi G. Riedel-Heller[3†] and A. Veronica Witte[5,6,7†]*

[1]German Center for Neurodegenerative Diseases (DZNE), RG Psychosocial Epidemiology and Public Health, Greifswald, Germany, [2]Center for Cognitive Science, University of Kaiserslautern, Kaiserslautern, Germany, [3]Institute of Social Medicine, Occupational Health and Public Health (ISAP), University of Leipzig, Leipzig, Germany, [4]LIFE—Leipzig Research Center for Civilization Diseases, University of Leipzig, Leipzig, Germany, [5]Max-Planck-Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, [6]Clinic for Cognitive Neurology, University Hospital Leipzig, Leipzig, Germany, [7]Collaborative Research Centre 1052 "Obesity Mechanisms," Subproject A1, Faculty of Medicine, University of Leipzig, Leipzig, Germany, [8]Edward Roybal Institute of Aging, University of Southern California, Los Angeles, CA, United States, [9]Institute for Medical Informatics, Statistics and Epidemiology (IMISE), University of Leipzig, Leipzig, Germany, [10]Institute of Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics (ILM), University Hospital Leipzig, Leipzig, Germany, [11]Faculty of Applied Social Sciences, University of Applied Sciences Erfurt, Erfurt, Germany

OPEN ACCESS

Edited by:
Tomáš Paus,
University of Toronto, Canada

Reviewed by:
Rosanna Kathleen Olsen,
Baycrest Hospital, Canada
Ranil De Silva,
University of Sri Jayewardenepura, Sri Lanka

***Correspondence:**
Francisca S. Rodriguez
francisca-saveria.rodriguez@dzne.de

†These authors have contributed equally to this work

Received: *28 October 2020*
Accepted: *21 December 2020*
Published: *13 January 2021*

Citation:
Rodriguez FS, Huhn S, Vega WA, Aranda MP, Schroeter ML, Engel C, Baber R, Burkhardt R, Löffler M, Thiery J, Villringer A, Luck T, Riedel-Heller SG and Witte AV (2021) Do High Mental Demands at Work Protect Cognitive Health in Old Age via Hippocampal Volume? Results From a Community Sample. Front. Aging Neurosci. 12:622321. doi: 10.3389/fnagi.2020.622321

As higher mental demands at work are associated with lower dementia risk and a key symptom of dementia is hippocampal atrophy, the study aimed at investigating the association between mental demands at work and hippocampal volume. We analyzed data from the population-based LIFE-Adult-Study in Leipzig, Germany (n = 1,409, age 40–80). Hippocampal volumes were measured *via* three-dimensional Magnetic resonance imaging (MRI; 3D MP-RAGE) and mental demands at work were classified *via* the O*NET database. Linear regression analyses adjusted for gender, age, education, APOE e4-allele, hypertension, and diabetes revealed associations between higher demands in "language and knowledge," "information processing," and "creativity" at work on larger white and gray matter volume and better cognitive functioning with "creativity" having stronger effects for people not yet retired. Among retired individuals, higher demands in "pattern detection" were associated with larger white matter volume as well as larger hippocampal subfields CA2/CA3, suggesting a retention effect later in life. There were no other relevant associations with hippocampal volume. Our findings do not support the idea that mental demands at work protect cognitive health *via* hippocampal volume or brain volume. Further research may clarify through what mechanism mentally demanding activities influence specifically dementia pathology in the brain.

Keywords: hippocampus, cognitive functioning, mental demands, intellectual activities, aging

INTRODUCTION

Mentally demanding activities at work seem to delay cognitive decline and lower dementia risk (Valenzuela and Sachdev, 2006a,b), as longitudinal studies (Smyth et al., 2004; Karp et al., 2009), as well as twin studies (Andel et al., 2005; Potter et al., 2007), have shown. Therein, it seems that the effect depends on the type of mental demand at work (MDW). A recent analysis of a prospective multi-center cohort study following 2,315 individuals up to 11 years has shown that higher MDW involving "information processing" (e.g., analyzing data) and "pattern detection" (e.g., detecting a figure, object, word, or sound that is hidden in other distracting material) relate to lower dementia risk in general, and high MDW involving "mathematics" and "creativity" delay dementia onset (Then et al., 2015). In that cohort, other MDW did not significantly affect dementia risk at all. To date, it is unclear why some demands may be more protective against dementia than others.

One possible explanation for this phenomenon might be related to potential mechanisms of how MDW and related factors alter dementia risk. However, these mechanisms, particularly at the neurobiological level, are largely unknown. Evidence on this subject matter is sparse. One study reported that, among young adults, the cognitive complexity of work demands is associated with better white matter integrity (Kaup et al., 2018). The only study with older adults, that we could identify, matched people by cognitive status and observed that job complexity was associated with a smaller hippocampal volume and more brain atrophy (Boots et al., 2015). Hippocampal volume declined with older age (Fjell et al., 2014; Fraser et al., 2015) and accelerated hippocampal atrophy is implicated in Alzheimer's dementia (AD; Kaye et al., 1997; den Heijer et al., 2010). Hence, possibly, the hippocampus may play a role in the protective effects of mental demands at work on cognitive health. First human interventional studies implementing high-resolution magnetic resonance imaging (MRI) suggested that modifiable factors such as cognitive and physical activity exert protective effects on cognitive health *via* improvements in hippocampus plasticity (Duzel et al., 2016). For example, two months of memory training compared to placebo increased hippocampal volume in a group of older adults (Engvig et al., 2014), and AD patients with high educational attainment seem to have a larger hippocampal volume (Shpanskaya et al., 2014). Exposure to higher demands may thus also work on this pathway. However, the evidence is still sparse and it remains unclear whether long-lasting mental stimulation preserves hippocampus plasticity.

The study aimed to explore whether mental stimulation at work protects cognitive health by preserving hippocampal volume. Specifically, we investigated whether five types of MDW (language and knowledge, information processing, pattern recognition, mathematics, and creativity at work) were associated with hippocampal volume in a cross-sectional analysis of the large community-based "Adult Study" of the Leipzig Research Centre for Civilization Diseases (LIFE). As the effect may be dependent on being active in the workforce, we conducted the analyses separately for those working and those retired. Also, we tested whether hippocampal and brain volume (HBV) mediates the association between MDW and cognitive performance.

MATERIALS AND METHODS

Study Design

We analyzed data of the "Adult Study" of the Leipzig Research Centre for Civilization DZNE (LIFE), a large population-based study investigating the prevalence, early onset markers, genetic predispositions, and the role of lifestyle factors in major civilization diseases. The details of the study have been described by Loeffler et al. (2015). Briefly, a random age- and a sex-stratified sample of residents of the city of Leipzig was obtained from the residents' registry office. A letter of invitation to participate in the study was sent to every individual on the list. The only exclusion criterion was being pregnant.

The LIFE-Adult Study was conducted between August 2011 and November 2014. The study was approved by the ethics committee of the Medical Faculty of the University of Leipzig and was carried out in conformity with the principles embodied in the Declaration of Helsinki. All participants signed written informed consent before participation.

The assessments included physical and medical examinations, self-administered questionnaires, and psychometric testing, which were administered by trained study assistants and monitored by experienced scientists following standardized study protocols (Loeffler et al., 2015). A subsample of participants completed MRI at a second examination date (random sample from population registry, n = 2,637, 18–80 years). From these, we included all participants age 40–80 years (n = 2,271) from the analysis. Three-hundred and four participants were excluded due to major brain pathology (e.g., stroke, multiple sclerosis, tumors) or bad image quality (e.g., motion artifact). Another 10 participants were excluded because they reported having been diagnosed with a neurological or psychiatric disorder (n = 4 substance-related disorder, n = 2 human immunodeficiency virus (HIV), n = 2 Parkinson's disease, n = 1 multiple sclerosis, n = 1 epilepsy). None of the participants included in the analyses had dementia or major neurocognitive disorder; we verified *via* cognitive testing (see "Cognitive Performance" section). We also excluded 74 participants because they were unemployed or retired with a total unemployment period of more than three years during their life. Further, individuals with missing data on important covariates were excluded: n = 128 missing data on apolipoprotein E (APOE) ε4 genotype, n = 26 missing data on having diabetes, n = 101 missing data on hypertension, and n = 117 occupational information could not be matched to O*NET database. And another n = 102 were excluded due to incomplete or invalid cognitive testing. The total number of participants in the analysis was n = n = 1,409.

Hippocampal and Brain Volume

To estimate volumes, we used the three-dimensional Magnetization-Prepared Rapid Gradient Echo sequence (3D MP-RAGE) anatomical T1-weighted images of the brain, acquired with a 3T Siemens Magnetom Verio Syngo MR B17 at

the University Clinic Leipzig. Generalized autocalibrating partially parallel acquisition parallel imaging technique (Griswold et al., 2002; according to the Alzheimer's Disease Neuroimaging Initiative standard protocol (Wang et al., 2014) was applied using following scanning variables: repetition time/ echo time 2,300 ms/ 2.98 ms; flip angle 9°; slice/ voxel size 1 mm/ 1.0 × 1.0 × 1.0 mm (x × y × z); slices 176; the field of view 256 mm; bandwidth 240 Hz/Px; base resolution 256; scanning time 5 min 10 s. Clinical MRI ratings were performed by neuroradiologists blind to further assessment data. Volumes of intracranial volume (ICV) and total gray and white matter volume derived from FreeSurfer analysis. FreeSurfer (FS) version 6 was used. FS is a free software package developed by the Athinoula A. Martinos Center for Biomedical Imaging of Harvard University[1]. The images were segmented into gray matter maps and matched to a study-specific cerebral-/cerebellar-specific template. For better tissue segmentation and interindividual alignment, the cerebrum and cerebellum were estimated separately. After deleting non-brain tissue, ICV was obtained by adding up the gray matter, white matter, and CSF. Three-dimensional sequences of both hippocampi were reconstructed. Automated segmentation of the hippocampal subfields was performed by an algorithm implemented in FS (Van Leemput et al., 2009). The subfields Cornu Ammonis (CA) 1, CA 2–3, CA 4-Dentate Gyrus, presubiculum, and subiculum were considered for further analysis (Erickson et al., 2011; Brickman et al., 2014). Total (left and right) whole hippocampal and subfield volumes (in mm^3), as well as ICV, were normally distributed. We adjusted volumes of the whole hippocampus and its subfields for ICV according to (Raz et al., 2005; Kerti et al., 2013) using the following formula: adjusted volume (in mm^3) = raw volume (in mm^3) − ß* (ICV − ICVmean) with ß being the slope of the regression of the respective volume on ICV. Manual quality control of FS labels was done by two experienced staff members individually according to standard operating procedures. Also, a script was programmed for validation of delineation. Analyses were rerun and unusable or inconsistent data were excluded from analyses.

Mental Demands at Work (MDW)

Mental demands at work (MDW) were investigated as predictors in this study. In standardized interviews, participants provided information on their present or, if retired, on their former occupation. The occupations were translated into English and coded according to the 2010 Standard Occupational Classification of the O*NET database[2]—a validated database containing standardized occupation—specific descriptors that were developed by the US Department of Labor/Employment and Training Administration (USDOL/ETA). The occupation code of every participant came with a great number of variables that describe details of the work tasks. Each variable is indicating on a continuous scale the level (from low to high) on which the person is facing the particular work characteristic. For purpose of analysis, we selected only those variables that describe "mental" demands at work (O*NET variables "Cognitive Abilities" 1.A.1.a–1.A.2.c.3 and "Skills and worker requirements" 2.A.1.a.1–4.A.4.c.3). By following the classification scheme of MDW in previous analyses (Then et al., 2015), relevant variables were combined in clusters of MDW: "language and knowledge," "information processing," "mathematics," "pattern detection," and "creativity." The value of each MDW is the mean of the included variables.

Cognitive Performance

Cognitive performance was used in this study to check whether hippocampal and brain volume mediates the association between mental demands at work and cognitive performance as a precondition for the effect under investigation. Cognitive performance was assessed by trained study assistants and was subject to regular quality control by experienced psychologists. Participants completed the Verbal Fluency Test, the Trail Making Test (TMT), and the Word List Test—subtests of the German version of the neuropsychological test battery of the Consortium to Establish a Registry for Alzheimer's disease; (CERADplus; Morris et al., 1989). The German version of the CERADplus has been validated (see Aebi et al., 2002). The Verbal Fluency Test is considered to measure verbal abilities, semantic fluency, and semantic memory (Kraan et al., 2013). The participants are instructed to name as many animals as possible in one minute. The participant's score equals the number of correctly named animals. The TMT is considered to measure working memory, task-switching ability (Salthouse, 2011), and executive control (Arbuthnott and Frank, 2010). The participants are instructed, first, to connect numbers in ascending order as fast as they can (version A) and, second, to connect numbers and letters alternatingly (version B). When an error is made, the participants had to return to the number where the error originated. The participant's score corresponds to the number of seconds needed to complete the test. The Word List Test is considered to measure memory. The participants are instructed to read out ten words and subsequently recall them. This was repeated three times with the same 10 words. The participant's score is the number of words remembered correctly.

Covariates

Age was calculated as the difference in years by subtracting the birth date from the interview date. Gender was estimated using the sex that was recorded in the population registry. Education, as reported by the participant, was categorized as "low" for having completed high school or less, "moderate" for having completed college or a professional training school, and "high" for having completed a university degree. Information on diabetes and hypertension was obtained by asking the participant "Have you ever been diagnosed with . . .?." The APOE genotype was identified from peripheral blood leukocytes using an automated protocol on the Qiagen Autopure LS (Qiagen, Hilden, Germany) and by following the method of Aslanidis and Schmitz (1999) *via* Roche Lightcycler 480 (Aslanidis and Schmitz, 1999).

Statistical Analysis

Statistical analyses were performed using STATA 16 and employed an alpha level for statistical significance of 5%

[1]https://surfer.nmr.mgh.harvard.edu/
[2]http://www.onetonline.org

($p < 0.05$, two-tailed). Bonferroni-correction for the clusters of MDW (five levels) would yield a significance level of $p < 0.01$. All analyses were conducted separately for those actively working and those retired, as associations might be different in people who have retired from their job compared to those who face MDW on a daily level.

Descriptive data analyses on differences in whole hippocampal and brain volume and MDW concerning characteristics of the study sample were conducted *via* analysis of variance (ANOVA) and Pearson's correlation. The association between hippocampal and brain volume and MDW on cognitive performance was analyzed *via* pairwise comparison correlations. The main research question, the association between MDW and hippocampal and brain volume, was analyzed *via* linear regression analyses, first univariate and then adjusted for factors that might introduce confounding due to known effects on hippocampus volume i.e., gender and age (Pruessner et al., 2001), education (Noble et al., 2012), diabetes (Gold et al., 2007), hypertension (Shih et al., 2016), and APOE e4-allele (Plassman et al., 1997).

The following sensitivity analyses were conducted: as education is an important confounder, we analyzed potential interaction effects using the same linear regression model with an additional interaction term for education and MDW. As hippocampal and brain volume deteriorates with aging, the association under investigation might be age-sensitive. Therefore, we analyzed potential interaction effects using the same linear regression model as before with an additional interaction term for age and MDW [for retired and not-retired individuals together adjusted for being retired (yes/no)]. Finally, we tested whether hippocampal and brain volume mediates the association between MDW and cognitive performance using partial least square structural equation modeling.

RESULTS

Hippocampal and Brain Volume

The mean whole hippocampus volume was 7.7 [standard deviation (SD) 0.8], the mean ICV 1,472,315 mm^3 (SD 147,104.7), the mean gray matter volume 585,876.9 mm^3 (SD 61,774.2), and the mean white matter 424,326.4 mm^3 (SD 50,995.6). Associations between personal characteristics and brain parameters are shown in **Table 1**. Individuals who were retired and those who were older had a smaller hippocampal volume, less ICV, less gray matter, and less white matter (see **Tables 1**, **2**). Individuals with hypertension had a smaller hippocampal volume and less gray matter. There was no significant difference concerning hippocampal volumes and education or APOE e4 allele (see **Table 1**). A smaller hippocampal volume was associated with poorer performance in the Word List Test and the TMT, and less gray matter was associated with poorer performance in all cognitive tests (see **Table 2**). In individuals that were not retired, the white matter was associated with performance in the Word List Test and the TMT A (see **Table 2**).

Mental Demands at Work (MDW)

The average level of the MDW "language and knowledge" was 3.31 (SD 0.63), of "information processing" was 3.69 (SD 0.80), of "mathematics" was 2.36 (SD 1.06), of "pattern detection" was 2.67 (SD 0.59), and of "creativity" was 2.99 (SD 1.24). Individuals who were retired had, on average, significantly higher levels of "language and knowledge" (3.35 vs. 3.27, $F = 5.88$, $p = 0.015$), "mathematics" (2.49 vs. 2.23, $F = 22.18$, $p < 0.001$), "pattern detection" (2.71 vs. 2.62, $F = 7.33$, $p = 0.007$), and "creativity" (3.09 vs. 2.89, $F = 9.54$, $p = 0.002$), but not of "information processing" (3.73 vs. 3.66, $F = 3.41$, $p = 0.065$). Individuals with higher education had significantly higher MDW levels compared to individuals with lower education (see **Table 3**). Retired men had significantly higher MDW than women; there was no significant difference in MDW between men and women that were still actively working (see **Table 3**). Further, there was no significant difference concerning APOE e4 or diabetes status.

Higher MDW were significantly associated with a better performance in cognitive tests that measured executive cognitive abilities (all $p < 0.001$; TMT B, Verbal Fluency Test), whilst the performance in the TMT A and the Word List Test was only associated with a higher level in the MDW "language and knowledge," "information processing," and "pattern detection" among those who were not yet retired (see **Table 3**, see also **Figure 1**).

Association of Demands With Hippocampal and Brain Volume

Higher MDW was associated with more white and gray matter (see **Table 2** and **Figure 2**). If adjusted for confounders, only higher MDW "information processing" was associated with larger white and gray matter volume, higher MDW "creativity" and "language and knowledge" only if not yet retired (see **Table 4**). Higher MDW "pattern detection" was only significantly associated with more white and gray matter volume if the persons had retired (see **Table 4**).

For hippocampal volume, there was only one significant association between hippocampal volume and MDW in the descriptive univariate analyses: Among those retired, a higher level of "creativity" correlated with a smaller hippocampal volume (see **Table 2**). This association was no longer present when accounting for confounders (see **Table 4**). We considered the different hippocampal subfields and observed that a higher level of the MDW "pattern detection" was associated with larger hippocampal subfields CA2/CA3 in retired individuals (see **Table 4**). The estimates point out that the difference between those with high and low demands is greater in younger age and decreases with older age.

Sensitivity Analyses

Interaction effects between education and MDW were found for retired individuals. Education interacts with "pattern detection" on white matter volume (moderate education $b = -19{,}831.3$, CI95 $-34{,}962.9$ to $-4{,}699.8$, $p = 0.010$; high education $b = -13{,}588.1$, CI95 $-27{,}207.3$ to -31.1, $p = 0.051$) and intracranial volume (moderate education $b = -49{,}895.7$, CI95 $-97{,}128.5$ to $-2{,}662.9$, $p = 0.038$), indicating that individuals

TABLE 1 | Mean hippocampal and brain volume (in mm^3) concerning the characteristics of the study sample (n = 1, 409).

		Hippocampal volume[2]		White matter		Gray matter		ICV	
		Not retired	**Retired**	**Not retired**	**Retired**	**Not retired**	**Retired**	**Not retired**	**Retired**
	N (%)	721 (47.7)	789 (52.3)	600 (46.3)	697 (53.7)	721 (47.7)	789 (52.3)	721 (47.7)	789 (52.3)
	Mean (SD)	8,110.5 (713.6)	7,343.9 (797.7)	440,912.5 (54,663.3)	410,048.5 (42,772.2)	616,548.2 (61,820.7)	557,849.0 (46,565.5)	1,494,801.0 (148,403.6)	1,451,766.7 (142,937.6)
	F^1		384.66	129.85		438.9		32.92	
	p^1		**<0.001**	**<0.001**		**<0.001**		**<0.001**	
Gender	Male	8,135.9 (742.9)	7,183.3 (792.4)	463,028.2 (47,883.9)	429.065.3 (41,027.6)	644,587.9 (56,991.8)	580,159.6 (43,521.0)	1,569,507.5 (129,104.9)	1,533,354.7 (121,579.1)
	Female	8,080.0 (676.5)	7,521.1 (766.2)	412,381.6 (49,463.4)	389,616.8 (34,403.8)	582,951.7 (49,438.5)	533,218.2 (36,300.4)	1,405,289.8 (117,332.2)	1,361,693.6 (105,930.0)
	F^1	1.10	3.89	160.42	187.75	235.6	267.6	314.12	442.97
	p^1	0.30	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**
Education	Low	8,120.4 (693.1)	7,398.8 (796.8)	437,954.5 (53,649.5)	405,310.3 (43,938.4)	614,713.5 (65,268.6)	550,588.1 (46,902.9)	1,479,338.1 (149,924.9)	1,418,773.8 (143,087.2)
	Moderate	8,193.2 (671.9)	7,337.2 (817.4)	434,644.2 (45,154.7)	404,216.5 (41,484.5)	612,291.5 (60,004.2)	549,241.7 (44,133.9)	1,480,419.5 (135,626.6)	1,433,737.0 (145,244.8)
	High	8,050.2 (753.4)	7,302.5 (785.7)	447,843.9 (60,363.1)	417,585.2 (41,603.8)	620,912.1 (59,457.9)	569,215.9 (45,528.4)	1,518,347.6 (151,877.8)	1,490,322.7 (132,331.5)
	F^1	2.20	1.07	3.14	7.66	1.23	16.91	5.88	21.48
	p^1	0.11	0.34	**0.044**	**0.001**	0.293	**<0.001**	**0.003**	**<0.001**
APOE e4	No	8,108.5 (715.9)	7,357.3 (805.5)	441,460.1 (52,466.5)	408,982.5 (42,553.7)	616,906.9 (62,950.9)	558,195.9 (46,903.5)	1,494,868.2 (15,271.1)	1,449,386.9 (140,586.4)
	Yes	8,117.2 (707.8)	7,302.1 (773.2)	438,952.2 (62,075.4)	413,485.7 (43,421.4)	615,320.1 (57,952.9)	556,770.2 (45,602.3)	1,494,571.1 (138,582.0)	1,459,166.4 (150,151.5)
	F^1	0.02	0.70	0.22	1.40	0.08	0.14	0.0	0.68
	P^1	0.89	0.40	0.643	0.238	0.773	0.712	0.982	0.409
Diabetes	No	8,138.2 (700.2)	7,379.1 (789.7)	441,458.6 (55,020.1)	411,409.6 (43,710.5)	617,428.5 (62,295.9)	559,631.1 (47,019.1)	1,494,128.4 (149,160.9)	1,450,186.9 (142,902.6)
	Yes	7,450.3 (723.4)	7,163.7 (816.7)	427,806.8 (44,191.6)	403,160.5 (37,075.9)	595,541.4 (45,017.6)	548,731.3 (43,203.9)	1,510,850.5 (130,274.5)	1,459,849.5 (143,399.9)
	F^1	26.79	7.94	1.44	3.58	3.50	5.95	0.35	0.49
	p^1	**<0.001**	**0.005**	0.231	0.059	0.062	**0.015**	0.553	0.483
Hypertension	No	8169.2 (699.7)	7426.5 (800.4)	442,415.7 (56,364.8)	411,163.6 (41,694.7)	623,744.7 (62,593.8)	563,508.4 (45,029.6)	1,497,694.8 (145,837.2)	1,449,627.3 (140,500.7)
	Yes	7926.3 (715.9)	7253.4 (793.3)	436,055.2 (49,324.3)	409,596.1 (43,795.8)	599,634.6 (56,702.8)	554,395.7 (48,057.6)	1,484,954.9 (156,973.3)	1,455,907.2 (146,564.7)
	F^1	14.6	8.42	1.51	0.22	19.01	6.73	0.90	0.34
	p^1	**<0.001**	**0.004**	0.219	0.638	**<0.001**	**0.009**	0.343	0.561

Notes: 1, as estimated by analysis of variance (ANOVA); 2, whole hippocampal volume in mm^3 including left and right hippocampi; ICV, intracranial volume; N, number of participants; p, level of significance; SD, standard deviation; SES, socioeconomic status. Bold indicates significance <0.05.

TABLE 2 | Pearson's pairwise correlation between hippocampal and brain volume with cognitive performance and mental demands at work (MDW, n = 1, 409).

	Hippocampal volume[1]		White matter		Gray matter		ICV	
	Retired r (p)	Not retired r (p)	Retired r (p)	Not retired r (p)	Retired r (p)	Not retired r (p)	Retired r (p)	Not retired r (p)
Age	**−0.349 (<0.001)**	**−0.277 (<0.001)**	**−0.174 (<0.001)**	**−0.086 (0.035)**	**−0.138 (<0.001)**	**−0.481 (<0.001)**	0.038 (0.293)	**0.123 (0.001)**
Verbal fluency	0.004 (0.99)	−0.057 (0.12)	0.069 (0.065)	**0.091 (0.026)**	**0.079 (0.028)**	**0.146 (<0.001)**	**0.079 (0.026)**	**0.112 (0.003)**
Word list	**0.199 (<0.001)**	**0.168 (<0.001)**	−0.055 (0.149)	−0.055 (0.181)	**−0.070 (0.048)**	**0.138 (<0.001)**	**−0.168 (<0.001)**	−0.036 (0.330)
Trail making Test A	**−0.154 (<0.001)**	**−0.170 (<0.001)**	−0.074 (0.051)	**−0.129 (0.002)**	**−0.101 (0.005)**	**−0.270 (<0.001)**	0.014 (0.699)	**−0.108 (0.004)**
Trail making Test B	**−0.132 (<0.001)**	**−0.165 (<0.001)**	**−0.085 (0.025)**	−0.077 (0.060)	**−0.108 (0.002)**	**−0.219 (<0.001)**	−0.007 (0.838)	−0.070 (0.059)
MDW language and knowledge	−0.053 (0.14)	0.052 (0.16)	**0.123 (0.001)**	**0.107 (0.009)**	**0.154 (<0.001)**	**0.110 (0.003)**	**0.180 (<0.001)**	**0.087 (0.019)**
MDW information processing	−0.066 (0.06)	0.013 (0.74)	**0.182 (<0.001)**	**0.126 (0.002)**	**0.198 (<0.001)**	**0.123 (0.001)**	**0.225 (<0.001)**	**0.107 (0.004)**
MDW mathematics	−0.052 (0.15)	0.011 (0.78)	**0.081 (0.034)**	**0.109 (0.007)**	**0.127 (<0.001)**	**0.124 (0.001)**	**0.139 (<0.001)**	**0.118 (0.002)**
MDW pattern detection	−0.035 (0.32)	0.028 (0.45)	**0.158 (<0.001)**	**0.128 (0.002)**	**0.188 (<0.001)**	**0.118 (0.002)**	**0.199 (<0.001)**	**0.095 (0.011)**
MDW creativity	**−0.111 (0.002)**	0.043 (0.25)	**0.153 (<0.001)**	**0.165 (<0.001)**	**0.196 (<0.001)**	**0.158 (<0.001)**	**0.225 (<0.001)**	**0.169 (<0.001)**

Notes: 1, whole hippocampal volume in mm^3 including left and right hippocampi; ICV, intracranial volume; p, level of significance; r, Pearson's correlation coefficient. Bold indicates significance <0.05.

with low education have a greater increase in white matter and intracranial volume if they work in jobs with higher MDW "pattern detection" than individuals with moderate education. For individuals who were not retired, education interacted with "creativity" on white matter volume (moderate education b = 11749.9, CI95 2982.3–20517.6, p = 0.009), indicating that low-educated individuals have a greater increase in the white matter if they work in jobs with higher MDW "creativity" than moderately-educated individuals.

Interaction with age was found for higher MDW "creativity" (b = −211.6, CI95 −352.9 to −70.2, p = 0.003) and "language and knowledge" (b = −271.9, CI95 −500.5 to −43.5, p = 0.020) on white matter. Higher MDW "creativity" interacted also with age on total hippocampal volume (b = −3.4, CI95 −5.5 to −1.4, p = 0.001), indicating that hippocampal volume declines faster in the age period 40–60 in individuals with higher MDW "creativity" than in individuals with low demands.

The potential mediating effects of hippocampal and brain volume on the association of HDW on cognitive performance were analyzed *via* structural equation modeling using latent factor variables. Results are shown in **Table 5**. For retired individuals, higher MDW was significantly associated with better performance in cognitive testing. For individuals not yet retired, higher MDW was significantly associated with better performance in cognitive testing as well as hippocampal and brain volume. For both groups, there were no indirect associations from MDW to cognitive performance *via* hippocampal and brain volume. Covariates were significantly associated with hippocampal and brain volume and, in this way, also indirectly associated with cognitive performance.

DISCUSSION

Since previous studies have shown that high levels of mental demands at work (MDW) are associated with lower dementia risk (Then et al., 2015), this study aimed to investigate whether higher MDW protects cognitive health by preserving hippocampal volume. Findings indicate only a significant association between higher MDW "pattern detection" (not any other MDW) with larger hippocampal subfields CA2 / CA3 in retired individuals. However, our sample comprised only community-based individuals without dementia and we cannot derive any conclusions on individuals with severe cognitive impairments. It is conceivable that associations between MDW and hippocampal volume are sensitive to the neurodegenerative processes of dementia, studies using clinical samples may observe stronger effects on hippocampal size.

The lack of a general association between MDW and hippocampal volume in our sample could be explained by multiple factors. On the one hand, a (neuro)protective effect of MDW may not be related to higher hippocampal plasticity. A study that investigated associations between occupational complexity and hippocampal volume in Alzheimer patients found no association (Boots et al., 2015). Then again, crude hippocampus volume might not be a valid biomarker for hippocampal plasticity. Besides volume, hippocampus microstructure and functional connectivity measured using

TABLE 3 | Characteristics of the study sample (n = 1,409).

		Retired					Not retired				
		MDW language and knowledge	**MDW information processing**	**MDW mathematics**	**MDW pattern detection**	**MDW creativity**	**MDW language and knowledge**	**MDW information processing**	**MDW mathematics**	**MDW pattern detection**	**MDW creativity**
		M (SD)	**M (SD)**	**M (SD)**	**M (SD)**	**M (SD)**	**M (SD)**	**M (SD)**	**M (SD)**	**M (SD)**	**M (SD)**
Sex	male	3.46 (0.67)	3.92 (0.78)	2.67 (1.19)	2.83 (0.59)	3.44 (1.24)	3.25 (0.67)	3.67 (0.81)	2.29 (1.11)	2.65 (0.63)	3.04 (1.26)
	Female	3.24 (0.58)	3.53 (0.80)	2.28 (0.97)	2.56 (0.53)	2.70 (1.19)	3.31 (0.53)	3.65 (0.77)	2.16 (0.84)	2.59 (0.57)	2.72 (1.09)
	F^1	23.68	47.32	25.08	45.84	72.20	1.85	0.11	2.90	2.04	13.16
	p^1	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	0.17	0.74	0.09	0.15	**<0.001**
Education	low	2.93 (0.49)	3.27 (0.67)	1.96 (0.73)	2.47 (0.45)	2.34 (1.02)	2.97 (0.57)	3.34 (0.78)	1.87 (0.80)	2.43 (0.61)	2.44 (1.08)
	moderate	3.24 (0.61)	3.59 (0.73)	2.34 (1.01)	2.61 (0.55)	2.90 (1.14)	3.24 (0.53)	3.61 (0.65)	2.07 (0.71)	2.62 (0.59)	2.77 (0.93)
	high	3.77 (0.49)	4.21 (0.70)	3.01 (1.19)	2.96 (0.59)	3.82 (1.13)	3.58 (0.55)	3.98 (0.73)	2.67 (1.14)	2.81 (0.56)	3.39 (1.25)
	F^1	195.39	138.10	83.62	63.81	136.28	80.57	52.06	54.21	29.00	51.23
	p^1	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**
APOE e4	No	3.36 (0.64)	3.73 (0.81)	2.49 (1.10)	2.72 (0.57)	3.09 (1.27)	3.28 (0.62)	3.65 (0.81)	2.26 (1.03)	2.62 (0.62)	2.89 (1.22)
	Yes	3.32 (0.65)	3.76 (0.83)	2.44 (1.12)	2.67 (0.59)	3.07 (1.28)	3.25 (0.59)	3.68 (0.73)	2.12 (0.87)	2.63 (0.54)	2.87 (1.10)
	F^1	0.50	0.20	0.41	1.20	0.06	0.29	0.11	2.30	0.02	0.05
	p^1	0.48	0.66	0.52	0.27	0.81	0.59	0.74	0.13	0.90	0.83
Diabetes	No	3.37 (0.64)	3.75 (0.81)	2.50 (1.11)	2.71 (0.59)	3.12 (1.26)	3.28 (0.61)	3.66 (0.79)	2.24 (0.98)	2.63 (0.60)	2.89 (1.18)
	Yes	3.28 (0.63)	3.66 (0.83)	2.39 (1.07)	2.69 (0.55)	2.93 (1.32)	3.18 (0.69)	3.55 (0.88)	1.99 (1.28)	2.53 (0.72)	2.95 (1.60)
	F^1	1.76	1.35	0.99	0.02	2.51	0.75	0.54	1.67	0.72	0.06
	p^1	0.18	0.25	0.32	0.89	0.11	0.39	0.47	0.20	0.40	0.81
Hypertension	No	2.02 (0.85)	3.74 (0.81)	2.45 (1.09)	2.69 (0.58)	3.06 (1.26)	1.97 (0.79)	3.67 (0.80)	2.23 (0.99)	2.63 (0.60)	2.89 (1.16)
	Yes	2.06 (0.83)	3.73 (0.81)	2.51 (1.13)	2.71 (0.58)	3.11 (1.28)	1.88 (0.79)	3.59 (0.76)	2.19 (1.07)	2.59 (0.63)	2.79 (1.24)
	F^1	0.38	0.02	0.39	0.22	0.25	1.56	1.19	0.26	0.51	1.00
	p^1	0.537	0.897	0.533	0.640	0.618	0.213	0.276	0.608	0.475	0.317
		r (p)^2	**r (p)^2**	**r (p)^2**	**r (p)^2**	**r (p)^2**	**r (p)^2**	**r (p)^2**	**r (p)^2**	**r (p)^2**	**r (p)^2**
Verbal Fluency		**0.177 (<0.001)**	**0.143 (<0.001)**	**0.118 (0.001)**	0.069 (0.051)	**0.154 (<0.001)**	**0.159 (<0.001)**	**0.120 (0.001)**	0.065 (0.079)	**0.086 (0.021)**	**0.117 (0.002)**
Wordlist		**0.102 (0.004)**	0.068 (0.06)	0.050 (0.16)	0.030 (0.40)	0.016 (0.66)	**0.128 (0.001)**	**0.098 (0.009)**	0.069 (0.06)	**0.075 (0.045)**	0.024 (0.53)
TMT A		−0.050 (0.16)	−0.040 (0.26)	−0.041 (0.26)	−0.050 (0.16)	−0.030 (0.40)	**−0.107 (0.004)**	**−0.093 (0.013)**	−0.037 (0.33)	**−0.085 (0.023)**	−0.051 (0.17)
TMT B		**−0.179 (<0.001)**	**−0.157 (<0.001)**	**−0.165 (<0.001)**	**−0.132 (<0.001)**	**−0.107 (0.003)**	**−0.244 (<0.001)**	**−0.165 (<0.001)**	**−0.146 (<0.001)**	**−0.142 (<0.001)**	**−0.136 (<0.001)**

Notes: 1, as estimated by analysis of variance (ANOVA); 2, as estimated by pairwise comparison correlations; CI, confidence interval 95%; M, mean; MDW, mental demands at work; N, number of participants, p, level of significance; r, regression coefficient; SD, standard deviation; TMT A, Trail Making Test A with higher scores indicating poorer cognitive performance; TMT B, Trail Making Test B with higher scores indicating poorer cognitive performance. Bold indicates significance <0.05.

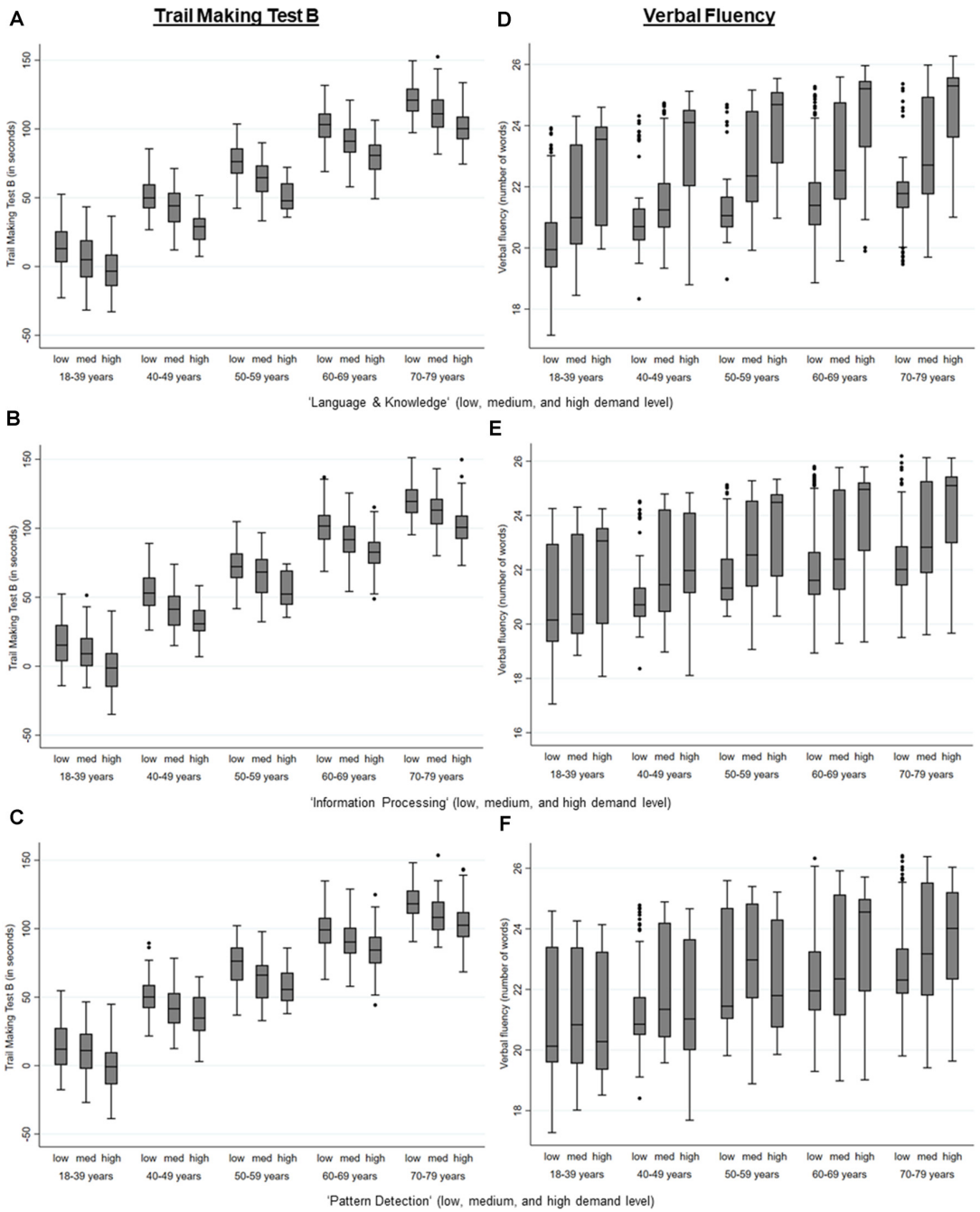

FIGURE 1 | Performance in the Trail Making Test B and the Verbal fluency test by the level of mental demands at work (MDW) and age, estimated *via* regression analyses separate for those that are retired and those that are not retired an adjusted for age, gender, education, APOE e4-allele, hypertension, and diabetes. **(A)** Level of the mental work demand, Language & Knowledge on performance in the Trail Making Test B. **(B)** Level of the mental work demand, Information Processing on performance in the Trail Making Test B. **(C)** Level of the mental work demand, Pattern Detection on performance in the Trail Making Test B. **(D)** Level of the mental work demand, Language & Knowledge on performance in the Verbal Fluency Test. **(E)** Level of the mental work demand, Information Processing on performance in the Verbal Fluency Test. **(F)** Level of the mental work demand, Pattern Detection on performance in the Verbal Fluency Test.

diffusion-weighted imaging and resting-state BOLD-fMRI has been implicated in cognitive functions and might be a more sensitive marker of cognitive impairment or dementia (Zhou et al., 2008; Fellgiebel and Yakushev, 2011). For example, the functional coupling of the hippocampus with dorsal sub-regions is essential for learning (Mattfeld and Stark, 2015) and mental demands at work that would strengthen functional connectivity of the hippocampus could therefore help to maintain learning

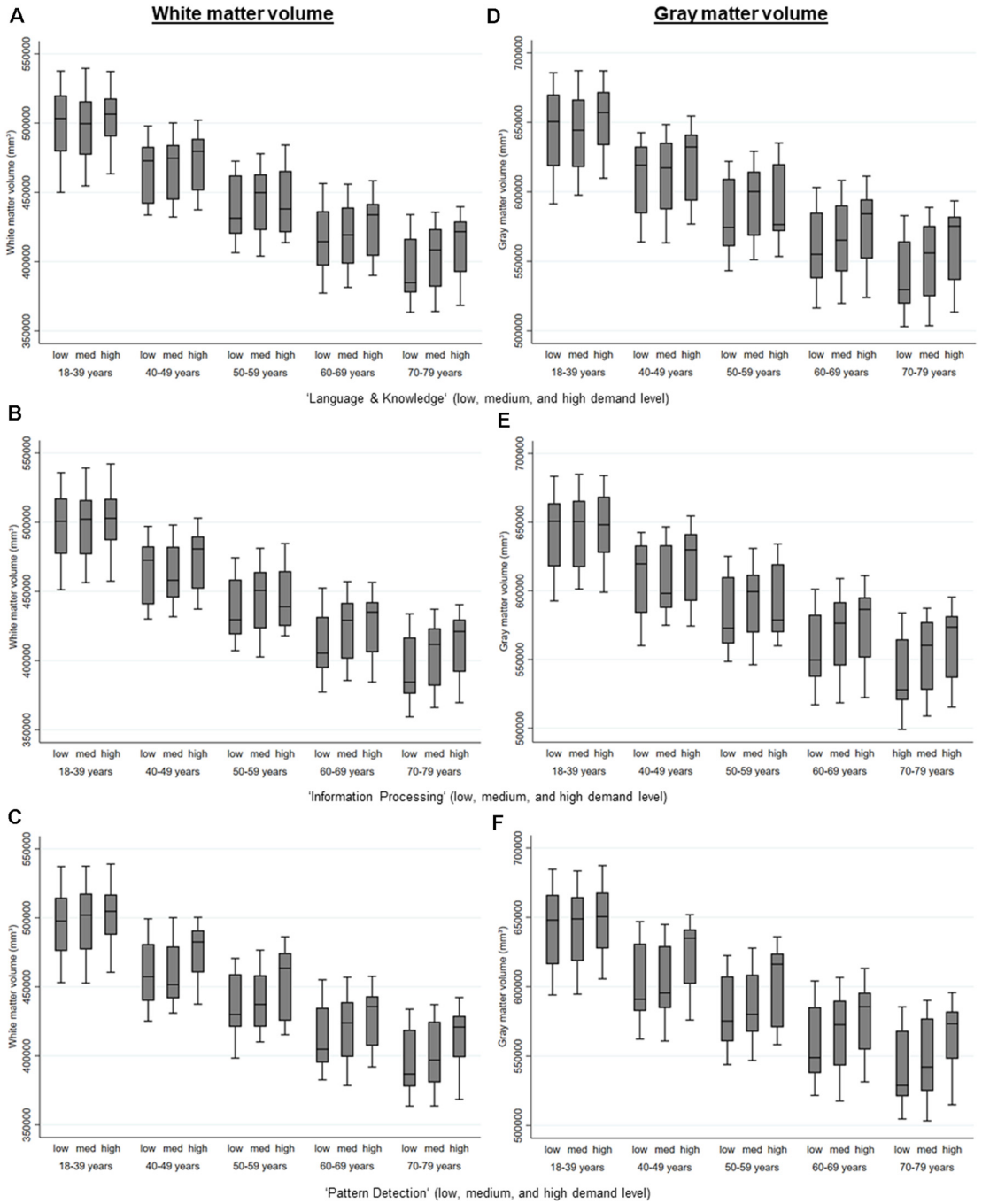

FIGURE 2 | White matter and gray matter volume by the level of MDW and age, estimated *via* regression analyses separate for those that are retired and those that are not retired and adjusted for age, gender, education, APOE e4-allele, hypertension, and diabetes. **(A)** Level of the mental work demand, Language & Knowledge on white matter volume. **(B)** Level of the mental work demand, Information Processing on white matter volume. **(C)** Level of the mental work demand, Pattern Detection on white matter volume. **(D)** Level of the mental work demand, Language & Knowledge on gray matter volume. **(E)** Level of the mental work demand, Information Processing on gray matter volume. **(F)** Level of the mental work demand, Pattern Detection on gray matter volume.

abilities longer in life. As cognitive training has shown to promote hippocampal functioning (Rosen et al., 2011; Kirchhoff et al., 2012), likely, working longer in a mentally stimulating job could at least maintain hippocampal functionality for a longer life period.

Findings from our study emphasize that higher MDW, especially in "language and knowledge," "information processing," and "creativity," are associated with better performance in cognitive tests. This confirms a possible protective association between higher levels of those MDW

TABLE 4 | Results of the regression analyses on the association between MDW and hippocampal and brain volume, separate for those that are retired and those that are not retired, adjusted for age, gender, education, APOE e4-allele, hypertension, and diabetes.

	Unadjusted				Adjusted [1]			
	Retired		Not retired		Retired[1]		Not retired[1]	
Hippocampal volume[2]								
	***b* (CI 95%)**	***P***	***b* (CI 95%)**	***P***	***b* (CI 95%)**	***P***	***b* (CI 95%)**	***P***
MDW language and knowledge	−65.5 (−156.5; 25.6)	0.16	61.1 (−26.9; 148.9)	0.17	−16.2 (−91.1; 58.7)	0.671	59.4 (−10.9; 129.7)	0.098
MDW information processing	−64.9 (−136.3; 6.3)	0.07	11.3 (−54.9; 77.6)	0.74	−10.1 (−86.5; 66.3)	0.795	−10.7 (−82.6; 61.2)	0.771
MDW mathematics	−37.5 (−88.9; 13.9)	0.15	7.5 (−48.7; 63.8)	0.79	1.6 (−51.1; 54.3)	0.952	−4.1 (−63.4; 55.3)	0.893
MDW pattern detection	−48.5 (−145.7; 48.6)	0.33	33.4 (−55.4; 122.3)	0.46	43.5 (−53.5; 140.5)	0.379	1.4 (−90.9; 93.7)	0.976
MDW creativity	−69.9 (−114.9; -24.9)	**0.002**	25.5 (−21.6; 72.6)	0.29	−21.1 (−69.8; 27.7)	0.396	34.4 (−13.9; 82.7)	0.162
Hippocampal subfields[2,3] CA2&3								
MDW pattern detection	123.3 (22.5; 224.2)	**0.017**	17.6 (−87.5; 122.8)	0.742	131.1 (22.9; 239.4)	**0.018**	−35.9 (−143.8; 72.1)	0.514
White matter volume								
MDW language and knowledge	8,242.8 (3,495.2; 12,990.4)	**0.001**	9,563.9 (2,126.8; 17,000.9)	**0.012**	3,773.7 (−323.7; 7,871.0)	0.071	8,356.7 (3,285.4; 13,428.1)	**0.001**
MDW information processing	9,657.0 (5,903.1; 13,410.9)	**<0.001**	8,781.2 (2,864.1; 14,698.4)	**0.004**	5,089.9 (983.9; 9,195.8)	**0.015**	5,680.8 (106.3; 11,255.3)	**0.046**
MDW mathematics	3,133.7 (410.6; 5,856.7)	**0.024**	6,004.3 (1,982.5; 10,026.0)	**0.003**	628.5 (−2,187.6; 3,444.5)	0.661	1,915.3 (−2,294.8; 6,125.5)	0.372
MDW pattern detection	11,809.8 6,388.5; 17,231.2)	**<0.001**	11,556.2 (4,108.8; 19,003.7)	**0.002**	5,491.9 (72.6; 10911.4)	**0.047**	5,923.1 (−1005.9; 12852.1)	0.094
MDW creativity	5,142.6 (2,787.8; 7,497.4	**<0.001**	7,603.6 (4,137.4; 11,069.8)	**<0.001**	1,694.5 (−925.4; 4,314.2)	0.205	3,827.9 (527.7; 7,128.1)	**0.023**
Gray matter volume								
MDW language and knowledge	11,172.7 (6,316.2; 16,029.3)	**<0.001**	11,125.8 (3,879.4; 18,372.2)	**0.003**	5,435.6 (1,252.9; 9,618.4)	**0.011**	6,654.2 (2,033.8; 11,274.7)	**0.005**
MDW information processing	11,376.7 (7,482.1; 15,271.4)	**<0.001**	9,549.5 (3,637.8; 15,461.2)	**0.002**	4,847.6 (508.1; 9,187.2)	**0.029**	4,728.1 (14.1; 9,442.1)	**0.049**
MDW mathematics	5,328.8 (2,706.5; 7,951.2)	**<0.001**	7,697.3 (3,159.7; 12,234.9)	**0.001**	1,821.6 (−956.4; 4,599.5)	0.198	3,222.3 (−522.4; 6,967.1)	0.092
MDW pattern detection	15,126.9 (9,957.7; 20,296.2)	**<0.001**	12,025.1 (4,776.6; 19,273.6)	**0.001**	6029.2 (732.8; 11,325.6)	**0.026**	4,171.3 (−1,699.7; 10,042.3)	0.163
MDW creativity	7,175.6 (4,732.4; 9,618.8)	**<0.001**	8,178.6 (4,126.6; 12,330.6)	**<0.001**	2,294.5 (−432.3; 5,021.4)	0.099	3,895.6 (657.0; 7,134.2)	**0.018**
ICV								
MDW language and knowledge	40,294.2 (25,402.5; 55,185.9)	**<0.001**	21,223.4 (3,182.7; 39,264.1)	**0.021**	12,026.3 (−330.9; 24,383.4)	0.056	5,435.6 (1,252.9; 9,618.4)	**0.011**
MDW information processing	39,658.5 (27,516.9; 51,800.2)	**<0.001**	20,100.5 (5,650.9; 34,550.0)	**0.006**	13,828.0 (790.4; 26,865.6)	**0.038**	10,287.7 (−2,802.9; 23,378.3)	0.123
MDW mathematics	17,967.4 (9,707.6; 26,227.4)	**<0.001**	17,513.9 (6,801.5; 28,226.4)	**0.001**	2,385.2 (−5994.1; 10,764.5)	0.576	6,324.8 (−4,400.9; 17,050.6)	0.247
MDW pattern detection	49,191.9 (32,572.5; 65811.2)	**<0.001**	23,164.2 (5,368.9; 40,959.4)	**0.011**	11287.2 (−5,167.2; 27,741.5)	0.178	5,461.1 (−11,188.2; 22,110.4)	0.520
MDW creativity	25,298.9 (17,895.6; 32,702.3)	**<0.001**	21,021.4 (11,754.3; 30,288.4)	**<0.001**	4,554.3 (−3,702.4; 12,811.1)	0.279	6,743.5 (−1,869.7; 15,368.8)	0.125

Notes: 1, model adjusted for age, gender, education; 2, whole hippocampal volume in mm^3 including left and right hippocampi; 3, significant findings only; APOE e4-allele, hypertension, and diabetes; b, regression coefficient; CA, cornu ammonis; CI, confidence interval 95%; ICV, intracranial volume; MDW, mental demands at work; P, level of significance. Bold indicates significance <0.05.

TABLE 5 | Results of the structural equation modeling on the association of MDW and hippocampal and brain volume (HBV) on cognition (COG).

	Retired				Not retired			
	Loading	***P***			**Loading**	***P***		
Mental demands at work (MDW)								
MDW language and knowledge	0.898	<0.001			0.891	<0.001		
MDW information processing	0.908	<0.001			0.912	<0.001		
MDW mathematics	0.873	<0.001			0.840	<0.001		
MDW pattern detection	0.826	<0.001			0.844	<0.001		
MDW creativity	0.886	<0.001			0.867	<0.001		
Hippocampal and brain volume (HBV)[1]								
Hippocampal volume	−0.085	0.026			0.460	<0.001		
White matter	0.900	<0.001			0.821	<0.001		
Gray matter	0.922	<0.001			0.944	<0.001		
Intracranial volume	0.959	<0.001			0.862	<0.001		
Cognition (COG)								
TMT B	0.776	<0.001			0.799	<0.001		
TMT A	0.539	<0.001			0.785	<0.001		
Verbal fluency	0.795	<0.001			0.514	<0.001		
Word list	0.208	<0.001			0.635	<0.001		
Covariates								
Education	0.562	<0.001			0.031	0.447		
Age	0.037	0.336			0.860	<0.001		
Gender	0.923	<0.001			0.344	<0.001		
APOE e4	0.045	0.236			0.147	<0.001		
Diabetes	0.085	0.026			0.341	<0.001		
Hypertension	0.059	0.121			0.552	<0.001		
	b	***P***	**Indirect (*b*)**	**Total (*b*)**	***b***	***p***	**Indirect (*b*)**	**Total (*b*)**
MDW > HBV	0.001	0.988			0.109	0.002		
MDW > COG	0.173	<0.001	0.000	0.173	0.100	0.008	−0.000	0.100
HBV > COG	0.080	0.082			−0.001	0.989		
Covariates > HBV	−0.580	<0.001			−0.503	<0.001		
Covariates > COG	−0.031	0.523	−0.046	−0.077	−0.424	<0.001	0.000	−0.423

Notes: 1, whole hippocampal volume in mm^3 including left and right hippocampi; b, regression coefficient; CA, cornu ammonis; CI, confidence interval 95%; ICV, intracranial volume; MDW, mental demands at work; P, level of significance; SE, standard error; TMT, trail making test.

and better cognitive health in old age, as a previous study has shown concerning dementia risk (Then et al., 2015). For those who were retired, only the association with executive cognitive abilities (Trail Making Test B, Verbal Fluency) retained significance, indicating that the effect might be stronger during the active workforce participation. Indeed, previous studies have shown that older age of retirement was associated with a reduced dementia risk (Dufouil et al., 2014) and that high complexity at work seems to facilitate cognitive functioning before retirement (Finkel et al., 2009). This observation is in line with the "use-it-or-lose-it" theory, which states that the active use of cognitive abilities prevents their deterioration in old age (Salthouse, 2006). Nonetheless, there is always the possibility of reverse causality so that individuals with more resilience to cognitive decline work in jobs with higher demands. Even though a great number of previous studies demonstrated that the training of cognitive abilities delays cognitive decline (Rizkalla, 2018). Irrespective of whether innate resilience or training effects establish a protective association between work demands and cognitive health in old age, it seems that the use of semantic long-term memory involved in complex information processing tasks plays a major role. There is no research so far explaining this observation.

As "language and knowledge" and "information processing," identified as possibly protective MDW in a previous study (Then et al., 2015), were both associated with gray matter in our analysis, it is conceivable that their effect on cognitive performance might be mediated through their effect on gray matter. In previous studies, cognitive training has been shown to lead to changes in gray matter (Ceccarelli et al., 2009; Kühn et al., 2014), which could counteract the deterioration of gray matter in aging (Alexander et al., 2006) and Alzheimer's disease (Karas et al., 2004). However, in our study, there was no indirect path to cognitive functioning. Accordingly, there seem to be independent associations between higher demands and brain volume and higher demands and cognitive performance as well as between brain volume and cognitive performance. A lack of association between brain and cognitive functioning in individuals with high intellectual stimulation during their life-course is commonly referred to as cognitive reserve (Stern, 2002) and is a well-known phenomenon. Hence, it is likely that intellectual stimulation at work, especially in the MDW "language and knowledge," "information processing," and "creativity," take a cognitive reserve-effect, in addition to effects that they have on the brain.

While it is intuitive to see high demands involving language and knowledge and more complex information processing as cognitive training at work, this understanding does not come so obviously for creativity. In our findings, there were associations

between higher MDW "creativity" and more gray matter as well as white matter volume among individuals that were still working. It also seems that low-educated individuals can benefit from the effects of MDW "creativity." There is not yet any consensus on how creative activity shows in or alters the brain (Arden et al., 2010). It might be domain-specific (Boccia et al., 2015); norepinephrine (Heilman, 2016) and the functional connectivity of the default mode network might play a role (Beaty et al., 2014). Thus, more research is necessary to understand potential effects.

Our study is not without limitations. First, it is only cross-sectional. Longitudinal analyses would reveal changes in hippocampal volume in each individual over time—this would give more information on a possible causal pathway. Moreover, we did not include clinical dementia cases. Our study might therefore underestimate the true effect size in the general elderly population. Further, even though we adjusted for important socioeconomic and lifestyle factors, there might be others affecting the results, which we did not consider. Therefore, novel model prediction statistics that are better capable to account for in-depths individual differences, collinearities, and multiple data points might be used in future analyses of large MRI datasets (Bzdok and Yeo, 2017). Finally, 1 mm^3 isotropic data lacks sufficient resolution and contrast for the visualization of the internal structure of the hippocampus (Elman et al., 2019; Wisse et al., 2020) so that the estimates for hippocampal subfields may not be as optimal as desired. Further studies are necessary to confirm our findings for the subfields. Strengths of our study include the large sample size, the well-defined population under study, and the state-of-the-art 3T neuroimaging protocol.

CONCLUSIONS

In trying to gain a better understanding of how mentally demanding activities at work (MDW) influence dementia risk, our analysis explored whether high demands could maintain or even enhance the plasticity of the hippocampus. Our findings do not support the notion that MDW protects cognitive health *via* hippocampal volume, neither *via* brain volume. Yet, we observed associations between higher demands in "language and knowledge," "information processing," and "creativity" at work on larger white and gray matter volume and better cognitive functioning. "Creativity," however, seems to be more relevant for individuals that are still actively working; the protective effect might disappear when the person retires. Among retired individuals, on the other hand, higher demands in "pattern detection" were associated with larger white matter volume as well as larger hippocampal subfields CA2/CA3, suggesting a retention effect later in life. This finding is so important because a previous study has identified this MDW as well as the other three previously mentioned MDWs to be associated with a lower dementia risk (Then et al., 2015). At this point, it is unclear what the biological pathways are connecting them to cognitive health in old age. Further studies are necessary to evaluate how individuals who have high levels of mental demands at work gain resilience against dementia.

DATA AVAILABILITY STATEMENT

The datasets of the "Adult Study" of the Leipzig Research Centre for Civilization Diseases (LIFE) are available upon request from the authors. Requests to access the datasets should be directed to Christoph Engel, christoph.engel@imise.uni-leipzig.de.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of the Medical Faculty of the University of Leipzig. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

CE, ML, JT, AV, and SR-H contributed to the conception and design of the study. FR, SH, MS, CE, RBa, RBu, TL, and AW contributed to the data collection, data preparation, and data quality control. FR performed the statistical analysis and wrote the first draft. SH, WV, MA, and AW interpreted the data and/or wrote sections of the draft. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the LIFE—Leipzig Research Center for Civilization Diseases, Universität Leipzig, which was funded through the European Social Fund and the Free State of Saxony, LIFE-103 P1. The funders did not influence the content of the study or the analysis of the data. FR reports a grant from the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG): grant # TH2137/3-1.

ACKNOWLEDGMENTS

We thank all the study participants and the research teams involved in making this study possible.

REFERENCES

Aebi, C., Monsch, A. U., Berres, M., Brubacher, D., and Staehelin, H. B. (2002). Validation of the German CERAD-neuropsychological assessment battery. *Neurobiol. Aging* 23, S27–S28.

Alexander, G. E., Chen, K., Merkley, T. L., Reiman, E. M., Caselli, R. J., Aschenbrenner, M., et al. (2006). Regional network of magnetic resonance imaging gray matter volume in healthy aging. *Neuroreport* 17, 951–956. doi: 10.1097/01.wnr.0000220135.16844.b6

Andel, R., Crowe, M., Pedersen, N. L., Mortimer, J., Crimmins, E., Johansson, B., et al. (2005). Complexity of work and risk of Alzheimer's disease: a population-based study of Swedish twins. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 60, P251–P258. doi: 10.1093/geronb/60.5.p251

Arbuthnott, K., and Frank, J. (2010). Trail making test, part B as a measure of executive control: validation using a set-switching paradigm. *J. Clin.*

Exp. Neuropsychol. 22, 518–528. doi: 10.1076/1380-3395(200008)22:4;1-0;FT518

Arden, R., Chavez, R. S., Grazioplene, R., and Jung, R. E. (2010). Neuroimaging creativity: a psychometric view. *Behav. Brain Res.* 214, 143–156. doi: 10.1016/j.bbr.2010.05.015

Aslanidis, C., and Schmitz, G. (1999). High-speed apolipoprotein E genotyping and apolipoprotein B3500 mutation detection using real-time fluorescence PCR and melting curves. *Clin. Chem.* 45, 1094–1097.

Beaty, R. E., Benedek, M., Wilkins, R. W., Jauk, E., Fink, A., Silvia, P. J., et al. (2014). Creativity and the default network: a functional connectivity analysis of the creative brain at rest. *Neuropsychologia* 64, 92–98. doi: 10.1016/j.neuropsychologia.2014.09.019

Boccia, M., Piccardi, L., Palermo, L., Nori, R., and Palmiero, M. (2015). Where do bright ideas occur in our brain? Meta-analytic evidence from neuroimaging studies of domain-specific creativity. *Front. Psychol.* 6:1195. doi: 10.3389/fpsyg.2015.01195

Boots, E. A., Schultz, S. A., Almeida, R. P., Oh, J. M., Koscik, R. L., Dowling, M. N., et al. (2015). Occupational complexity and cognitive reserve in a middle-aged cohort at risk for Alzheimer's disease. *Arch. Clin. Neuropsychol.* 30, 634–642. doi: 10.1093/arclin/acv041

Brickman, A. M., Khan, U. A., Provenzano, F. A., Yeung, L. K., Suzuki, W., Schroeter, H., et al. (2014). Enhancing dentate gyrus function with dietary flavanols improves cognition in older adults. *Nat. Neurosci.* 17, 1798–1803. doi: 10.1038/nn.3850

Bzdok, D., and Yeo, B. T. T. (2017). Inference in the age of big data: future perspectives on neuroscience. *NeuroImage* 155, 549–564. doi: 10.1016/j.neuroimage.2017.04.061

Ceccarelli, A., Rocca, M. A., Pagani, E., Falini, A., Comi, G., and Filippi, M. (2009). Cognitive learning is associated with gray matter changes in healthy human individuals: a tensor-based morphometry study. *NeuroImage* 48, 585–589. doi: 10.1016/j.neuroimage.2009.07.009

den Heijer, T., van der Lijn, F., Koudstaal, P. J., Hofman, A., van der Lugt, A., Krestin, G. P., et al. (2010). A 10-year follow-up of hippocampal volume on magnetic resonance imaging in early dementia and cognitive decline. *Brain* 133, 1163–1172. doi: 10.1093/brain/awq048

Dufouil, C., Pereira, E., Chene, G., Glymour, M. M., Alperovitch, A., Saubusse, E., et al. (2014). Older age at retirement is associated with decreased risk of dementia. *Eur. J. Epidemiol.* 29, 353–361. doi: 10.1007/s10654-014-9906-3

Duzel, E., van Praag, H., and Sendtner, M. (2016). Can physical exercise in old age improve memory and hippocampal function? *Brain* 139, 662–673. doi: 10.1093/brain/awv407

Elman, J. A., Panizzon, M. S., Gillespie, N. A., Hagler, D. J., Fennema-Notestine, C., Eyler, L. T., et al. (2019). Genetic architecture of hippocampal subfields on standard resolution MRI: how the parts relate to the whole. *Hum. Brain Mapp.* 40, 1528–1540. doi: 10.1002/hbm.24464

Engvig, A., Fjell, A. M., Westlye, L. T., Skaane, N. V., Dale, A. M., Holland, D., et al. (2014). Effects of cognitive training on gray matter volumes in memory clinic patients with subjective memory impairment. *J. Alzheimers Dis.* 41, 779–791. doi: 10.3233/JAD-131889

Erickson, K. I., Voss, M. W., Prakash, R. S., Basak, C., Szabo, A., Chaddock, L., et al. (2011). Exercise training increases size of hippocampus and improves memory. *Proc. Natl. Acad. Sci. U S A* 108, 3017–3022. doi: 10.1073/pnas.1015950108

Fellgiebel, A., and Yakushev, I. (2011). Diffusion tensor imaging of the hippocampus in MCI and early Alzheimer's disease. *J. Alzheimers Dis.* 26, 257–262. doi: 10.3233/JAD-2011-0001

Finkel, D., Andel, R., Gatz, M., and Pedersen, N. L. (2009). The role of occupational complexity in trajectories of cognitive aging before and after retirement. *Psychol. Aging* 24, 563–573. doi: 10.1037/a0015511

Fjell, A. M., McEvoy, L., Holland, D., Dale, A. M., and Walhovd, K. B. (2014). What is normal in normal aging? Effects of aging, amyloid and Alzheimer's disease on the cerebral cortex and the hippocampus. *Prog. Neurobiol.* 117, 20–40. doi: 10.1016/j.pneurobio.2014.02.004

Fraser, M. A., Shaw, M. E., and Cherbuin, N. (2015). A systematic review and meta-analysis of longitudinal hippocampal atrophy in healthy human ageing. *NeuroImage* 112, 364–374. doi: 10.1016/j.neuroimage.2015.03.035

Gold, S. M., Dziobek, I., Sweat, V., Tirsi, A., Rogers, K., Bruehl, H., et al. (2007). Hippocampal damage and memory impairments as possible early brain complications of type 2 diabetes. *Diabetologia* 50, 711–719. doi: 10.1007/s00125-007-0602-7

Griswold, M. A., Jakob, P. M., Heidemann, R. M., Nittka, M., Jellus, V., Wang, J. M., et al. (2002). Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn. Reson. Med.* 47, 1202–1210. doi: 10.1002/mrm.10171

Heilman, K. M. (2016). Possible brain mechanisms of creativity. *Arch. Clin. Neuropsychol.* 31, 285–296. doi: 10.1093/arclin/acw009

Karas, G., Scheltens, P., Rombouts, S., Visser, P., Van Schijndel, R., Fox, N., et al. (2004). Global and local gray matter loss in mild cognitive impairment and Alzheimer's disease. *NeuroImage* 23, 708–716. doi: 10.1016/j.neuroimage.2004.07.006

Karp, A., Andel, R., Parker, M. G., Wang, H. X., Winblad, B., and Fratiglioni, L. (2009). Mentally stimulating activities at work during midlife and dementia risk after age 75: follow-up study from the Kungsholmen Project. *Am. J. Geriatr. Psychiatry* 17, 227–236. doi: 10.1097/JGP.0b013e318190b691

Kaup, A. R., Xia, F., Launer, L. J., Sidney, S., Nasrallah, I., Erus, G., et al. (2018). Occupational cognitive complexity in earlier adulthood is associated with brain structure and cognitive health in midlife: the CARDIA study. *Neuropsychology* 32, 895–905. doi: 10.1037/neu0000474

Kaye, J. A., Swihart, T., Howieson, D., Dame, A., Moore, M. M., Karnos, T., et al. (1997). Volume loss of the hippocampus and temporal lobe in healthy elderly persons destined to develop dementia. *Neurology* 48, 1297–1304. doi: 10.1212/wnl.48.5.1297

Kerti, L., Witte, A. V., Winkler, A., Grittner, U., Rujescu, D., and Floel, A. (2013). Higher glucose levels associated with lower memory and reduced hippocampal microstructure. *Neurology* 81, 1746–1752. doi: 10.1212/01.wnl.0000435561.00234.ee

Kirchhoff, B. A., Anderson, B. A., Smith, S. E., Barch, D. M., and Jacoby, L. L. (2012). Cognitive training-related changes in hippocampal activity associated with recollection in older adults. *NeuroImage* 62, 1956–1964. doi: 10.1016/j.neuroimage.2012.06.017

Kraan, C., Stolwyk, R. J., and Testa, R. (2013). The abilities associated with verbal fluency performance in a young, healthy population are multifactorial and differ across fluency variants. *Appl. Neuropsychol. Adult.* 20, 159–168. doi: 10.1080/09084282.2012.670157

Kühn, S., Gleich, T., Lorenz, R. C., Lindenberger, U., and Gallinat, J. (2014). Playing Super Mario induces structural brain plasticity: gray matter changes resulting from training with a commercial video game. *Mol. Psychiatry* 19, 265–271. doi: 10.1055/a-1186-0441

Loeffler, M., Engel, C., Ahnert, P., Alfermann, D., Arelin, K., Baber, R., et al. (2015). The LIFE-Adult-Study: objectives and design of a population-based cohort study with 10,000 deeply phenotyped adults in Germany. *BMC Public Health* 15:691. doi: 10.1186/s12889-015-1983-z

Mattfeld, A. T., and Stark, C. E. (2015). Functional contributions and interactions between the human hippocampus and subregions of the striatum during arbitrary associative learning and memory. *Hippocampus* 25, 900–911. doi: 10.1002/hipo.22411

Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., et al. (1989). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology* 39, 1159–1165. doi: 10.1212/wnl.39.9.1159

Noble, K. G., Grieve, S. M., Korgaonkar, M. S., Engelhardt, L. E., Griffith, E. Y., Williams, L. M., et al. (2012). Hippocampal volume varies with educational attainment across the life-span. *Front. Hum. Neurosci.* 6:307. doi: 10.3389/fnhum.2012.00307

Plassman, B. L., Welsh-Bohmer, K. A., Bigler, E. D., Johnson, S. C., Anderson, C. V., Helms, M. J., et al. (1997). Apolipoprotein E epsilon 4 allele and hippocampal volume in twins with normal cognition. *Neurology* 48, 985–989. doi: 10.1212/wnl.48.4.985

Potter, G. G., Helms, M. J., Burke, J. R., Steffens, D. C., and Plassman, B. L. (2007). Job demands and dementia risk among male twin pairs. *Alzheimers Dement.* 3, 192–199. doi: 10.1016/j.jalz.2007.04.377

Pruessner, J. C., Collins, D. L., Pruessner, M., and Evans, A. C. (2001). Age and gender predict volume decline in the anterior and posterior hippocampus in

early adulthood. *J. Neurosci.* 21, 194–200. doi: 10.1523/JNEUROSCI.21-01-00194.2001

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., et al. (2005). Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. *Cereb. Cortex* 15, 1676–1689. doi: 10.1093/cercor/bhi044

Rizkalla, M. N. (2018). Cognitive training in the elderly: a randomized trial to evaluate the efficacy of a self-administered cognitive training program. *Aging Ment. Health* 22, 1384–1394. doi: 10.1080/13607863.2015.1118679

Rosen, A. C., Sugiura, L., Kramer, J. H., Whitfield-Gabrieli, S., and Gabrieli, J. D. (2011). Cognitive training changes hippocampal function in mild cognitive impairment: a pilot study. *J. Alzheimers Dis.* 26, 349–357. doi: 10.3233/JAD-2011-0009

Salthouse, T. A. (2006). Mental exercise and mental aging evaluating the validity of the "use it or lose it" hypothesis. *Perspect. Psychol. Sci.* 1, 68–87. doi: 10.1111/j.1745-6916.2006.00005.x

Salthouse, T. A. (2011). What cognitive abilities are involved in trail-making performance? *Intelligence* 39, 222–232. doi: 10.1016/j.intell.2011.03.001

Shih, Y.-H., Tsai, S.-F., Huang, S.-H., Chiang, Y.-T., Hughes, M. W., Wu, S.-Y., et al. (2016). Hypertension impairs hippocampus-related adult neurogenesis, CA1 neuron dendritic arborization and long-term memory. *Neuroscience* 322, 346–357. doi: 10.1016/j.neuroscience.2016.02.045

Shpanskaya, K. S., Choudhury, K. R., Hostage, C. Jr., Murphy, K. R., Petrella, J. R., Doraiswamy, P. M., et al. (2014). Educational attainment and hippocampal atrophy in the Alzheimer's disease neuroimaging initiative cohort. *J. Neuroradiol.* 41, 350–357. doi: 10.1016/j.neurad.2013.11.004

Smyth, K. A., Fritsch, T., Cook, T. B., McClendon, M. J., Santillan, C. E., and Friedland, R. P. (2004). Worker functions and traits associated with occupations and the development of AD. *Neurology* 63, 498–503. doi: 10.1212/01.wnl.0000133007.87028.09

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *J. Int. Neuropsychol. Soc.* 8, 448–460. doi: 10.1017/s1355617702813248

Then, F. S., Luck, T., Luppa, M., Konig, H. H., Angermeyer, M. C., and Riedel-Heller, S. G. (2015). Differential effects of enriched environment at work on cognitive decline in old age. *Neurology* 84, 2169–2176. doi: 10.1212/WNL.0000000000001605

Valenzuela, M. J., and Sachdev, P. (2006a). Brain reserve and cognitive decline: a non-parametric systematic review. *Psychol. Med.* 36, 1065–1073. doi: 10.1017/S0033291706007744

Valenzuela, M. J., and Sachdev, P. (2006b). Brain reserve and dementia: a systematic review. *Psychol. Med.* 36, 441–454. doi: 10.1017/S0033291705006264

Van Leemput, K., Bakkour, A., Benner, T., Wiggins, G., Wald, L. L., Augustinack, J., et al. (2009). Automated segmentation of hippocampal subfields from ultra-high resolution *in vivo* MRI. *Hippocampus* 19, 549–557. doi: 10.1002/hipo.20615

Wang, J., He, L., Zheng, H., and Lu, Z. L. (2014). Optimizing the magnetization-prepared rapid gradient-echo (MP-RAGE) sequence. *PLoS One* 9:e96899. doi: 10.1371/journal.pone.0096899

Wisse, L. E. M., Chételat, G., Daugherty, A. M., de Flores, R., la Joie, R., Mueller, S. G., et al. (2020). Hippocampal subfield volumetry from structural isotropic 1 mm3 MRI scans: a note of caution. *Hum. Brain Mapp.* doi: 10.1002/hbm.25234 [Epub ahead of print].

Zhou, Y., Dougherty, J. H. Jr., Hubner, K. F., Bai, B., Cannon, R. L., and Hutson, R. K. (2008). Abnormal connectivity in the posterior cingulate and hippocampus in early Alzheimer's disease and mild cognitive impairment. *Alzheimers Dement.* 4, 265–270. doi: 10.1016/j.jalz.2008.04.006

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 03 February 2021
doi: 10.3389/fnagi.2021.625931

Differential Patterns of Gyral and Sulcal Morphological Changes During Normal Aging Process

Hsin-Yu Lin[1,2†], *Chu-Chung Huang*[2,3*†], *Kun-Hsien Chou*[2,4], *Albert C. Yang*[5,6], *Chun-Yi Zac Lo*[7], *Shih-Jen Tsai*[5,8] *and Ching-Po Lin*[2,4*]

[1] *Centre for Research and Development in Learning, Nanyang Technological University, Singapore, Singapore,* [2] *Institute of Neuroscience, National Yang-Ming University, Taipei, Taiwan,* [3] *School of Psychology and Cognitive Science, East China Normal University, Institute of Cognitive Neuroscience, Shanghai, China,* [4] *Brain Research Center, National Yang-Ming University, Taipei, Taiwan,* [5] *Department of Psychiatry, Taipei Veterans General Hospital, Taipei, Taiwan,* [6] *Division of Interdisciplinary Medicine and Biotechnology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, United States,* [7] *Institute of Science and Technology for Brain Inspired Intelligence, Fudan University, Shanghai, China,* [8] *School of Medicine, National Yang-Ming University, Taipei, Taiwan*

The cerebral cortex is a highly convoluted structure with distinct morphologic features, namely the gyri and sulci, which are associated with the functional segregation or integration in the human brain. During the lifespan, the brain atrophy that is accompanied by cognitive decline is a well-accepted aging phenotype. However, the detailed patterns of cortical folding change during aging, especially the changing age-dependencies of gyri and sulci, which is essential to brain functioning, remain unclear. In this study, we investigated the morphology of the gyral and sulcal regions from pial and white matter surfaces using MR imaging data of 417 healthy participants across adulthood to old age (21–92 years). To elucidate the age-related changes in the cortical pattern, we fitted cortical thickness and intrinsic curvature of gyri and sulci using the quadratic model to evaluate their age-dependencies during normal aging. Our findings show that comparing to gyri, the sulcal thinning is the most prominent pattern during the aging process, and the gyrification of pial and white matter surfaces were also affected differently, which implies the vulnerability of functional segregation during aging. Taken together, we propose a morphological model of aging that may provide a framework for understanding the mechanisms underlying gray matter degeneration.

Keywords: aging, gyrification, cortical thickness, intrinsic curvature, gyri and sulci

OPEN ACCESS

Edited by:
Tomáš Paus,
University of Toronto, Canada

Reviewed by:
Hugo Schnack,
Utrecht University, Netherlands
Yash Patel,
University of Toronto, Canada

***Correspondence:**
Chu-Chung Huang
czhuang@psy.ecnu.edu.cn
Ching-Po Lin
chingpolin@gmail.com

[†] These authors have contributed equally to this work

Received: *04 November 2020*
Accepted: *08 January 2021*
Published: *03 February 2021*

Citation:
Lin H-Y, Huang C-C, Chou K-H, Yang AC, Lo C-YZ, Tsai S-J and Lin C-P (2021) Differential Patterns of Gyral and Sulcal Morphological Changes During Normal Aging Process. Front. Aging Neurosci. 13:625931. doi: 10.3389/fnagi.2021.625931

INTRODUCTION

The morphology and function of the human brain change throughout the aging process. However, the mechanisms underlying how the structures of gyri and sulci are altered with age remain unclear. In recent decades, *in vivo* magnetic resonance imaging (MRI) has been widely utilized to investigate the effects of aging on the human brain (Good et al., 2001; Jernigan et al., 2001; Raz et al., 2004; Salat et al., 2004), and these studies have provided information on how the structures of the human brain change during the course of a lifespan. The decrease in cortical volume is the most dramatic change that occurs during aging (Scahill et al., 2003). Nevertheless, the highly convoluted and complex nature of the cerebral cortex implies aspects such as folding and thickness of the cortex may have different influences on cortical morphology and brain function. In that case, in addition to volumetric measurements, surface area, gyrification, and thickness measurements also provide

detailed information for brain morphology (Panizzon et al., 2009; Winkler et al., 2010; Gautam et al., 2015). The cerebrum attains its folding structures through a complex orchestrated set of systematic mechanisms including differential proliferation, mechanical buckling, and differential expansion (Richman et al., 1975; Van Essen et al., 2018; Llinares-Benadero and Borrell, 2019). Previous studies have suggested that the development of cortical anatomy is dominated by genetic factors rather than random convolutions (Peng et al., 2016). These genetic factors could contribute to changes of cortical structure throughout the lifespan (Fjell et al., 2015; Ronan and Fletcher, 2015). As such, previous findings suggest that cortical features might also degenerate in a non-random, systematic way.

It is worth noting that the locations of specific gyri and sulci (e.g., central sulcus, visual cortex…) are consistent in different individuals and correspond to regional functions, and these structures are considered structural-functional entities (Brodmann, 1909; Welker, 1990). With this point of view, gyrification and cortical thickness are thought to reflect the functional aspects of the cortex, such as intelligence, behavior complexity, and cognition (Kaas, 2013; Gautam et al., 2015). Gyrification is thought to be a mechanical process that causes the surface to buckle (Ronan et al., 2014; Van Essen et al., 2018), although the primary underlying mechanism is still under debate (Xu et al., 2010; Ronan and Fletcher, 2015; Llinares-Benadero and Borrell, 2019). Traditionally, local gyrification index (LGI) (Schaer et al., 2008) has been used as a proxy to probe the regional folding degree of the human brain (Nanda et al., 2014, Zhang et al., 2014). However, due to the ratio-principle of LGI, it has less sensitivity compared to the intrinsic curvature on describing the complex pattern of cortical surface, especially in the deep brain regions, e.g., the insula (Griffin, 1994; Ronan et al., 2011). It has been suggested that intrinsic curvature is a more sensitive index to describe cortical folding and is presumed to reflect the differential development and cortical connectivity in different areas of the cortex (Ronan et al., 2014). Intrinsic curvature of the pial and white matter surface has also been found to be related to the structural changes of both superficial or deeper layers of cortex (Ronan et al., 2012; Wagstyl et al., 2016). Cortical thickness is a common feature reflecting total neuronal body in the cortex (Nadarajah and Parnavelas, 2002). Sulci, which are thinner than gyri, communicate locally with neighboring structures, while gyri act as functional centers that connect remote gyri to neighboring sulci (Fischl and Dale, 2000; Deng et al., 2014). Stronger functional connectivity was found in gyri than sulci, as supported by a series of experiments (Deng et al., 2014). Previous studies have shown that cognitive performance is associated with pattern changes in regional gyri (Jones et al., 2006; Turner and Spreng, 2012; Gregory et al., 2016), and that the variability of sulcal volume is found to be related to the progression of the neurological diseases in patients (Mega et al., 1998; Sullivan et al., 1998; Im et al., 2008). In the healthy neurodevelopmental process, cortex nonuniformly thins and demonstrates increased thinning in sulci compared to gyri (Vandekar et al., 2015). Moreover, gyri and sulci exhibit opposite trends in curvature changes during aging (Magnotta et al., 1999) and show different kinds of specialized organizations and connectivity (Welker, 1990). Together, the effects of aging on the regional thickness and gyrification of the brain result in decreased cognitive functions (Salat et al., 2004; Gregory et al., 2016). However, sulci are recently deemed to be more functionally segregated in structural connectivity and the increasing age is accompanied by decreasing segregation in large-scale brain systems (Chan et al., 2014; Liu et al., 2017). The morphological evidence for the mechanism underlying such degeneration of the human brain reported by the studies mentioned above and others is not unified. Thus, investigations of distinguishing gyral and sulcal morphology and comparing them are critical.

Previous studies have suggested that the relationship between cortical degeneration and advancing age is not linear, and the degeneration trajectories of cortical features might also differ (Klein et al., 2014; Storsve et al., 2014; Cao et al., 2017). Therefore, understanding the variations in brain morphology and its trajectories throughout the lifetime is a crucial step toward revealing the epicenter of aging-related degeneration of cortical morphology, because the pattern may reflect the configuration of connectivity of the brain (Ronan et al., 2011). By using the image dataset that was collected by a single scanner with a large sample size and a broad age range, this study aimed to investigate the following three issues at both a whole-brain and a regional level: (1) to reassure whether the effects of age on the cortical morphological features are nonlinear across adulthood to old age, (2) whether age differentially and/or systematically affects gyri and sulci, (3) whether the pial and white matter surfaces display distinct variations in their patterns during aging. We believe that measuring detailed morphological features to depict the degenerative pattern of the cerebral cortex could help elucidate underlying degeneration mechanisms.

MATERIALS AND METHODS

Participant Characteristics and Image Data Acquisition

A total of 417 healthy Chinese controls were recruited from northern Taiwan, and their ages ranged from 21 to 92 (Male/Female: 211/206). All of the included participants had sufficient visual and auditory acuity to undergo basic cognitive assessment. This research was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of Taipei Veterans General Hospital. Written informed consent was obtained from all participants after they had been given an overview of this study. A trained research assistant used the diagnostic structured Mini-International Neuropsychiatric Interview (M.I.N.I.) to evaluate each subject (Sheehan et al., 1998). The participants' cognitive functions were assessed using the Mini-Mental State Examination (MMSE) (Folstein et al., 1975), Wechsler Digit Span tasks (forward and backward), and Clinical Dementia Rating (CDR) scale (Hughes et al., 1982) to avoid enrolling participants with possible dementia or cognitive impairments. Subjects who met any of the following exclusion criteria were not enrolled in the study: (1) any Axis I psychiatric diagnoses on the Diagnostic and Statistical Manual of Mental Disorders-IV (American Psychiatric, 2000), such as mood

TABLE 1 | Demographic of included healthy participants.

Demographic variables	Subjects ($N = 417$)
Age (years)	51.90 (18.5)
Gender (male/female)	211/206
Education (years)	13.82 (5.5)
Handedness (left/right)	14/403
MMSE	28.47 (1.6)
Digits Span Forward	13.87 (2.3)
Digits Span Backward	8.29 (3.7)
TIV (cm^3)	1352.8 (130.58)
GMV (cm^3)	659.2 (77.84)
WMV (cm^3)	422.6 (51.23)
CSF (cm^3)	271.1 (92)

The variables are demonstrated as mean (standard deviation).
MMSE Mini-Mental Status Examination/TIV Total intracranial volume.
GMV Gray matter volume/WMV White matter volume/CSF Cerebrospinal fluid.

disorders or psychotic disorders; (2) any neurological disorders, such as dementia, head injury, stroke, or Parkinson's disease; (3) illiterate; (4) an MMSE score less than 24; and/or (5) elderly (age of 65 or over) with a CDR score over 0.5. The demographic information for the subjects is listed in **Table 1**.

All MRI scans were performed on a 3 Tesla Siemens scanner (MAGNETOM Trio TIM system, Siemens AG, Erlangen, Germany) with a 12-channel head coil at National Yang-Ming University. High-resolution structural T1-weighted (T1w) MRI scans were acquired with a three-dimensional magnetization-prepared rapid gradient echo sequence (repetition/echo time = 2,530/3.5, inversion time = 1,100 ms, field of view = 256 mm, flip angle =7°, matrix size = 256 × 256, 192 sagittal slices, isotropic voxel size= 1 mm^3, and no gap). Each participant's head was immobilized with cushions inside the coil to minimize the generation of motion artifacts during image acquisition.

Cortical Reconstruction

The pial and white matter surfaces of each subject were reconstructed using an automated cortical surface reconstruction approach in FreeSurfer 5.3 (http://surfer.nmr.mgh.harvard.edu) according to the following steps: registration, skull-stripping, segmentation of the gray matter, white matter and cerebrospinal fluid, tessellation of the gray-white matter boundary, automated topology correction, and surface deformation according to intensity gradients to optimally place the gray-white and gray-cerebrospinal fluid borders at the locations of the greatest shifts in intensity, which defines the transition to the other tissue class (Dale et al., 1999). The vertices were arranged in a triangular grid with the spacing of approximately 1 mm (∼16,000 grid points) in each hemisphere. For any inaccuracies, the qualities of the segmentation and surface reconstruction were checked carefully by four researchers using the double-blinded method (The dataset was divided into four parts. Each of the researchers checked through two parts of the datasets, and every part was checked twice by two different people.) The data of the subjects were excluded if two people marked it as poor quality. For those only marked once by the researchers, we checked the data again and decided whether it should be removed from further processing and analyses. A total of 15 subjects were excluded in the end because of the poor quality. We separated both the left and right hemispheres into gyral and sulcal regions using FreeSurfer-generated mean curvature (concave-convex division, Sulci: mean curvature value > 0/Gyri: mean curvature value < 0) for further comparison.

Cortical Thickness

By calculating the shortest distance between the pial surface and the gray matter-white matter boundary of the tessellated surface, we obtained the vertex-wise cortical thickness. To first validate the aging dataset, we tested the relationship between thickness and age across the whole-brain vertices. The surface was smoothed by a 15-mm Gaussian kernel (Fischl et al., 1999). The effects of smoothing with 10 and 20 mm kernels can be seen in **Supplementary Figures 2, 3**. To investigate the differential effects of aging between two cortical features, we further calculated the average thicknesses of the gyral and sulcal regions and the ratio of the gyral and sulcal thicknesses (Gyri/Sulci ratio = Gyral thickness divided by sulcal thickness).

Intrinsic Curvature

Intrinsic curvature is a fundamental property of a surface, and its measurements reflect higher complexity intrinsic information of surfaces and may provide a more sensitive measure of cortex than other larger-scale gyrification measures (Ronan et al., 2012). In addition to the curve values, intrinsic curvature also contains the shape information which gives rise to an un-uniform expansion of the surfaces and is demonstrated to have a greater spatial frequency when quantified at a millimeter-scale (Ronan et al., 2011). The degree of intrinsic curvature is dependent on the degree of the differential, with a bigger differential resulting in a greater degree of curvature, which might reflect the underlying connectivity of the human cortex (Ronan et al., 2014). Thus, it has been hypothesized that the higher the intrinsic curvature value, the higher the complexity, underlying connectivity, and folding or curves are in one region, although the direct evidence of the relationship between intrinsic curvature and connectivity was not presented in literature yet.

The vertex-wise intrinsic curvature was calculated using Caret software (v5.65, https://www.nitrc.org/projects/caret/) as the product of the principal curvatures (Ronan et al., 2011, 2012). The post-processing and filtration of the curvature were done in MATLAB (The MathWorks, Inc., Natick, MA, USA). We took the absolute values of intrinsic curvature for the pial and white matter surfaces. A low-pass filter (threshold = 2 mm^{-2}) was applied to minimize error, keep the curvature values compatible with the resolution of the cortical reconstruction and remove the abnormal values from further analysis (Ronan et al., 2011, 2012). Additionally, the intrinsic curvature of the pial and white matter surfaces was separately extracted for the statistical analyses. To first validate the aging dataset, a vertex-wise analysis was applied to look at the relationship between gyrification and age. The surface was smoothed by a 15-mm Gaussian kernel (Fischl et al., 1999). The effects of smoothing with 10 and 20 mm kernels can

be seen in **Supplementary Figures 2, 3**. Next, we calculated the average intrinsic curvatures for gyral and sulcal regions and the ratio of the gyri/sulci (Gyri/Sulci ratio = Gyral intrinsic curvature divided by sulcal intrinsic curvature) at the pial and white matter surfaces to investigate the differential aging effects.

Regional Analysis

We performed a regional analysis by dividing the cortical surface of each subject into five lobes: frontal, temporal, parietal, occipital, and cingulate. Insula was excluded in the analysis because it is situated deep in the lateral sulcus. The lobes of the brain were defined by the Desikan-Killiany atlas (Desikan et al., 2006). The average gyri, sulci, and gyri/sulci ratio of thickness and intrinsic curvature were calculated in each lobe.

Statistical Analyses and Curve Fitting

For each hemisphere, an age correlation analysis of cortical thickness and intrinsic curvature (on the pial and white matter surfaces separately) was first tested using the General Linear Model vertex-by-vertex with gender and total intracranial volume as covariates in order to reveal any general variation trends on the brain surface.

The linear, quadratic and cubic models were applied in the regression analyses across age to determine the age-dependencies of cortical thickness (gyral thickness, sulcal thickness, and gyri/sulci thickness ratio) and intrinsic curvature (gyral intrinsic curvature, sulcal intrinsic curvature, and gyri/sulci intrinsic curvature ratio on the pial and white matter surfaces separately). For model selection, the linear ($y = p^*age + p1$), quadratic ($y = p^*age^2 + p1^*age + p2$), and cubic ($y = p^*age^3 + p1^*age^2 + p2^*age + p3$) models were all tested using averaged whole-brain measurements, and we chose the better model according to the properties of goodness-of-fit, including the Akaike information criterion (AIC) (Akaike, 1974), root mean square error (RMSE), and R^2. To test the significance of each fitted model and minimize the type-I errors, we applied permutation-based multiple testing on all age-dependencies, reassigning age randomly 10,000 times. All p-values were adjusted by the Bonferroni correction ($p < 0.05$).

Relationships Between Cognitive Performance and Structural Measures

We examined whether the general cognitive performance (MMSE/Digit span tasks) can be explained by the structural measures of gyri and sulci on the pial and white matter surfaces, after controlling for covariates. We further hypothesized that different cortical measures may contribute different amounts of effect to cognitive performance. Therefore, hierarchical multiple regression analysis was performed in SPSS to investigate the relationship between general cognitive performance (MMSE/Digit span tasks, as dependent variable) and the structural measures of gyri and sulci on the pial and white matter surfaces (independent variables), while age, gender, TIV and education were used as covariates of non-interest in the regression model. To test the significance of the regression models, we used a p-value threshold of 0.05.

Stepwise Regression Predicting Age Using the Structural Measurements as Predictors

Stepwise regression analysis was performed to determine the relative contribution of the structural measurement to chronological aging. Sex and TIV were forced to enter in the stepwise model as covariates, and the six structural measurements including gyral/sulcal thickness and intrinsic curvature on the pial or white surface were then all entered to the stepwise model at once using a selection criterion of $p < 0.05$.

RESULTS

Aging Effect in Brain Tissue Volume

To ensure that the recruited samples are in the aging process without muddling obvious development of brain tissues, we firstly examined the relationship between brain tissue volume (GMV, WMV, CSFV, and TIV) and age across 21–92 years old. The GMV, WMV, CSFV, and TIV are plotted as a function of age in **Supplementary Figure 1**. We confirmed that GMV and WMV decrease with age, CSFV increase with age and no correlation between TIV and age was observed in the analyzed sample.

Vertex-Wise Linear Correlations Between Age and Cortical Measurements

Across the age range of 21–92 years, vertex-wise cortical thickness showed a global negative linear correlation with age in both hemispheres (**Figure 1A**). The vertex-wise age correlation results showed a regional decline (uncorrected $p < 0.01$) of the intrinsic curvature on the pial surface (**Figure 1B**). However, the intrinsic curvature had an overall increasing pattern (uncorrected $p < 0.01$) on the white matter surface (**Figure 1C**).

Goodness-of-Fit Tests of the Polynomial Regression Models

We examined the goodness-of-fit of the linear, quadratic, and cubic models, including looking at the two parameters: Root Mean Square Error (RMSE) and Akaike Information Criterion (AIC). For cortical thickness, quadratic models stood out to be the best fit (**Supplementary Table 1**). On the pial surface, bigger differences were found between the linear and quadratic models in both the RMSE and AIC results (**Supplementary Table 2**). The parameters of the quadratic and cubic models showed nearly the same, but the cubic values were smaller in most cases. The goodness-of-fit profile on the white matter surface was identical to that of the pial surface (**Supplementary Table 3**). The RMSE and AIC values were steady and smaller in the quadratic and cubic models compared with those of the linear model. Although some parameters were smaller using the cubic model compared with the quadratic model, those of cubic and quadratic models were very similar. Hence, we integrated the findings and only looked at the quadratic effects in the following analyses, which illustrated the aging process better but were not over-fit (For the linear results, see **Supplementary Figure 4** and **Supplementary Table 4**).

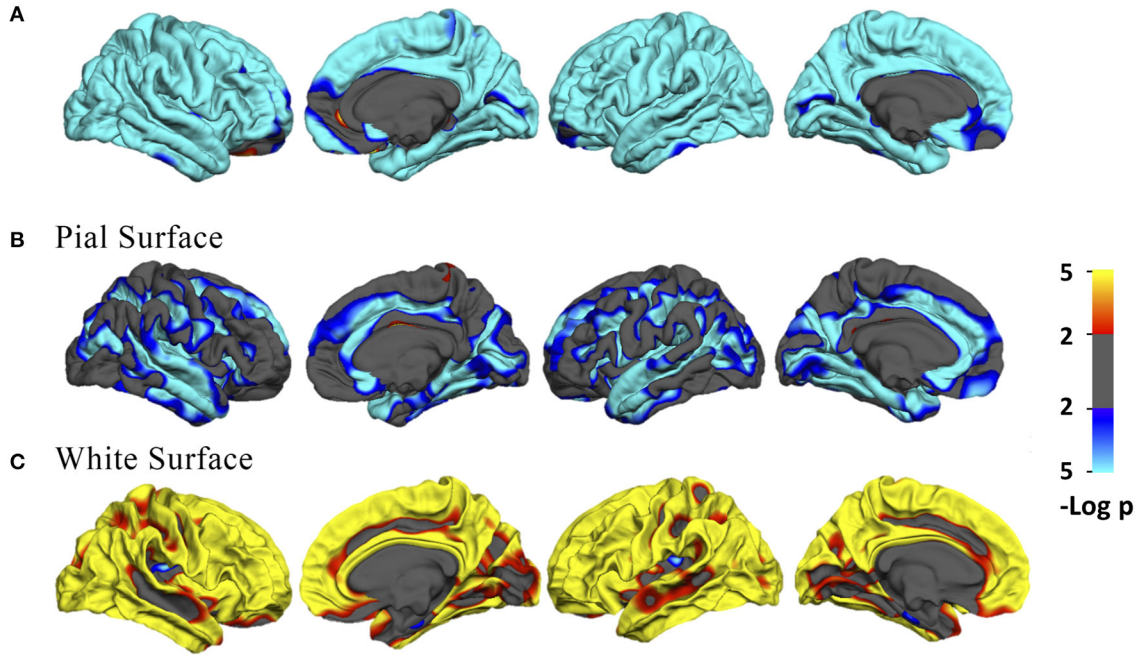

FIGURE 1 | (A) Vertex-wise age correlation in cortical thickness. **(B,C)** Vertex-wise age correlation in intrinsic curvature on the pial and white matter surfaces. The threshold was set to 2. Vertices with $p < 0.01$, uncorrected, are colored. Blue indicates a negative correlation; red indicates a positive correlation.

Quadratic Relationship Between Age and Whole-Brain Cortical Measurements

All fitted curves were tested using permutation tests, and the results were all significant ($p < 0.05$). According to the quadratic regression results (**Figure 2**, **Table 2**), both the cortical thicknesses of the gyri and sulci decreased with age [Gyral R^2 (lh/rh): 0.400/0.370; Sulcal R^2 (lh/rh): 0.562/0.526]. The gyri/sulci thickness ratio increased with age [R^2 (lh/rh): 0.287/0.231], which implied that the degree of decrease in the sulci was larger than that of the gyri.

For gyrification, the aging process affected the gyri and sulci in opposite ways on the pial surface: a negative quadratic correlation was found with age in the sulcal region [R^2 (lh/rh): 0.563/0.539], while a positive quadratic correlation was found with age in the gyral region [R^2 (lh/rh): 0.244/0.212]. Both the gyral and sulcal intrinsic curvatures of the white matter surface increased quadratically with age [Gyral R^2 (lh/rh): 0.465/0.447; Sulcal R^2 (lh/rh): 0.317/0.326]. The correlations between age and gyri/sulci intrinsic curvature ratio on the white matter surface were insignificant. The results were all similar and consistent between the right and left hemispheres. The average thickness and intrinsic curvature values are shown in **Supplementary Table 5**. A correlation matrix between all the structural measurements for gyri, sulci, and gyri/sulci ratio can be seen in **Supplementary Figure 5**.

Quadratic Relationship Between Age and Cortical Measurements of the Five Lobes

The relationship between cortical measurements and age for the five lobes including frontal, parietal, temporal, occipital lobe, and cingulate were examined. Most of the age-cortical measure relationships in the frontal, parietal, and temporal lobes were consistent with the whole-brain results. The quadratic relationships between age and the sulcal intrinsic curvature on the white matter surface of the occipital lobe, gyral intrinsic curvature on the pial surface of occipital lobe and cingulate and gyri/sulci ratio of thickness in cingulate were not significant ($p > 0.05/108$) (**Table 3**, **Supplementary Figures 6–10**).

Relationships Between Cognitive Performance and Structural Measures

The hierarchical regression models were found significant between MMSE and sulcal cortical thickness and gyral intrinsic curvature on the white matter surface. In the first step, the regression model including age, gender, education, and TIV was significant ($R^2 = 0.236$, $p < 0.001$). In the second step, adding the whole brain sulcal cortical thickness into the model explained additional 0.8% of variance, and the model stayed significant ($R^2 = 0.244$, F-change $p = 0.041$). In the third step, adding the whole brain gyral intrinsic curvature into the model also explained additional 0.8% of variance, and the model was significant ($R^2 = 0.252$, F-change $p = 0.039$). The models at the second and third steps revealed that MMSE was positively correlated with the cortical measurements (**Figure 3**; sulcal cortical thickness: $\beta = 0.141$, $p = 0.029$; gyral intrinsic curvature: $\beta = 0.109$, $p = 0.039$). All p-values reported here were uncorrected, and after FDR correction, all models were non-significant.

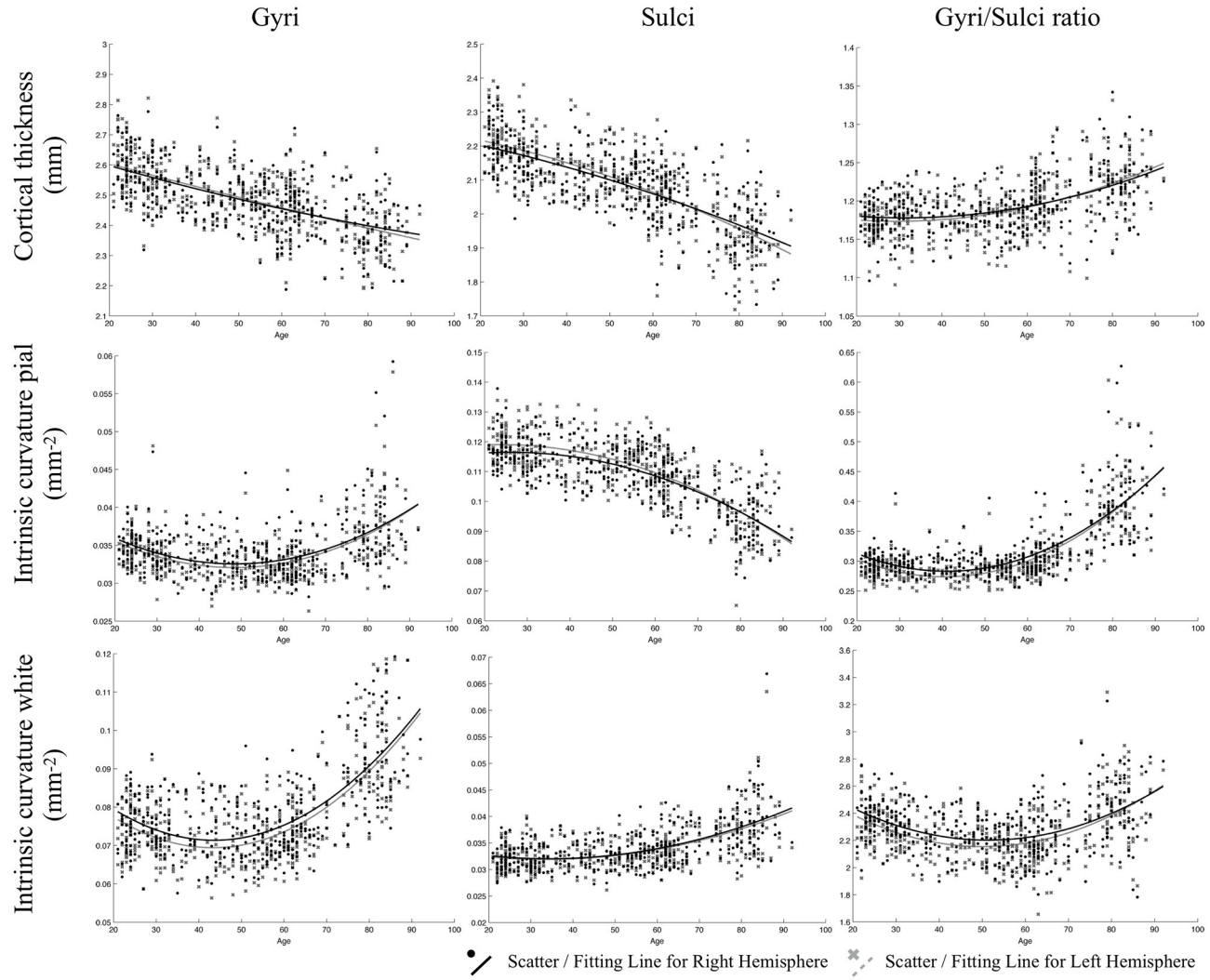

FIGURE 2 | Results of the regressions of age with gyri, sulci, and the gyri/sulci ratio of cortical thickness and intrinsic curvature on the pial and white matter surfaces. The lines refer to the fitted curve for the age and measurements, and the dots indicate the distribution of that data for the subjects.

TABLE 2 | Results of quadratic regression analyses.

Age correlation	Cortical thickness	Intrinsic curvature (lh/rh)	
	(lh/rh)	Pial surface	White surface
Gyri	–(0.400/0.370)	+(0.244/0.212)	+(0.465/0.447)
Sulci	–(0.562/0.526)	–(0.563/0.539)	+(0.317/0.326)
Gyri/Sulci ratio	+(0.287/0.231)	+(0.617/0.560)	N.A.

Only the R^2 for the significant curves were shown in the table ($p < 0.05/18$; adjusted by the Bonferroni correction for multiple comparisons). The variables are adjusted R^2; + represents positive correlation, – represents negative correlation.

Different Contribution of the Structural Measurements in Chronological Age-Dependencies

By using stepwise regression that adjusted for the sex and TIV effect, four of the six measurements were selected into the final model with adjusted, including mean thickness of sulci (standardized $\beta = -0.214$, $p < 0.001$), mean intrinsic curvature of the pial (standardized $\beta = -0.524$, $p < 0.001$), and white surface of sulci (standardized $\beta = 0.544$, $p < 0.001$), and mean intrinsic curvature of pial surface of gyri (standardized $\beta = -0.284$, $p < 0.001$). All p-values were significant after FDR correction.

The Aging Model

To summarize the various measurements and structural findings in the current and previous studies, we suggest a putative model of the general pattern of cortical aging (**Figure 4**). In this model, the gyral intrinsic curvature of the pial surface increases slightly with age, while the tips of the gyri stay close to the skull [**Figure 4(1)**]. The gyral intrinsic curvature on the white matter surface increases with age, so we hypothesized that the surface may move outward, resulting in a decrease in cortical thickness [**Figure 4(2)**]. The sulcal intrinsic curvature of the pial surface

TABLE 3 | Results of quadratic regression analyses of the 5 lobes.

Lobe	Age correlation	Cortical thickness	Intrinsic curvature (lh/rh)	
		(lh/rh)	Pial surface	White surface
Frontal	Gyri	–(0.414/0.259)	+(0.245/0.271)	+(0.448/0.457)
	Sulci	–(0.284/0.271)	–(0.366/0.315)	+(0.387/0.391)
	Gyri/Sulci ratio	+(0.078/ N.A.)	+(0.494/0.485)	N.A.
Parietal	Gyri	–(0.414/0.259)	+(0.192/0.185)	+(0.461/0.446)
	Sulci	–(0.483/0.478)	–(0.443/0.436)	+(0.223/0.201)
	Gyri/Sulci ratio	+(0.261/0.248)	+(0.561/0.545)	+(0.243/0.187)
Temporal	Gyri	–(0.350/0.353)	+(0.228/0.197)	+(0.408/0.388)
	Sulci	–(0.623/0.594)	–(0.609/0.561)	+(0.220/0.198)
	Gyri/Sulci ratio	+(0.493/0.426)	+(0.585/0.498)	+(0.202/N.A.)
Occipital	Gyri	–(0.117/0.142)	N.A.	+(0.324/0.251)
	Sulci	–(0.561/0.583)	–(0.386/0.358)	N.A.
	Gyri/Sulci ratio	+(0.513/0.490)	+(0.418/0.314)	+(0.299/0.263)
Cingulate	Gyri	–(0.414/0.259)	N.A.	+(0.358/0.220)
	Sulci	–(0.284/0.271)	–(0.378/0.368)	+(0.143/0.295)
	Gyri/Sulci ratio	N.A.	+(0.413/0.218)	+(N.A./0.121)

The significance of the curve fitting is evaluated using the permutation test, only the R^2 for the significant curves were shown in the table ($p < 0.05/108$; adjusted by the Bonferroni correction for multiple comparisons). The variables are adjusted R^2; + represents positive correlation, – represents negative correlation.

decreases with age [**Figure 4(3)**], which indicates that the sulci are getting wider and flatter on the surface with increasing age. The pial border might also move outward, resulting in a decline in the sulcal depth and increase in sulcal width during aging which was reported in previous literature (Kochunov et al., 2005; Liu et al., 2013). On the white matter surface, the sulcal intrinsic curvature increases with age, and the gray-white matter boundary may move outward and becomes cusped [**Figure 4(4)**]. Finally, the decreasing extent of the sulcal thickness appears larger than that of the gyral thickness. Additionally, we found that the trends of all correlations were consistent for both gender (see **Supplementary Figure 11**), indicating that the overall pattern of the current putative model reflects common mechanisms of change during aging.

DISCUSSION

This study aims to reveal the pattern of the gyral and sulcal changes on the pial and white matter surfaces during the normal aging process using a relatively large cohort dataset. We investigated the whole-brain vertex-wise pattern of the structural changes during the aging process, in which the white matter and pial surfaces showed a different association of gyrification with age. The gyri/sulci ratio was used to highlight the difference of cortical thinning and curvature changes between gyri and sulci spanning across aging. Current findings suggest that, instead of gyri, the changes of sulcal thickness and curvature contribute more during the normal aging process. Finally, we also concluded a putative model of aging in this study based on previous evidence for the fundamental shape of the cortex that accompanied by the current results, which provides a better understanding of the cortical structure degeneration across adulthood to old age.

As the advance of age, we observed that the curvature decreased on the sulcal-pial surface and increased on the sulcal-white matter surface (**Figure 2**). Previous studies have found that sulcal width increases while sulcal depth decreases with age (Kochunov et al., 2005; Liu et al., 2013). This finding partially supports the fact that the sulcal-pial surface might flatten and move outward instead of shrinking toward the white matter surface, which is mainly caused by the steady production of CSF from the choroid plexus that inflates the cerebrum during the loss of brain parenchyma (Miller et al., 1987; Matsumae et al., 1996; Scahill et al., 2003). Moreover, the increased curvature of the WM surface is associated with imbalanced WM-GM shrinkage (Deppe et al., 2014). The specific changes of sulcal morphology in WM-GM surface can be linked with the loss of short association fibers (U-fiber) underneath sulcus that contribute to the impaired local clustering of the brain connections (Toro and Burnod, 2005; Gao et al., 2014; Van Essen et al., 2018). Third, by analyzing the gyri/sulci ratio, we found that the degree of sulcal thickness thinning was larger than that of gyral thickness during the normal aging process. These findings implied that the changes of sulcal morphology are more prominent than gyral regions during aging, which resulted in the variation in the following ways: (1) For sulcal morphology, while the sulcal pial surface moved outward and flatten, the GM-WM surface became cusped and the thickness of sulci decreases. (2) For gyral regions, the curvature on both surfaces became cusped, with mildly decreased gyral thickness. All inferences and current evidence were integrated and graphed in the putative model of aging (**Figure 4**).

One of the noticeable findings in the current study is that gyri and sulci altered differently during the normal aging process.

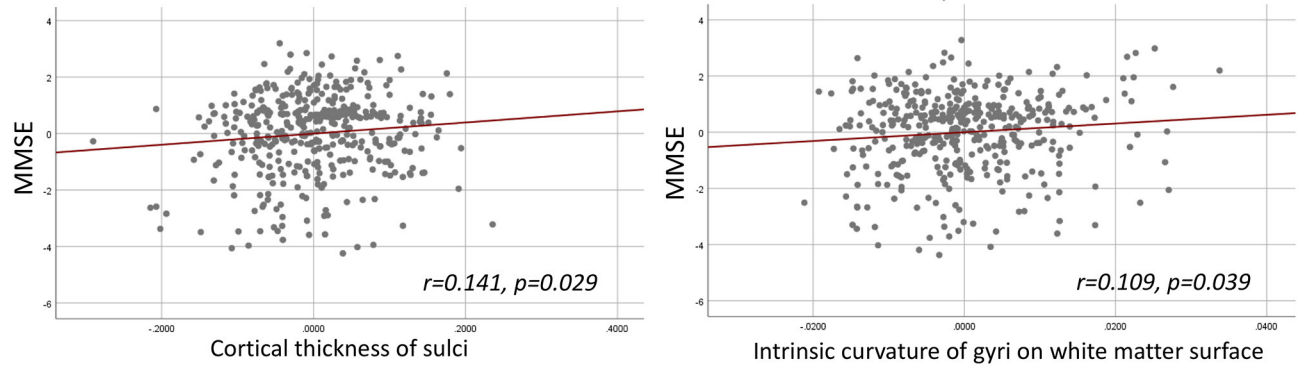

FIGURE 3 | Hierarchical multiple regression result. **Left**: Partial regression plot of MMSE and sulcal cortical thickness. **Right**: Partial regression plot of MMSE and gyral intrinsic curvature of the white matter surface.

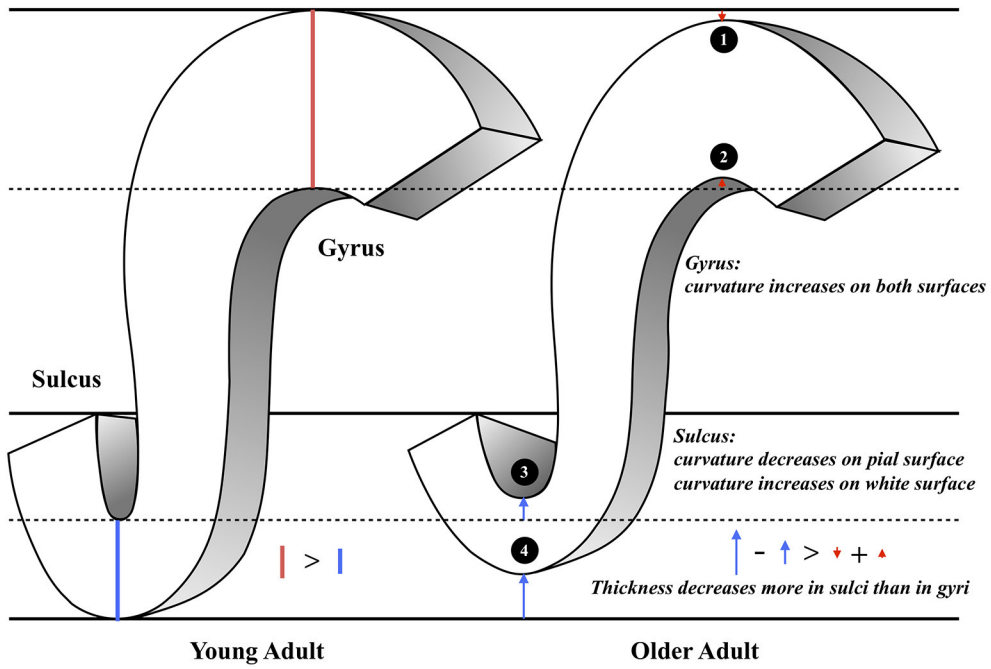

FIGURE 4 | The illustration of the putative morphological model of aging. This figure concludes our main findings as the following: **(1)** Gyral intrinsic curvature on the pial surface slightly increases with age. **(2)** Gyral intrinsic curvature on the white matter surface increases with age. **(3)** Sulcal intrinsic curvature on the pial surface decreases with age; sulcal width increases and sulcal depth decreases with age **(4)** Sulcal intrinsic curvature on the white matter surface increases with age. **(1–4)** Sulcal thickness declines more than gyral thickness (see the length of the red and blue arrows).

Gyral crowns were reported having specialized and enhanced connections and organization between cortices (Brodmann, 1909; Welker, 1990). Although common functional regions were mostly defined in the gyral region and its adjacent sulci, our findings suggested that the sulci itself greatly altered and may be responsible for the decrease of functional segregation during aging. Previous studies using diffusion-weighted imaging have indicated that gyral regions show denser white matter fibers than sulcal regions (Nie et al., 2012; Chen et al., 2013). Deng et al. (2014) further supported these findings that gyri are functional connection centers, while sulci are likely to be local functional units connecting neighboring gyri through inter-column cortico-cortical fibers. Moreover, Gao et al. (2014) found association between the disintegrity in short-range fibers and lower cognitive efficiency on prospective memory, and the loss of the clustering coefficient has also been found to correlate with inferior intelligence quotient (IQ) (Li et al., 2009). Taking together, we suggest that morphological degeneration in sulcal regions could be more vulnerable to the effects of aging. The loss of local interconnectivity in the brain might be more pronounced in the normal aging process. In the current study, we found mild trends that MMSE was positively correlated with sulcal cortical thickness and gyral intrinsic curvature on the white matter surface. Therefore, we assumed that the sulcal thickness

declines suggest that part of the short-ranged connectivity changes and may influence intra-cortical brain function (Schuz and Palm, 1989; Elston, 2003; Cullen et al., 2010; Wagstyl et al., 2015). Moreover, targeting the cognitive correlates of sulcal degeneration and its impact to brain topology should be investigated in future studies.

We also found different trends of curvature changing on the WM-GM boundary and the pial surface during aging. Based on the retrogenesis mechanism, the differential proliferation hypothesis (Richman et al., 1975) may support differential degeneration process of gyrification of the pial and white matter surfaces during aging. Changes in myelin and synapses in the cortex could be the reason for the morphology development and degeneration, and thereby resulting in cognitive decline (Bartzokis, 2004; Masliah et al., 2006; Fjell and Walhovd, 2010; Whitaker et al., 2016). Nonetheless, a direct link between cortical myelination and gyrification still needs to be found. Although several hypotheses of the mechanisms underlying gyrification are currently being debated by neuroscientists, some have suggested that gyrification is shaped or influenced by multiple mechanisms (Ronan and Fletcher, 2015). The alterations in aging-related gyrification could reflect the underlying cortical connections and the functions of our brain by either pruning or degenerative processes (Ronan et al., 2011; Jockwitz et al., 2017). The vulnerability of sulcal cortical thinning that was discovered in the current study could shed light on the structure-function relationship in the human brain.

The aging model of cerebral cortex we proposed in this study is a global effect spanning across most of the brain regions. However, while most of the regions showing greater gyri/sulci ratio, frontal lobe and cingulate showed similar degree of a decrease between gyral and sulcal thickness, where the amount of decrease was non-significant correlating with age (**Table 3**). Our findings implied that the gyral and sulcal thickness in the frontal and cingulate regions decreased to the same extent during the normal aging process. Several studies have found accelerated regional gray matter volume decline in the frontal lobe compared with other lobes (Tisserand et al., 2002; Resnick et al., 2003) and have demonstrated decreases in cortical volume specifically in the frontal and cingulate regions. These findings consistently suggest that the frontal and cingulate cortices may be the key regions in the brain aging process. Another study reported a higher increase of sulcal width in the superior frontal sulci and a lower correlation between age and decreased sulcal depth in the inferior and orbitofrontal sulci (Kochunov et al., 2005). This report might support our findings that both gyral and sulcal thickness decreased to the same degree so that sulcal width increased significantly and sulcal depth decreased because of the high atrophy in gyri. Moreover, the trend for sulcal gyrification changes on the white matter surface in the occipital lobe did not increase as we found in the whole-brain examination. Small variations exist in different brain regions because the development of the lobes and their functions are diverse.

In this study, we characterized the effects of age on different structural progressions in a large sample of healthy adults. However, this study had several limitations. First, the causality of cortical thickness and intrinsic curvature affecting brain function was not investigated. Due to the nature of the cross-sectional design, we were unable to avoid cohort effect or indicate which structure degenerated earlier or had a higher impact on the brain function. The structure-function relationships among gyri, sulci, and the pial and white matter surfaces and how they impact each other still need to be examined. In this case, the envisioned model of the degeneration of the cerebral cortex may need to be investigated with longitudinal data. Second, to generalize the degenerative process, the current model focused on the trends for comprehensive changes in the cortex and brain lobes. However, cortical variations during aging, including changes in volume and thickness, have been found to decline regionally (Thambisetty et al., 2010; Westlye et al., 2010). Therefore, although we posited general aging-related alterations in the gyri and sulci of the cortex in the model, regional variations still need to be specified. Third, we used a simple concave-convex concept as a division of gyri and sulci which may seem to be an arbitrary classification. Adopting the convexity map generated by FreeSurfer to classify cortex into gyral, sulcal and undefined regions (Liu et al., 2019; Yang et al., 2019; Zhang et al., 2019) is recommended for future analysis especially when looking at regional changes in the analysis. Next, our aging model is partly based on previous literature. The measure of sulcal width and depth was not directly quantified in the current study. The completeness of the aging model could be tested and reproduced with all the measurements together. Lastly, participants with mild or severe cognitive functions were excluded in this study, and only general cognition assessments were conducted for current participants. Thus, the generalizability of current findings is limited to healthy population, moreover, its cognitive implications should be further examined with detailed cognitions such as verbal memory, visual executive in future studies.

CONCLUSION

This study illustrates a cortical degeneration model from the perspective of brain morphology which provides an overview for the brain aging process using multiple structural measurements. We found systematical and nonuniform cortical thinning during normal aging, that the overall degree of sulcal degeneration is greater than gyri in terms of thickness and gyrification. These degeneration mechanisms might relate to pruning, life-long reshaping and neurodegenerative processes, associating with differential brain functional degeneration and the underlying neuronal tension. We suggest that the cortical features of gyri, sulci, the pial and white matter surfaces should be considered independently in future studies, which could be associated with segregation and integration alterations in brain connectome during aging.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board of Taipei Veterans General Hospital. All participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

H-YL, C-CH, and C-PL conceived and planned the experiments. AY and S-JT contributed to data collection. H-YL and C-CH conducted the experiments and data analysis. H-YL, C-CH, K-HC, AY, C-YL, S-JT, and C-PL contributed to the interpretation of the results. H-YL and C-CH took the lead in writing the manuscript. All authors provided feedback and helped with the analysis and manuscript.

FUNDING

C-PL was supported by the Ministry of Science and Technology (MOST) of Taiwan (Grant Nos. MOST 108-2420-H-010-001; MOST 108-2321-B-010-010-MY2; MOST 109-2634-F-010-001; and MOST 108-2321-B-010-013-MY2); Taipei Veterans General Hospital (VGHUST109-V1-2-2); Shanghai Science and Technology Innovation Plan Grant Nos. 17JC1404105 and 17JC1404101. K-HC was supported by grant from the Ministry of Science and Technology of Taiwan (MOST 107-2221-E-010-010-MY3). AY and S-JT were supported by grants (109-2628-B-010-011; 109-2321-B-010-006; and 109-2634-F-075-001) from the Ministry of Science and Technology of Taiwan. C-YL was partially supported by Shanghai Municipal Science and Technology Major Project (No. 2018SHZDZX01), the National Key Research and Development Program of China (No. 2018YFC0910503), the Young Scientists Fund of the National Natural Science Foundation of China (No. 81801774), the Natural Science Foundation of Shanghai (No. 18ZR1403700), ZJ Lab, and Shanghai Center for Brain Science and Brain-Inspired Technology. This work was supported by the Brain Research Center, National Yang-Ming University from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE), Taipei, Taiwan.

ACKNOWLEDGMENTS

We acknowledge Yi-Hua Huang, I-Ting Lee, Yi-Hui Lin for helping with the surface reconstruction quality assurance. We thank Daniel Cheong and Maddie Kiszewski for proofreading the manuscript. We thank Chih-Chin Heather Hsu and I-Ting Lee for data access technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2021.625931/full#supplementary-material

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Automatic Control* 19, 716–723. doi: 10.1109/TAC.1974.1100705

American Psychiatric, A. (2000). *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR*. Washington, DC: American Psychiatric Association.

Bartzokis, G. (2004). Age-related myelin breakdown: a developmental model of cognitive decline and Alzheimer's disease. *Neurobiol Aging.* 25, 5–18; author reply 49-62. doi: 10.1016/j.neurobiolaging.2003.03.001

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Großhirnrinde: in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Leizig: Barth.

Cao, B., Mwangi, B., Passos, I. C., Wu, M. J., Keser, Z., Zunta-Soares, G. B., et al. (2017). Lifespan gyrification trajectories of human brain in healthy individuals and patients with major psychiatric disorders. *Sci Rep.* 7:511. doi: 10.1038/s41598-017-00582-1

Chan, M. Y., Park, D. C., Savalia, N. K., Petersen, S. E., and Wig, G. S. (2014). Decreased segregation of brain systems across the healthy adult lifespan. *Proc. Natl. Acad. Sci. U.S.A.* 111, E4997–E5006. doi: 10.1073/pnas.1415122111

Chen, H., Zhang, T., Guo, L., Li, K., Yu, X., Li, L., et al. (2013). Coevolution of gyral folding and structural connection patterns in primate brains. *Cereb Cortex.* 23, 1208–1217. doi: 10.1093/cercor/bhs113

Cullen, D. K., Gilroy, M. E., Irons, H. R., and Laplaca, M. C. (2010). Synapse-to-neuron ratio is inversely related to neuronal density in mature neuronal cultures. *Brain Res.* 1359, 44–55. doi: 10.1016/j.brainres.2010.08.058

Dale, A. M., Fischl, B., and Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage.* 9, 179–194. doi: 10.1006/nimg.1998.0395

Deng, F., Jiang, X., Zhu, D., Zhang, T., Li, K., Guo, L., et al. (2014). A functional model of cortical gyri and sulci. *Brain Struct Funct.* 219, 1473–1491. doi: 10.1007/s00429-013-0581-z

Deppe, M., Marinell, J., Krämer, J., Duning, T., Ruck, T., Simon, O. J., et al. (2014). Increased cortical curvature reflects white matter atrophy in individual patients with early multiple sclerosis. *NeuroImage* 6, 475–487. doi: 10.1016/j.nicl.2014.02.012

Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage.* 31, 968–980. doi: 10.1016/j.neuroimage.2006.01.021

Elston, G. N. (2003). Cortex, cognition and the cell: new insights into the pyramidal neuron and prefrontal function. *Cereb Cortex.* 13, 1124–1138. doi: 10.1093/cercor/bhg093

Fischl, B., and Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc. Natl. Acad. Sci. U.S.A.* 97, 11050–11055. doi: 10.1073/pnas.200033797

Fischl, B., Sereno, M. I., Tootell, R. B., and Dale, A. M. (1999). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum. Brain Mapp.* 8, 272–284. doi: 10.1002/(SICI)1097-0193(1999)8:4<272::AID-HBM10>3.0.CO;2-4

Fjell, A. M., Grydeland, H., Krogsrud, S. K., Amlien, I., Rohani, D. A., Ferschmann, L., et al. (2015). Development and aging of cortical thickness correspond to genetic organization patterns. *Proc. Natl. Acad. Sci. U.S.A.* 112, 15462–15467. doi: 10.1073/pnas.1508831112

Fjell, A. M., and Walhovd, K. B. (2010). Structural brain changes in aging: courses, causes and cognitive consequences. *Rev. Neurosci.* 21, 187–221. doi: 10.1515/REVNEURO.2010.21.3.187

Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198. doi: 10.1016/0022-3956(75)90026-6

Gao, J., Cheung, R. T. F., Chan, Y. S., Chu, L. W., Mak, H. K. F., and Lee, T. M. C. (2014). The relevance of short-range fibers to cognitive

efficiency and brain activation in aging and dementia. *PLoS ONE*. 9:e90307. doi: 10.1371/journal.pone.0090307

Gautam, P., Anstey, K. J., Wen, W., Sachdev, P. S., and Cherbuin, N. (2015). Cortical gyrification and its relationships with cortical volume, cortical thickness, and cognitive performance in healthy mid-life adults. *Behav. Brain Res*. 287, 331–339. doi: 10.1016/j.bbr.2015.03.018

Good, C. D., Johnsrude, I. S., Ashburner, J., Henson, R. N., Friston, K. J., and Frackowiak, R. S. (2001). A voxel-based morphometric study of ageing in 465 normal adult human brains. *Neuroimage*. 14, 21–36. doi: 10.1006/nimg.2001.0786

Gregory, M. D., Kippenhan, J. S., Dickinson, D., Carrasco, J., Mattay, V. S., Weinberger, D. R., et al. (2016). Regional variations in brain gyrification are associated with general cognitive ability in humans. *Curr. Biol*. 26, 1301–1305. doi: 10.1016/j.cub.2016.03.021

Griffin, L. D. (1994). The intrinsic geometry of the cerebral cortex. *J. Theor. Biol*. 166, 261–273. doi: 10.1006/jtbi.1994.1024

Hughes, C. P., Berg, L., Danziger, W. L., Coben, L. A., and Martin, R. L. (1982). A new clinical scale for the staging of dementia. *Br. J. Psychiatry*. 140, 566–572. doi: 10.1192/bjp.140.6.566

Im, K., Lee, J. M., Seo, S. W., Hyung Kim, S., Kim, S. I., and Na, D. L. (2008). Sulcal morphology changes and their relationship with cortical thickness and gyral white matter volume in mild cognitive impairment and Alzheimer's disease. *Neuroimage*. 43, 103–113. doi: 10.1016/j.neuroimage.2008.07.016

Jernigan, T. L., Archibald, S. L., Fennema-Notestine, C., Gamst, A. C., Stout, J. C., Bonner, J., et al. (2001). Effects of age on tissues and regions of the cerebrum and cerebellum. *Neurobiol. Aging*. 22, 581–594. doi: 10.1016/S0197-4580(01)00217-2

Jockwitz, C., Caspers, S., Lux, S., Jutten, K., Schleicher, A., Eickhoff, S. B., et al. (2017). Age- and function-related regional changes in cortical folding of the default mode network in older adults. *Brain Struct. Funct*. 222, 83–99. doi: 10.1007/s00429-016-1202-4

Jones, B. F., Barnes, J., Uylings, H. B., Fox, N. C., Frost, C., Witter, M. P., et al. (2006). Differential regional atrophy of the cingulate gyrus in Alzheimer disease: a volumetric MRI study. *Cereb Cortex*. 16, 1701–1708. doi: 10.1093/cercor/bhj105

Kaas, J. H. (2013). The evolution of brains from early mammals to humans. *Wiley Interdiscip Rev. Cogn. Sci*. 4, 33–45. doi: 10.1002/wcs.1206

Klein, D., Rotarska-Jagiela, A., Genc, E., Sritharan, S., Mohr, H., Roux, F., et al. (2014). Adolescent brain maturation and cortical folding: evidence for reductions in gyrification. *PLoS ONE* 9:e84914. doi: 10.1371/journal.pone.0084914

Kochunov, P., Mangin, J. F., Coyle, T., Lancaster, J., Thompson, P., Riviere, D., et al. (2005). Age-related morphology trends of cortical sulci. *Hum Brain Mapp*. 26, 210–220. doi: 10.1002/hbm.20198

Li, Y., Liu, Y., Li, J., Qin, W., Li, K., Yu, C., et al. (2009). Brain Anatomical Network and Intelligence. *PLOS Comp. Biol*. 5:e1000395. doi: 10.1371/journal.pcbi.1000395

Liu, H., Jiang, X., Zhang, T., Ren, Y., Hu, X., Guo, L., et al. (2017). Elucidating functional differences between cortical gyri and sulci via sparse representation HCP grayordinate fMRI data. *Brain Res*. 1672, 81–90. doi: 10.1016/j.brainres.2017.07.018

Liu, H., Zhang, S., Jiang, X., Zhang, T., Huang, H., Ge, F., et al. (2019). The cerebral cortex is bisectionally segregated into two fundamentally different functional units of gyri and sulci. *Cereb. Cortex* 29, 4238–4252. doi: 10.1093/cercor/bhy305

Liu, T., Sachdev, P. S., Lipnicki, D. M., Jiang, J., Cui, Y., Kochan, N. A., et al. (2013). Longitudinal changes in sulcal morphology associated with late-life aging and MCI. *Neuroimage*. 74, 337–342. doi: 10.1016/j.neuroimage.2013.02.047

Llinares-Benadero, C., and Borrell, V. (2019). Deconstructing cortical folding: genetic, cellular and mechanical determinants. *Nat. Rev. Neurosci*. 20, 161–176. doi: 10.1038/s41583-018-0112-2

Magnotta, V. A., Andreasen, N. C., Schultz, S. K., Harris, G., Cizadlo, T., Heckel, D., et al. (1999). Quantitative in vivo measurement of gyrification in the human brain: changes associated with aging. *Cereb Cortex*. 9, 151–160. doi: 10.1093/cercor/9.2.151

Masliah, E., Crews, L., and Hansen, L. (2006). Synaptic remodeling during aging and in Alzheimer's disease. *J. Alzheimers Dis*. 9, 91–99. doi: 10.3233/JAD-2006-9S311

Matsumae, M., Kikinis, R., Morocz, I. A., Lorenzo, A. V., Sandor, T., Albert, M. S., et al. (1996). Age-related changes in intracranial compartment volumes in normal adults assessed by magnetic resonance imaging. *J. Neurosurg*. 84, 982–991. doi: 10.3171/jns.1996.84.6.0982

Mega, M. S., Thompson, P. M., Cummings, J. L., Back, C. L., Xu, M. L., Zohoori, S., et al. (1998). Sulcal variability in the Alzheimer's brain: correlations with cognition. *Neurology*. 50, 145–151. doi: 10.1212/WNL.50.1.145

Miller, J. D., Peeler, D. F., Pattisapu, J., and Parent, A. D. (1987). Supratentorial pressures. Part I: Differential intracranial pressures. *Neurol. Res*. 9, 193–197. doi: 10.1080/01616412.1987.11739794

Nadarajah, B., and Parnavelas, J. G. (2002). Modes of neuronal migration in the developing cerebral cortex. *Nat. Rev. Neurosci*. 3, 423–432. doi: 10.1038/nrn845

Nie, J., Guo, L., Li, K., Wang, Y., Chen, G., Li, L., et al. (2012). Axonal fiber terminations concentrate on gyri. *Cereb Cortex*. 22, 2831–2839. doi: 10.1093/cercor/bhr361

Panizzon, M. S., Fennema-Notestine, C., Eyler, L. T., Jernigan, T. L., Prom-Wormley, E., Neale, M., et al. (2009). Distinct genetic influences on cortical surface area and cortical thickness. *Cereb Cortex*. 19, 2728–2735. doi: 10.1093/cercor/bhp026

Peng, Q., Schork, A., Bartsch, H., Lo, M. T., Panizzon, M. S., Pediatric Imaging, N., et al. (2016). Conservation of distinct genetically-mediated human cortical pattern. *PLoS Genet*. 12:e1006143. doi: 10.1371/journal.pgen.1006143

Raz, N., Gunning-Dixon, F., Head, D., Rodrigue, K. M., Williamson, A., and Acker, J. D. (2004). Aging, sexual dimorphism, and hemispheric asymmetry of the cerebral cortex: replicability of regional differences in volume. *Neurobiol. Aging*. 25, 377–396. doi: 10.1016/S0197-4580(03)00118-0

Resnick, S. M., Pham, D. L., Kraut, M. A., Zonderman, A. B., and Davatzikos, C. (2003). Longitudinal magnetic resonance imaging studies of older adults: a shrinking brain. *J. Neurosci*. 23, 3295–3301. doi: 10.1523/JNEUROSCI.23-08-03295.2003

Richman, D. P., Stewart, R. M., Hutchinson, J. W., and Caviness, V. S. Jr. (1975). Mechanical model of brain convolutional development. *Science*. 189, 18–21. doi: 10.1126/science.1135626

Ronan, L., and Fletcher, P. C. (2015). From genes to folds: a review of cortical gyrification theory. *Brain Struct. Funct*. 220, 2475–2483. doi: 10.1007/s00429-014-0961-z

Ronan, L., Pienaar, R., Williams, G., Bullmore, E., Crow, T. J., Roberts, N., et al. (2011). Intrinsic curvature: a marker of millimeter-scale tangential cortico-cortical connectivity? *Int. J. Neural. Syst*. 21, 351–366. doi: 10.1142/S0129065711002948

Ronan, L., Voets, N., Rua, C., Alexander-Bloch, A., Hough, M., Mackay, C., et al. (2014). Differential tangential expansion as a mechanism for cortical gyrification. *Cereb. Cortex*. 24, 2219–2228. doi: 10.1093/cercor/bht082

Ronan, L., Voets, N. L., Hough, M., Mackay, C., Roberts, N., Suckling, J., et al. (2012). Consistency and interpretation of changes in millimeter-scale cortical intrinsic curvature across three independent datasets in schizophrenia. *Neuroimage*. 63, 611–621. doi: 10.1016/j.neuroimage.2012.06.034

Salat, D. H., Buckner, R. L., Snyder, A. Z., Greve, D. N., Desikan, R. S., Busa, E., et al. (2004). Thinning of the cerebral cortex in aging. *Cereb. Cortex*. 14, 721–730. doi: 10.1093/cercor/bhh032

Scahill, R. I., Frost, C., Jenkins, R., Whitwell, J. L., Rossor, M. N., and Fox, N. C. (2003). A longitudinal study of brain volume changes in normal aging using serial registered magnetic resonance imaging. *Arch. Neurol*. 60, 989–994. doi: 10.1001/archneur.60.7.989

Schaer, M., Cuadra, M. B., Tamarit, L., Lazeyras, F., Eliez, S., and Thiran, J. P. (2008). A surface-based approach to quantify local cortical gyrification. *IEEE Trans. Med. Imag*. 27, 161–170. doi: 10.1109/TMI.2007.903576

Schuz, A., and Palm, G. (1989). Density of neurons and synapses in the cerebral cortex of the mouse. *J. Comp. Neurol*. 286, 442–455. doi: 10.1002/cne.902860404

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., et al. (1998). The mini-international neuropsychiatric interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry*. 59(Suppl. 20), 22-33;quiz 34-57.

Storsve, A. B., Fjell, A. M., Tamnes, C. K., Westlye, L. T., Overbye, K., Aasland, H. W., et al. (2014). Differential longitudinal changes in cortical thickness, surface area and volume across the adult life span:

regions of accelerating and decelerating change. *J. Neurosci*. 34, 8488–8498. doi: 10.1523/JNEUROSCI.0391-14.2014

Sullivan, E. V., Lim, K. O., Mathalon, D., Marsh, L., Beal, D. M., Harris, D., et al. (1998). A profile of cortical gray matter volume deficits characteristic of schizophrenia. *Cereb. Cortex*. 8, 117–124. doi: 10.1093/cercor/8.2.117

Thambisetty, M., Wan, J., Carass, A., An, Y., Prince, J. L., and Resnick, S. M. (2010). Longitudinal changes in cortical thickness associated with normal aging. *NeuroImage*. 52, 1215–1223. doi: 10.1016/j.neuroimage.2010.04.258

Tisserand, D. J., Pruessner, J. C., Sanz Arigita, E. J., van Boxtel, M. P. J., Evans, A. C., Jolles, J., et al. (2002). Regional frontal cortical volumes decrease differentially in aging: an MRI study to compare volumetric approaches and voxel-based morphometry. *NeuroImage*. 17, 657–669. doi: 10.1006/nimg.2002.1173

Toro, R., and Burnod, Y. (2005). A morphogenetic model for the development of cortical convolutions. *Cereb Cortex*. 15, 1900–1913. doi: 10.1093/cercor/bhi068

Turner, G. R., and Spreng, R. N. (2012). Executive functions and neurocognitive aging: dissociable patterns of brain activity. *Neurobiol Aging*. 33, 826.e1–13. doi: 10.1016/j.neurobiolaging.2011.06.005

Van Essen, D. C., Donahue, C. J., and Glasser, M. F. (2018). Development and evolution of cerebral and cerebellar cortex. *Brain Behav Evol*. 91, 158–169. doi: 10.1159/000489943

Vandekar, S. N., Shinohara, R. T., Raznahan, A., Roalf, D. R., Ross, M., DeLeo, N., et al. (2015). Topologically dissociable patterns of development of the human cerebral cortex. *J Neurosci*. 35, 599–609. doi: 10.1523/JNEUROSCI.3628-14.2015

Wagstyl, K., Ronan, L., Goodyer, I. M., and Fletcher, P. C. (2015). Cortical thickness gradients in structural hierarchies. *Neuroimage*. 111, 241–250. doi: 10.1016/j.neuroimage.2015.02.036

Wagstyl, K., Ronan, L., Whitaker, K. J., Goodyer, I. M., Roberts, N., Crow, T. J., et al. (2016). Multiple markers of cortical morphology reveal evidence of supragranular thinning in schizophrenia. *Transl. Psychiatry*. 6:e780. doi: 10.1038/tp.2016.43

Welker, W. (1990). Why does cerebral cortex fissure and fold? a review of determinants of gyri and sulci. *Cereb. Cortex*. 8B, 3–136. doi: 10.1007/978-1-4615-3824-0_1

Westlye, L. T., Walhovd, K. B., Dale, A. M., Bjornerud, A., Due-Tonnessen, P., Engvig, A., et al. (2010). Differentiating maturational and aging-related changes of the cerebral cortex by use of thickness and signal intensity. *Neuroimage*. 52, 172–185. doi: 10.1016/j.neuroimage.2010.03.056

Whitaker, K. J., Vertes, P. E., Romero-Garcia, R., Vasa, F., Moutoussis, M., Prabhu, G., et al. (2016). Adolescence is associated with genomically patterned consolidation of the hubs of the human brain connectome. *Proc Natl Acad Sci U.S.A*. 113, 9105–9110. doi: 10.1073/pnas.1601745113

Winkler, A. M., Kochunov, P., Blangero, J., Almasy, L., Zilles, K., Fox, P. T., et al. (2010). Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. *Neuroimage*. 53, 1135–1146. doi: 10.1016/j.neuroimage.2009.12.028

Xu, G., Knutsen, A. K., Dikranian, K., Kroenke, C. D., Bayly, P. V., and Taber, L. A. (2010). Axons pull on the brain, but tension does not drive cortical folding. *J. Biomech. Eng*. 132:071013. doi: 10.1115/1.4001683

Yang, S., Zhao, Z., Cui, H., Zhang, T., Zhao, L., He, Z., et al. (2019). Temporal variability of cortical gyral-sulcal resting state functional activity correlates with fluid intelligence. *Front. Neural Circuits*. 13:36. doi: 10.3389/fncir.2019.00036

Zhang, S., Liu, H., Huang, H., Zhao, Y., Jiang, X., Bowers, B., et al. (2019). Deep learning models unveiled functional difference between cortical gyri and sulci. *IEEE Trans. Biomed. Engin*. 66, 1297–1308. doi: 10.1109/TBME.2018.2872726

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurology

SYSTEMATIC REVIEW
published: 09 April 2021
doi: 10.3389/fneur.2021.627824

Premature Mortality, Risk Factors, and Causes of Death Following Childhood-Onset Neurological Impairments: A Systematic Review

Jonathan A. Abuga[1,2*], *Symon M. Kariuki*[1,3,4], *Samson M. Kinyanjui*[1,3,5], *Michael Boele van Hensbroek*[2] *and Charles R. Newton*[1,3,4]

[1] *Kenya Medical Research Institute (KEMRI-Wellcome Trust Research Programme), Clinical Research (Neurosciences), Kilifi, Kenya,* [2] *Global Child Health Group, Emma Children's Hospital, Academic Medical Centre, University of Amsterdam, Amsterdam, Netherlands,* [3] *Department of Public Health, Pwani University, Kilifi, Kenya,* [4] *Department of Psychiatry, University of Oxford, Oxford, United Kingdom,* [5] *Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom*

OPEN ACCESS

Edited by:
Tomáš Paus, University of Toronto, Canada

Reviewed by:
Paola Costa-Mallen, Bastyr University, United States
Maura Pugliatti, University of Ferrara, Italy

***Correspondence:**
Jonathan A. Abuga
jabuga@kemri-wellcome.org; abugajn@gmail.com

Specialty section:
This article was submitted to Neuroepidemiology, a section of the journal Frontiers in Neurology

Received: *10 November 2020*
Accepted: *11 March 2021*
Published: *09 April 2021*

Citation:
Abuga JA, Kariuki SM, Kinyanjui SM, Boele van Hensbroek M and Newton CR (2021) Premature Mortality, Risk Factors, and Causes of Death Following Childhood-Onset Neurological Impairments: A Systematic Review. Front. Neurol. 12:627824. doi: 10.3389/fneur.2021.627824

Background: Neurological impairment (NI) and disability are associated with reduced life expectancy, but the risk and magnitude of premature mortality in children vary considerably across study settings. We conducted a systematic review to estimate the magnitude of premature mortality following childhood-onset NI worldwide and to summarize known risk factors and causes of death.

Methods: We searched various databases for published studies from their inception up to 31st October 2020. We included all cohort studies that assessed the overall risk of mortality in individuals with childhood-onset epilepsy, intellectual disability (ID), and deficits in hearing, vision and motor functions. Comparative measures of mortality such as the standardized mortality ratio (SMR), risk factors and causes were synthesized quantitatively under each domain of impairment. This review is registered on the PROSPERO database (registration number CRD42019119239).

Results: The search identified 2,159 studies, of which 24 studies were included in the final synthesis. Twenty-two (91.7%) studies originated from high-income countries (HICs). The median SMR was higher for epilepsy compared with ID (7.1 [range 3.1–22.4] vs. 2.9 [range 2.0–11.6]). In epilepsy, mortality was highest among younger age groups, comorbid neurological disorders, generalized seizures (at univariable levels), untreatable epilepsy, soon after diagnosis and among cases with structural/metabolic types, but there were no differences by sex. Most deaths (87.5%) were caused by non-epilepsy-related causes. For ID, mortality was highest in younger age groups and girls had a higher risk compared to the general population. Important risk factors for premature mortality were severe-to-profound severity, congenital disorders e.g., Down Syndrome, comorbid neurological disorders and adverse pregnancy and perinatal events. Respiratory infections and comorbid neurological disorders were the leading causes of death in ID. Mortality is infrequently examined in impairments of vision, hearing and motor functions.

Summary: The risk of premature mortality is elevated in individuals with childhood-onset NI, particularly in epilepsy and lower in ID, with a need for more studies for vision, hearing, and motor impairments. Survival in NI could be improved through interventions targeting modifiable risk factors and underlying causes.

Keywords: neurological, neurodevelopmental, disability, impairment, mortality, children

INTRODUCTION

Neurological impairments (NI) are a group of disorders resulting from damage to or dysfunction of the central nervous system (1–3). The most prevalent domains of NI include epilepsy, cognitive, sensorineural, and motor impairments (4, 5). The burden of NI varies greatly between and within regions and countries, which is attributed to epidemiological and demographic transitions (6–10). For example, improved child survival and persistence or emergence of risk factors for NI have increased the burden in older children and adolescents in LMICs (11–13).

Current evidence suggests an increased risk of premature mortality or reduced life expectancy among individuals with NI and disability (14–19). For instance, the risk of premature mortality is 2–3 times higher among people with epilepsy compared with the general population (16, 19); the risk is highest in LMICs (16) and childhood-onset seizures (19). The risk of mortality is also higher in structural/metabolic, untreated and intractable epilepsy. The causes of death in epilepsy include: (i) sudden unexplained death in epilepsy (SUDEP); (ii) accidents/burns; or (iii) acute/chronic infectious or non-infectious disease (16, 19), but it is unclear if mortality for other domains of NI is related to these or other different causes. Cohort studies are logistically intensive to conduct, which can influence the extent to which mortality is examined following NI across the world, particularly in LMICs.

We conducted a systematic review to: (i) estimate the magnitude of premature mortality in individuals with childhood-onset NI; (ii) summarize known risk factors; and (iii) describe causes of premature mortality among individuals who died. This evidence is required to inform medical and community-based interventions that might improve the survival and quality of life for individuals with NI and disability.

METHODS

We used the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines (20) and the Centre for Reviews and Dissemination (CRD) recommendations for undertaking reviews in healthcare (21) for the searches, identification, appraisal of eligible studies, synthesis, and reporting of findings in this review. A protocol is registered in the international prospective register of systematic reviews (PROSPERO-registration number CRD42019119239) (22).

Search Strategy and Eligibility Criteria

A search was conducted in the PubMed, EMBASE, and Scopus databases for cohort studies from database inception up to 31st October 2020 using terms in three groups: (i) neurologic impairments, cognitive impairments or intellectual disability, motor impairments, visual impairments, hearing impairments and epilepsy; (ii) mortality, death or survival; (iii) risk factors, predictors or causes of death (see **Table 1** for details of the search strategy in PubMed).

TABLE 1 | Searches in PubMed.

Database	Search terms
PubMed	(((excess mortality OR long-term survival OR life expectancy OR premature death OR death OR premature mortality OR survival) AND (neurodevelopmental disorders [MeSH Terms] or neurologic impairment OR cognitive disability OR motor impairment OR visual impairment OR hearing impairment OR epilepsy)) AND (risk factors OR causes of death OR predictors of mortality)) AND (cohort*) **Filters:** Humans

Two reviewers reviewed the retrieved citations in a two-stage process. In the first stage, the first reviewer (JA) reviewed all titles and available abstracts to identify relevant studies. The second reviewer (SK) independently reviewed 30% of the titles and abstracts; both reviewers compared their lists and resolved disagreements by consensus. In the second stage, both reviewers assessed whether the identified articles met the inclusion criteria.

Inclusion and Exclusion Criteria

We included: (i) original cohort studies of mortality following children with NI in five domains (epilepsy and impairments in cognitive, hearing, vision, and motor functions); (ii) studies with a childhood-onset or diagnosis of impairment (between the ages 0–19 years); (iii) studies reporting all-cause mortality as the primary outcome; and (iv) studies with an appropriate comparison group such as matched controls or the general population. We excluded studies with adulthood-onset of NI ($\geq$20 years), studies of mental and psychiatric problems, studies without an appropriate comparison group, those reporting the same data in different papers, reviews, editorials and studies reported in other languages that could not easily be translated into English.

Definition of Neurological Impairments and Data Extraction

An assessment was done on whether the definitions of NI in each study were in alignment with the International Classification of Diseases (ICD), the Diagnostic and Statistical Manual for Mental Disorders (DSM), or both with the versions dependent on the year of study. Childhood-onset epilepsy was defined according to the International League Against Epilepsy (ILAE)

as the presence of two or more unprovoked seizures occurring within 12 months identified before the age of 18 years (23). Individuals with cognitive impairment, hereafter referred to as intellectual disability (ID), refer to those who had IQ scores <70 or z-scores <−3 based on age-appropriate neuropsychological tests and age-inappropriate adaptive skills with a childhood-onset (24). Historically, cognitive impairment has been conceptualized as structural or functional limitations based on the medical model (2, 25); however, recent definitions have used the term ID to depict the misfit between contextual demands and the person's capabilities (3, 26). Motor impairments referred to limitations in muscle control, movement, or mobility, or complete absence of motor functioning based on valid criteria such as the Gross Motor Function Classification System (GMFCS) (27). Hearing impairment was defined as hearing loss >25–30 dB in the best hearing hear (28) and vision impairment as a deficit in sight presenting with visual acuity worse than 6/12 (29). We extracted data on study setting, population characteristics, cohort sizes, duration of follow-up, comparative measures of mortality risk such as mortality rate ratios (MRR), hazard ratios (HR) and standardized mortality ratios (SMR), and risk factors and causes of death as reported in the individual studies.

Quality of Studies

Guidelines from the Joanna Briggs Institute's (JBI) critical appraisal checklist for cohort studies were used (30) and the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist (31) to appraise the methodological quality of the included studies. Our emphasis on quality was focused on: (i) reliability/sensitivity of NI diagnosis and case ascertainment; (ii) sensitivity/reliability of mortality case and cause of death ascertainment; (iii) representativeness of the study population; (iv) risk of bias e.g., selection bias; (v) follow-up duration; and (vi) use of appropriate analytic methods. We grouped each report into an aggregate score of four classes: class 1 studies represent those studies with an overall score between 75 and 100%; class 2 studies scored between 50 and 74%; class 3 studies scored between 25 and 49% and class 4 studied scored below 25% representing the weakest evidence.

Synthesis of Evidence

We estimated the overall risk of premature mortality separately for each domain of NI because of heterogeneity in the diagnosis of each, and that each impairment has a unique underlying process and prognosis. Summary measures such as the median and range were used for descriptive analysis of SMR, MRR, and HR reported in the included studies. Similarly, we summarized the measures of effect on mortality for each risk factor and cause-specific proportionate mortality per domain of NI. Overall estimates of mortality from primary studies could not be combined in a meta-analysis because there was very high heterogeneity within and between the included studies (32).

RESULTS

Search Results

The results of the systematic search are in **Figure 1**. A total of 24 studies met the inclusion criteria of which 9 (37.5%) were on epilepsy, 10 (41.7%) on ID, 3 (12.5%) on motor-related impairments or cerebral palsy (CP), 1 (4.2%) on vision impairment and 1 on multiple domains of NI (**Supplementary Table 1**). Fourteen studies (58.3%) reported findings from population-based cohorts and 10 (41.7%) from clinical cohorts. Of the 24 studies, 5 (20.8%) were prospective in design while the rest (79.2%) identified study participants retrospectively. Europe provided 11 (45.8%) studies, North America 7 (31.8%) studies, Australia 4 (18.2%) studies, Asia one (4.5%) study, and Africa one study (**Supplementary Table 1**).

Quality of Studies

Most studies (63%) were classified as class 1 or excellent quality (**Supplementary Table 1**) and the median quality score for all studies was 91% (range [55–100]). The median quality score was similar for population-based studies and clinical cohort studies (91% [55–100] vs. 87% [64–100]; $p = 0.76$). The median quality score for retrospective cohort studies was also comparable with the median score for prospective studies (91% [55–100] vs. 100% [82–100], $p = 0.21$).

Epilepsy

Overall Risk of Mortality

The median SMR was 7.1 (range 3.1–22.4) for children with epilepsy (**Table 2**), and none of these studies originated from LMICs. The median SMR for clinical cohort studies (33–35) was 7.5 (range 7.0–22.4) and 6.8 (range 3.1–9.0) for population-based studies (36, 37, 39, 41). One study (38) reported a MRR of 14.9 (95% CI 13.9–16.1) and the other study (40) a hazard ratio of 3.8 (95% CI 3.1–4.7).

Risk of Mortality by Age and Sex

Four studies reported a significantly higher risk of mortality in younger children (33, 36, 38, 39), and mortality was highest among teenagers only in one study (40). Age was not associated with mortality in two reports (37, 41), and the age-mortality association was not reported in two studies (**Table 2**). The risk of mortality comparing boys and girls with epilepsy was similar in most studies (**Table 2**). Only one study (40) reported a significantly higher risk among boys (HR 1.3, [95% CI 1.1–1.5]).

Risk of Mortality by Epilepsy Factors

Most deaths (68.8%) occurred in the first 10 years after epilepsy diagnosis, and mortality declined significantly in the subsequent decades but remained higher compared with the general populations (36, 38, 39). Mortality was higher in structural/metabolic epilepsy compared with epilepsy of genetic or unknown etiology in three studies (34, 35, 41). The risk of mortality was also very high in epilepsy with comorbid brain disorders, some of which were possible causes of epilepsy (**Supplementary Table 2**).

Generalized seizures were associated with an increased risk of mortality at the univariable level, but not at the

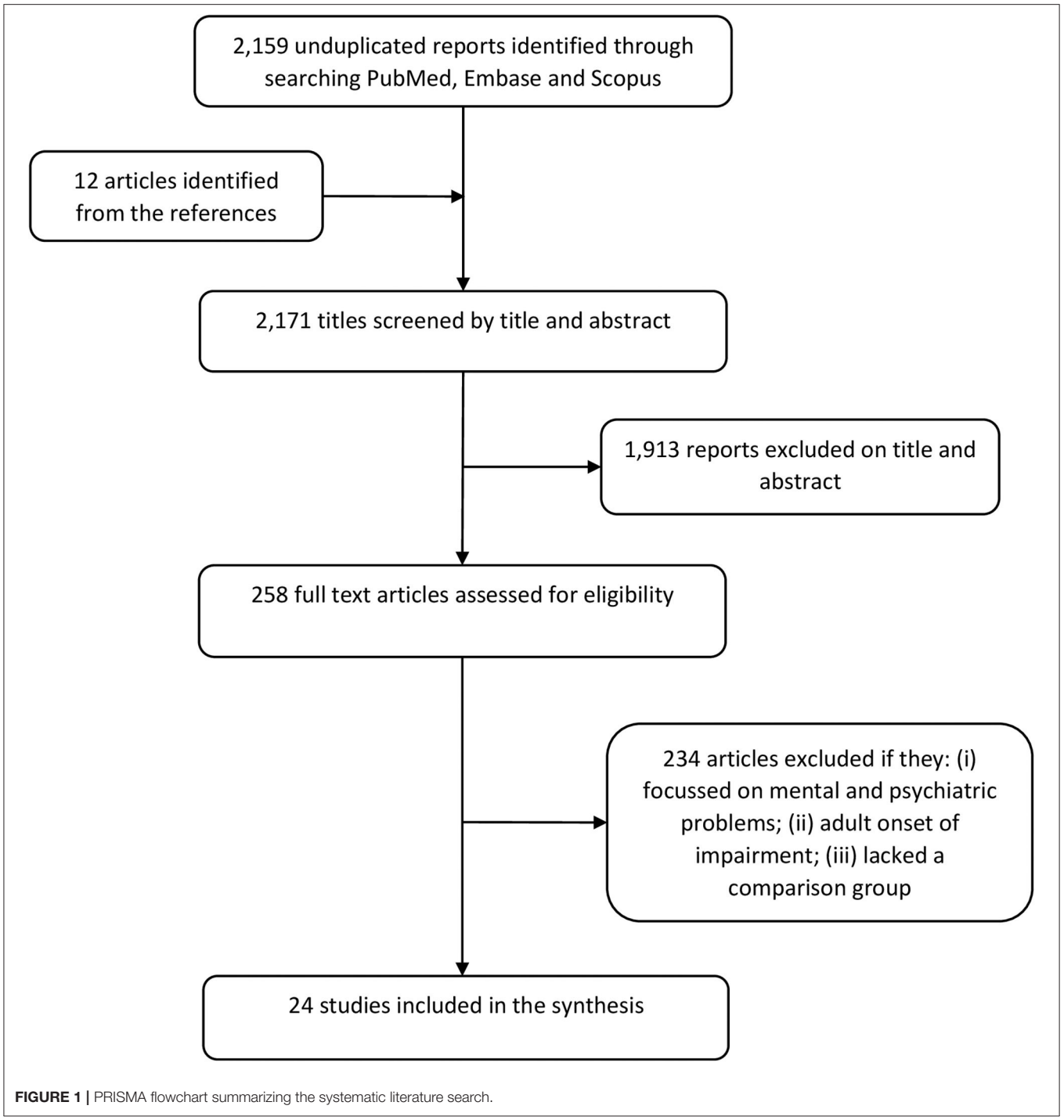

FIGURE 1 | PRISMA flowchart summarizing the systematic literature search.

multivariate level (41). Mortality was increased in epilepsy syndromes such as Lennox-Gastaut syndrome and infantile spasms in one report (36). Most reports included in this review, however, did not report the effect of seizure type on mortality (**Supplementary Table 2**).

Mortality was also higher in: (i) epilepsy patients using more than 2 antiepileptic drugs (AEDs) (33, 39); (ii) patients without a 5 year terminal remission after treatment with AEDs (41); (iii) patients from rural settings (40); and (iv) patients with Medicare health insurance plan compared with private/commercial insurance in one US-based study (40).

Causes of Death in Epilepsy

Most deaths (median percentage 87.5% [range 74.9–90.6]) were caused by non-epilepsy-related causes such as underlying neurological disorders or respiratory problems (**Table 3**); but

TABLE 2 | General study characteristics, overall risk, and mortality by age and sex in children with epilepsy.

Study	Country or State	Population characteristic	Percent quality score (Class)	Cohort size	Follow-up (years)	Measure of mortality (95% CI)	Age-specific mortality ratios (95% CI)	Sex-specific mortality ratios (95% CI)	
								Males	Females
Clinical-based cohort studies									
Ackers et al. (33)	England and Wales	Incident cases aged 0–18 years	64 (2)	6,190	[a]13	SMR 22.4 (18.9–26.2)	SMR of 42.4 (95% CI 33.3–53.2) for 2–11 years; 13.8 (10.4–18.0) for 12–18 years; 20.9 (13.2–31.3) for <2 years	SMR 19.4 (15.5–23.9)	27.1 (20.9–34.5)
Berg et al. (34)	Connecticut, USA	Incident cases aged <16 years	82 (1)	613	[b]7.9	SMR 7.5 (4.38–12.99)	Not reported	2.0%	2.3% (p = 0.77, SMR not estimated)
Callenbach et al. (35)	The Netherlands	Incident cases	99 (1)	472	[a]5	SMR 7.0 (2.4–11.5)	Not reported	SMR 6.6 (2.2–15.5)	7.4 (2.0–19.0)
Population-based cohort studies									
Autry et al. (36)	Atlanta, USA	Incident cohort 10 years	99 (1)	688	[a]26	SMR 3.1 (2.39–3.98)	SMR of 0.3 (0.07–0.93) for <1 year; 6.0 (3.0–10.7) for 1–4 years; 10.1 (5.2–17.6) for 5–9 years; 16.7 (9.5–27.1) for 10–14 years; 3.1 (1.5–5.7) for 15–19 years; 3.3 (1.6–5.9) for 20–24 years; and 1.7 (0.0–9.7) for 25–34 years.	SMR 3.0 (2.1–4.0)	3.4 (2.1–5.1)
Camfield et al. (37)	Nova Scotia, Canada	Incident cases <17 years	99 (1)	686	[b]13·9	SMR 7.1 (3.2–10.9)	Age at onset 1–5 years vs. <1 year (RR_{ADJ} 1.5, 95% CI 0.6–3.9), and 6–16 years (RR_{ADJ} 1.7, 95% CI 0.5–6.1).	Multivariable relative risk (girl vs. boy) 1.3 (0.6–2.9)	
Christensen et al. (38)	Aarhus, Denmark	Incident cases	99 (1)	25,244	[b]13.7	MRR 14.9 (13.9–16.1)	Short-term mortality (<1 year) in epilepsy with an onset before 5 years (MRR 41.5, 95% CI 35.4–48.3); long-term mortality (>1 year) (MRR 21.6, 95% CI 19.5–23.8).	Cumulative mortality 20 years after first epilepsy diagnosis 7.6% (6.8–8.4)	Cumulative mortality 5.8% (5.1–6.5)
Nickels et al. (39)	Rochester, MN, USA	Incident cases <18 years	73 (2)	467	[b]7.9	SMR 9.0 (5.4–14.4)	11 of 16 (69.8%) deaths occurred among children aged 1 month−10 years; 4 (25%) deaths occurred in the 10–19 years age group; one (6.3%) death occurred in the 20 years or older age group.	Not reported	Not reported
Selassie et al. (40)	South Carolina, USA	Incident cases ages <19 years	91 (1)	13,098	[a]11	HR 3.8 (3.1–4.7)	Mortality in children aged 13–18 years vs. 0–5 years (HR 1.5, 95% CI 1.2–1.9); 6–12 years (HR 0.9, 95% CI 0.7–1.1).	Males vs. females HR 1.28 (1.08–1.51)	
Sillanpaa and Shinnar (41)	Turku, Finland	Incident and prevalent cases <16 years	99 (1)	245	[b]40	SMR 6.4 (5.9–7.0)	Age at onset (<2 years vs. ≥ 2 years): HR 1.7 (0.8–3.5).	Mortality rate 7.3 deaths/1,000 (5.2–10.2)	6.41 deaths/1,000 (4.4–9.4)

[a] *range.*
[b] *mean or median.*

TABLE 3 | Proportionate mortality or cause-specific mortality rates/ratios for epilepsy and non-epilepsy related causes.

Study	Epilepsy-related	Epilepsy-unrelated	Epilepsy-unrelated sub-categories	Unknown
Ackers et al. (33)	18 (11.9%)	110 (72.8%)	Underlying neurological disorders (110)	23 (15.2%)
Berg et al. (34)	2 (15.4%)	10 (76.9%)		1 (7.7%)
Callenbach et al. (35)	None	9 (88.9%)	Respiratory problems (8) transtentoria and brain herniation (1)	1 (11.1%)
Camfield et al. (37)	2 (7.7%)	24 (92.3%)	Pneumonia (14); infection or sepsis (3), suicide (2); shunt malfunction (1); pulmonary embolism (1); congestive heart failure (1); gastroesophageal reflux and failure to thrive (1); and homicide (1)	None
Christensen et al. (38)	None	766 (95.4%)		37 (4.6%)
Nickels et al. (39)	2 (12.5%)	14 (87.5%)		None
Sillanpaa and Shinnar (41)	33 (55%)	26 (43%)	Pneumonia (12)	1 (2.0%)
Autry et al. (36)	Number/proportion not reported	Not reported	Cause specific SMR for neurological causes 19.4 (10.0–33.8); Infections and tumors 7.4 (5.4–9.9); Cardiac deaths 3.4 (0.7–9.9)	Not reported
Selassie et al. (40)	2.8/1,000 person-years of observation (pyo)	Not reported	Developmental conditions: 5.9/1,000 pyo; cardiovascular disorders excluding congenital malformations 4.4/1,000 pyo; injuries from external causes 3.8/1,000 pyo	Not reported

infections, tumors, cardiovascular disorders, and injuries were also reported in other studies (36, 40). Status epilepticus and sudden unexplained death in epilepsy (SUDEP) were the most common causes of epilepsy-related mortality (**Table 3**).

Intellectual Disability

Overall Risk of Mortality

The median SMR was 2.9 (range 2.0–11.6) for ID; however, one study reported a HR of 6.1 (95% CI 5.3–7.0) after 25 years of follow-up, and two studies reported crude mortality ratios of 1.8 and 1.7, respectively (**Table 4**). There was no difference in the overall risk of mortality between studies classifying ID using the International Classification of Disease Ninth Revision (ICD-9) or earlier versions and the Diagnostic and Statistical Manual of Mental Disorders Fourth Revision (DSM-IV) or earlier versions compared with studies classified using the ICD-10 or DSM-5. Studies of ID mainly utilized data from population-wide or state-wide disability service providers linked retrospectively with registries of mortality. Information about the definition of ID and sources of mortality data for each report are provided in **Supplementary Table 3**.

Mortality by Age and Sex

Mortality was highest in younger age-groups in all studies (**Table 4**). The median SMR for the ages 0–19 years was 13.8 (range 6.7–21.6). Most studies reported a monotonic decline of mortality ratios with increasing age, and mortality was slightly higher than the respective general populations in older age groups (60+ years), suggesting a healthy survivor effect (**Table 4**). The median SMR was 4.1 (range 2.6–16.6) for females and 2.5 (range 1.6–9.8) for males. Mortality was similarly high in males and females in a study excluding ID cases with significant physical impairment or comorbid/underlying degenerative conditions (49).

Mortality by Factors Related to Intellectual Disability

The risk of premature mortality was consistently higher in severe or profound ID compared with mild or moderate ID (**Table 5**). Three studies, however, did not report mortality by the severity of ID (44, 46, 50). Mortality in ID was significantly increased by genetic disorders (Down syndrome and Fragile X), maternal alcohol use, low-birth weight and postnatal injury, and the presence and number of comorbid neurological conditions such as epilepsy and cerebral palsy (43, 45, 48, 51) (**Table 5**).

Causes of Death in Intellectual Disability

Respiratory infections (34%), accidents (18%), and epilepsy (10.7%) were the most common causes of death in ID in one study from Western Australia (43); while, the leading causes of death in a Swedish study were congenital malformations (SMR 46.3 [32.9–65.0]), neurological diseases (SMR 9.7 [5.5–17.0]), mental disorders (SMR 4.0 [1.9–8.4]), and respiratory diseases (SMR 3.3 [2.0–5.5]) (45). Diseases of the respiratory system (21.8%), the circulatory system (19.1%) and the nervous system (13.0%) were the most common causes of death in the Scottish study (51). Another Scottish study (50) identified diseases of the nervous system (33%), congenital malformations, deformations and chromosomal anomalies (22%), and nutritional, metabolic and endocrinal diseases (8%) as the most frequent causes of death. Six out of 10 (60%) studies did not report the causes of death in ID.

TABLE 4 | General study characteristics, overall risk, and sociodemographic risk factors for mortality in people with intellectual disability.

Author and Year of publication	Country or state	Percent quality score (Class)	Cohort size	Follow-up* (years)	Overall measure of mortality (95% CI)	Risk of mortality by age	Risk of mortality by sex	
							Females	Males
Arvio et al. (42)	Finland	64 (2)	–	15	SMR 2.9 (2.9–3.0)	SMR 11.6 (95% CI 9.6–13.8) for <15 years; SMRs decreased with increasing age; SMR 2.0 (1.95–2.14) for >60 years	SMR 4.1 (4.0–4.3)	2.4 (2.3–2.4)
Bourke et al. (43)	Western Australia	91 (1)	10,593	25	Adjusted HR 6.1 (5.3–7.0)	aHR 6.0 (95% CI 4.8–7.6) for ages 1–5 years, 12.6 (9.0–17.7) for 6–10 years, and 4.9 (3.9–6.1) for 11–25 years.		aHR 0.8 (0.6–1.0) (male vs. female)
Florio and Trollor (44)	New South Wales, Australia	55 (2)	40,705	6	SMR 2.5 (2.3–2.6)	0–19 years mortality rate ratio (MRR)= 4.6, 20–49 years = 4.7, 50–69 years=2.6, 70–79 years=1.7, and 80+ years MR=0.8	SMR 4.3 (3.8–4.7)	2.5 (2.3–2.8)
Forsgren et al. (45)	Vasterbotten, Sweden	99 (1)	1,478	7	SMR 2.0 (1.7–2.3)	SMR 16 (10-24) for those 0–19 years	SMR 2.6 (2.0–3.3)	1.6 (1.2–2.0)
Lauer and McCallion (46)	New York, United States (USA)	73 (2)	–	3	MR 1.8	MRR was 5.9 for 18–24 years and 1.8 for those 75+ years.	Mortality rate 11.2/1,000	10.9/1,000
McCarron et al. (47)	Ireland	82 (1)	31,943	10	SMR 3.9 (3.7–4.0)	SMR 6.7 (5.9–7.5) for 0–19 years; SMR decreased with increasing age; SMR 2.7 (2.4–3.0) for 80+ years	SMR 4.9 (4.6–5.2)	3.1 (2.9–3.3)
Tyrer et al. (48)	Leister shire and Rutland, United Kingdom	91 (1)	2436	10	SMR 3.2 (2.9–3.6)	SMR 11.5 (8.1–15.8) in the 20s; diminished in older ages; SMR 1.5 (1.2–1.8) for 70+ years	SMR 3.6 (3.1–4.2)	2.9 (2.5–3.3)
Shavelle et al. (49)	California, USA	91 (1)	64,207	30	MR 1.7	5–19 years (MRR = 1.4); 20–39 years = 2.0; 40–59 years = 1.8; 60+ years = 1.3	Mortality ratio = 1 (females and males)	
Smith et al. (50)	Scotland	73 (2)	18,278	5	SMR 11.6 (9.6–14.0)	5–14 years SMR = 21.6 (16.6–28.2); ≥ 15 years but <25 years SMR = 7.7(5.9–10.2)	SMR of 16.6 (12.2–22.6)	9.8 (7.7–12.5)
Cooper et al. (51)	Glasgow, Scotland	91 (1)	961	17	SMR 2.2 (2.0–2.5)	SMR of 18.7 (0.4–37.1) for 15–25 years; 2 (1.3–7.1) for 26–35 years; 3.9 (2.3–5.4) for 36–45 years; 3.8 (2.9–4.7) for 46–55 years; and 1.9 (1.6–2.1) for >55 years	SMR 3.5 (2.9–4.1)	1.7 (1.4–2.0)

range.

Cerebral Palsy

Overall Risk of Mortality

We identified 3 studies of CP (52–54) which assessed the effect of motor impairment on the risk of mortality in children. The risk of mortality in CP was highest between the ages of 2–15 years compared to the general population rates (SMRs>25) and declined steadily to 2–3 times higher than the population rates by the age of 40 years (52, 54). The study from Bangladesh did not estimate the SMR but crude mortality rates (19.5 per 1,000 person-years) and mortality was highest in the youngest age group (<5 years).

Motor Impairment and Mortality in Cerebral Palsy

Lack of independent ambulation was the strongest predictor of mortality (adjusted HR 6.1, [3.3–11.8]) in the Australian study (54). However, motor impairment was a weaker predictor of premature mortality (MRR 1.4 [1.1–1.7]) than ID (MRR 2.1, [1.9–2.4]) in the Western Australian study

TABLE 5 | Mortality by severity and aetiology of intellectual disability.

Author and Year of publication	Overall estimate of mortality (95% CI)	Risk of mortality by level of disability (95% CI)						Mortality by cause of ID or comorbid neurological disorders (95% CI)
		Mild	Mild or moderate	Moderate	Severe	Severe or profound	Profound	
Arvio et al. (42)	SMR 2.9 (2.9–3.0)	2.3 (2.2–2.4)	–	–	3.4 (3.3–3.5)	–	–	–
Bourke et al. (43)	Adjusted HR 6.1 (5.3–7.0)	–	3.2 (2.6–3.9)	–	40.6 (33.4–49.2)	–	–	Biomedical causes 24.4 (20.7–28.7); unknown causes 1.8 (1.3–2.3); Autism 2.0 (0.8–4.7)
Florio and Trollor (44)	SMR 2.5 (2.3–2.6)	–	–	–	–	–	–	–
Forsgren et al. (45)	SMR 2.0 (1.7–2.3)	1.8 (1.1–2.7)		1.5 (1.1–2.0)	2.0 (1.5–2.6)		8.1 (5.6–11.7)	Mental Retardation 1.7 (1.4–2.0); MR + Epilepsy 5.0 (3.3–7.5); MR + Epilepsy + CP 5.8 (3.4–9.8)
Lauer and McCallion (46)	MR 1.8	–	–	–	–	–	–	–
McCarron et al. (47)	SMR 3.9 (3.7–4.0)	5.0%	–	7.8%	14.6%		24.8%	–
Tyrer et al. (48)	*SMR 3.2 (2.9–3.6)	–	–	–	–	–	–	Comorbidity with Downs Syndrome 7.60
Shavelle et al. (49)	Mortality ratio was 167%		165%	–	–	185%	–	–
Cooper et al. (51)	SMR 2.2 (2.0–2.5)	1.6 (1.3–1)	–	2.1 (1.6–2.6)	2.8 (2.1–3.4)	–	4.1 (3.1–5.2)	ID with Downs Syndrome 5.3 (4.0–6.6); ID without Downs Syndrome 1.9 (1.7–2.2)
Smith et al. (50)	SMR 11.6 (9.6–14.0)	–	–	–	–	–	–	–

**The study by Tyrer et al. (48) excluded participants with mild ID; those with moderate to profound ID included as one group in the analysis.*

(52). The latter study further reported that the overall disability score was a better predictor of mortality than motor impairment and ID, separately. While mortality was elevated among those with hearing impairment (adjusted HR 2.9 [1.2–6.7]) and swallowing difficulties (adjusted HR 2.3 [1.0–4.9]) in the study from Bangladesh (53), the risk of severe motor impairment (GMFCS levels III-V) on mortality did not reach statistical significance (adjusted HR 2.4 [0.7–8.4]).

Causes of Death in Cerebral Palsy

CP as an underlying cause of death was accountable for 79% of deaths in the Western Australian study with 59% of the fatalities directly resulting from respiratory problems (52). Similarly, respiratory causes were the most common direct causes of death in the Australian study (54). Meningitis (31.0%) and pneumonia (27.6%) were the leading causes of death in the study from Bangladesh (53) and most children who died were either severely malnourished or had feeding problems.

Vision Impairment and Mortality

The study from Malmöhus, Sweden (55) reported an odds ratio for mortality of 60.11 (95% CI 35.2–97.9) in visually impaired children and adolescents compared with an age- and sex-matched sample from the population. Most of the visually impaired children had additional impairments such as CP and ID, and respiratory diseases were the most common cause of death in this study.

Mortality in Multiple Domains of Childhood Neurological Impairment

The study from Kenya (14) investigated the long-term risk of premature mortality in children aged 6–9 years with NI in 5 domains (epilepsy and impairments in cognitive, hearing, vision and motor functions) compared with an age-matched sample from the general population. The overall risk of mortality was >3 times higher among those with any impairment compared to the general population (SMR 3.2 [1.7–5.5]). Developmental delay (adjusted HR 18.9, [2.2–160.4]) and severe malnutrition (20.9, [3.14–139.11]) increased the risk of mortality, and infections such as HIV and accidents were the most common causes of death.

DISCUSSION

The studies in this systematic review reported that the measures of mortality were significantly greater in children with NI

compared with the general population. The estimates were greatest for CP (SMR > 25) and lowest for ID (SMR = 2.9), with few studies reporting mortality outcomes for visual impairment, and no data for hearing impairment. Clinical cohort studies had higher estimates than population-based cohort studies, probability of severity bias in the former. These estimates should be interpreted carefully since the methodology and follow-up periods differed across the studies, complicating the combining of estimates across studies. The risk of mortality following NI depends on younger age, the severity of the primary impairment and the number and severity of comorbid disorders. Most causes of the mortality including infections, cardiovascular diseases, and tumors were unrelated to NI.

Interpretation for Epilepsy

Mortality was significantly higher in children with epilepsy compared with the reference population similar to a previous review of mortality in pediatric epilepsy (56), and so were the factors underlying mortality in NI (16, 19, 57). Early-onset epilepsy occurs in a period of vulnerability to adversities such as infections and severe epileptic encephalopathy, which can increase the risk of mortality (58). Structural epilepsy complicated by head injuries and infections of the brain has poor outcomes e.g., mortality, particularly in LMICs (16). Mortality risks are highest in intractable epilepsy with reduced responsiveness to AEDs (19). Mortality was generally similar in both males and females, although few studies reported high rates in males, probably related to their vulnerability to accidents from injuries and violence (40). Mortality risk reduces in subsequent decades (59); implying that those surviving may have responded better to AEDs or had the fewest risk factors for mortality. There are other context-specific factors for increased mortality such as residence and type of insurance which are related to social, economic and cultural disadvantages that influence access to AED treatment.

Lastly, our findings concur with previous studies that most deaths are caused by underlying comorbid brain disorders and respiratory infections and not by seizures or epilepsy (57, 60). Comorbidity conditions occur more frequently in those with severe epilepsy, who are more likely to die (61). Reducing mortality requires the management of both epilepsy and other comorbid conditions. Risk is elevated in generalized tonic seizures, which are often reported by many who die from SUDEP (62, 63), a condition that is not accurately documented in LMICs because coroner's autopsy reports are unavailable. Like SUDEP, status epilepticus was an infrequently identified cause of mortality compared to non-epilepsy causes, yet it occurs in 25–30% of people with epilepsy (64) and is known to increase mortality in children (16, 65). In many LMICs, status epilepticus is a complication of endemic infections that should be additionally prevented or managed (66) and may be due to poor access to AEDs, inappropriate treatment and delayed initiation of treatment (67). Most reported causes of death were non-epilepsy related e.g., cardiovascular problems, and respiratory problems, and so a comprehensive care and public health plan for children with epilepsy that includes other medical conditions is advised, especially in settings where these are endemic.

Interpretation for Intellectual Disability

This review corroborates findings from a previous review (18) that the overall risk of mortality in people with childhood-onset ID is higher than the general population. Mortality risk was, however, lower than that for epilepsy, which might imply better care for ID patients in HICs, use of cohorts of milder ID, or misreported due to stigma. Underreporting of deaths due to stigma related to ID intellectual disability not only underestimates the prevalence in general but also reduces the SMRs or relative risk for deaths related to ID if these deaths are classified under the general population. There was no single study of ID from LMICs, where prevalence and associated mortality may be higher (68). A median SMR of 2.9 for ID is probably an underestimate because a majority of the studies utilized a retrospective design which may be subject to three types of bias: (i) over-representation of severe ID in clinic samples; (ii) loss to follow-up bias and; (iii) bias due to incomplete data linkage for follow up which might under-uncertain the risk of mortality artificially lowering the SMRs (69). The risk of mortality is highest in the youngest age-groups, decreasing steadily with age, which is well-appreciated in the literature (70–72). The SMRs are higher in females than males, probably consistent with inequality in access to care (18). The risk of mortality is higher in severe-profound ID, which expectedly would be associated with significant functional limitation and disability.

Genetic causes of ID were important risk factors for mortality, and Fragile X syndrome was not exceptional in this review, with a remarkable reduction in life expectancy compared with general populations. It is known that neuropsychiatric problems like epilepsy are very common in ID (73–75) and can increase mortality. Pregnancy and perinatal factors were important risk factors suggesting they will not only cause ID but will worsen its prognosis including premature mortality. Noteworthy was maternal drug abuse and mortality in ID which can be explained by poor parenting of the affected children (76). Children with ID are susceptible to accidents that can be fatal and safety measures and close supervision are encouraged. Respiratory infections are important causes of morbidity and mortality which should be prevented and managed to improve outcomes in children with ID.

Interpretation for Motor Impairment in Cerebral Palsy

Cerebral palsy (CP) had the greatest risk of mortality, which is well-recognized in other studies (77–79). Motor impairments and ID were the strongest predictors of premature mortality in CP in the studies included in this review; both comorbidities are debilitating complications of CP. Two previous studies (80, 81), concur that additional neurological comorbidity worsens the disability score significantly reducing the chances of survival among children with CP. Disability scores, therefore, offer a better prediction of survival compared with motor impairment and/or ID, separately. Rehabilitative therapies to manage physical impairment while optimizing the mobility

of children with CP may improve the quality of life and reduce the risk of mortality (82, 83). Respiratory infections are the most common direct causes of death perhaps because CP impairs breathing and respiratory hygiene. Studies from LMICs settings highlight the significance of preventing other infections such as HIV, pneumonia, and meningitis as well as severe malnutrition to improve CP outcomes.

Interpretation for Vision and Hearing Impairment

Mortality is higher among visually impaired children having comorbid and severe neurological disorders, but the study did not provide SMRs, affecting comparisons with the general population. Only cause-specific estimates of mortality would separate the effects of loss of sight on mortality from the comorbid disorders, but these data were not available for this review. The risk of mortality for visual impairment alone was not significantly higher in the Kenyan study, where children with five domains of NI were followed for over 14 years to determine mortality, which was due to small numbers. There was no single study on hearing impairments and associated all-cause mortality, which could be due to publication bias (4) or because this impairment is often overlooked to follow up of mortality outcomes. In this review, children with CP had hearing impairments which increased the risk of premature mortality (53), suggesting the need to give attention to morbidity and disability of deaf children.

Limitations

The distribution of the studies identified by our searches was sparse, with most from HICs such as Europe and North America with fewest in Asia and Africa and none from South America. This may affect the generalizability of these estimates across continents or other countries, where specific studies are needed. The follow-up periods were variable, whereby studies with shorter follow-ups may underestimate the true burden of premature mortality. Follow-up times were reported differently in the primary studies, for instance, most studies neither reported the mean nor the median follow-up duration in years. This differential reporting of the follow-up duration hindered the estimation of weighted median SMRs. There is often incomplete documentation and certification of deaths due to a lack of functional vital registration systems, which affects case ascertainments yielding lower estimates of mortality. Because the grading process involved qualitative judgment by the reviewers, the interpretation of study methods and the application of criteria might be inconsistent and unreliable. The study populations from which we obtained the data vary greatly in terms of population characteristics, health, and social systems. Clinical-based estimates may overrepresent severe forms of NI that have additional risk for mortality; most LMICs have limited resources to do follow-up studies of mortality, often doing these studies in high-risk zones that are not representative of other low-risk areas. Despite these limitations, our review obtained critical evidence that might increase the survival of children with childhood-onset NI and disability.

SUMMARY

The risk of premature mortality is elevated in children with NI, or adults with childhood-onset NI, it being higher in CP and epilepsy, and lower in ID. There are few SMR studies for visual, hearing, and motor impairments. We recommend future population-based follow-up studies for multiple domains of NI in children, especially for those with visual, hearing, and motor impairments, and in LMICs where there is a dearth of evidence. Few studies in this review originated from LMICs, yet these countries have a concentration of risk factors and the highest burden of childhood NI and disability. The similarity of risk factors and causes of death across the five domains of NI provides an opportunity for integration of preventive, curative and rehabilitative services. Integrated interventions targeting modifiable risk factors, for instance, improving access to AEDs and prompt treatment of childhood epilepsy as well as caregiver/parental training, child supervision, and prevention of respiratory infections for children with ID, immunization, and improved nutrition in LMICs, are required to improve survival and quality of life among the affected children and families. It worth advising the families that premature death among children cannot be preventable in the presence of genetic causes for NI conditions that are untreatable and progressive such as mucopolysaccharidosis.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

JA: protocol development, searches, identification, appraisal of eligible studies, data synthesis, and writing the first draft of the manuscript. SKa: protocol development, searches, identification, and subsequent review of drafts. SKi: protocol development and review of subsequent drafts. MB: protocol development and review of subsequent drafts. CN: protocol development and review of subsequent drafts. All authors: have read and approved the final manuscript.

FUNDING

JA is a Ph.D. fellow funded by the DELTAS (Developing Excellence in Leadership, Training, and Science) Africa Initiative. The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS)'s Alliance for Accelerating Excellence in Science in Africa (AESA) and supported by the New Partnership for Africa's

Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust [grant 107754/Z/15/Z- DELTAS Africa Sub-Saharan Africa Consortium for Advanced Biostatistics (SSACAB) programme] and The Initiative To Develop African Research Leaders [DEL-15-003] and the UK government. SKa was supported by MQ: Transforming Mental Health (MQF17-18 Kariuki). The views expressed in this review are those of the author(s) and not necessarily those of AAS, NEPAD Agency, Wellcome Trust or the UK government. The authors had full access to the study data and had the final responsibility for the decision to submit this review.

ACKNOWLEDGMENTS

We acknowledge René Spijker-an information specialist at the Cochrane Netherlands, and Mary Bitta-a researcher at the KEMRI-Wellcome Research Programme, for their invaluable ideas in developing the search strategy for this systematic review.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2021.627824/full#supplementary-material

REFERENCES

1. Mung'ala-Odera V, Newton CR. Identifying children with neurological impairment and disability in resource-poor countries. *Child Care Health Dev.* (2007) 33:249–56. doi: 10.1111/j.1365-2214.2006.00714.x
2. WHO. *International Classification of Impairments, Disabilities, and Handicaps: A Manual of Classification Relating to the Consequences of disease, Published in Accordance With Resolution WHA29.35 of the Twenty-ninth World Health ASSEMBLY, May 1976.* World Health Organization 1980. Available online at: https://apps.who.int/iris/handle/10665/41003
3. WHO. *International Classification of Functioning, Health and Disability (ICF).* (2001). Available online at: https://www.who.int/classifications/icf/en/
4. Bitta M, Kariuki S, Abubakar A, Newton C. Burden of neurodevelopmental disorders in low and middle-income countries: a systematic review and meta-analysis. *Wellcome Open Res.* (2017) 2:121. doi: 10.12688/wellcomeopenres.13540.1
5. Mung'ala-Odera V, Meehan R, Njuguna P, Mturi N, Alcock KJ, Newton CR. Prevalence and risk factors of neurological disability and impairment in children living in rural Kenya. *Int J Epidemiol.* (2006) 35:683–8. doi: 10.1093/ije/dyl023
6. Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study (2010). *Lancet.* (2012) 380:2197–223. doi: 10.1016/S0140-6736(12)61689-4
7. Newton CR. Global burden of pediatric neurological disorders. *Semin Pediatr Neurol.* (2018) 27:10–5. doi: 10.1016/j.spen.2018.03.002
8. Olusanya BO, Davis AC, Wertlieb D, Boo N-Y, Nair MKC, Halpern R, et al. Developmental disabilities among children younger than 5 years in 195 countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study (2016). *Lancet Global Health.* (2018) 6:e1100–21. doi: 10.1016/S2214-109X(18)30309-7
9. Olusanya BO, de Vries PJ. Nurturing care for children with developmental disabilities: a moral imperative for sub-Saharan Africa. *Lancet Child Adolesc Health.* (2018) 2:772–4. doi: 10.1016/S2352-4642(18)30281-5
10. Wang H, Liddell C, Coates M, Mooney M, Levitz C, Schumacher A, et al. Global, regional, and national levels of neonatal, infant, and under-5 mortality during 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet.* (2014) 384:957–79. doi: 10.1016/S0140-6736(14)60497-9
11. Kassebaum N, Kyu HH, Zoeckler L, Olsen HE, Thomas K, Pinho C, et al. Child and Adolescent health from 1990 to 2015: findings from the global burden of diseases, injuries, and risk factors 2015 study. *JAMA Pediatr.* (2017) 171:573–92. doi: 10.1001/jamapediatrics.2017.0250
12. Newton C. Neurodevelopmental disorders in low- and middle-income countries. *Dev Med Child Neurol.* (2012) 54:1072. doi: 10.1111/j.1469-8749.2012.04384.x
13. Scherzer AL, Chhagan M, Kauchali S, Susser E. Global perspective on early diagnosis and intervention for children with developmental delays and disabilities. *Dev Med Child Neurol.* (2012) 54:1079–84. doi: 10.1111/j.1469-8749.2012.04348.x
14. Abuga JA, Kariuki SM, Kinyanjui SM, Boele Van Hensbroek M, Newton CRJC. Premature mortality in children aged 6–9 years with neurological impairments in rural Kenya: a cohort study. *Lancet Global Health.* (2019) 7:e1728–35. doi: 10.1016/S2214-109X(19)30425-5
15. Charlson FJ, Baxter AJ, Dua T, Degenhardt L, Whiteford HA, Vos T. Excess mortality from mental, neurological, and substance use disorders in the Global Burden of Disease study 2010. In: Patel V, Chisholm D, Dua T, Laxminarayan R, Medina-Mora MEE, editors. *Mental, Neurological, and Substance Use Disorders: Disease Control Priorities,* 3rd ed. Washington, DC: International Bank for Reconstruction and Development/the World Bank (2016). p. 4.
16. Levira F, Thurman DJ, Sander JW, Hauser WA, Hesdorffer DC, Masanja H, et al. Premature mortality of epilepsy in low- and middle-income countries: a systematic review from the Mortality Task Force of the International League Against Epilepsy. *Epilepsia.* (2017) 58:6–16. doi: 10.1111/epi.13603
17. Ngugi AK, Bottomley C, Fegan G, Chengo E, Odhiambo R, Bauni E, et al. Premature mortality in active convulsive epilepsy in rural Kenya: causes and associated factors. *Neurology.* (2014) 82:582–9. doi: 10.1212/WNL.0000000000000123
18. O'Leary L, Cooper SA, Hughes-McCormack L. Early death and causes of death of people with intellectual disabilities: a systematic review. *J Appl Res Intellect Disabil.* (2018) 31:325–42. doi: 10.1111/jar.12417
19. Thurman DJ, Logroscino G, Beghi E, Hauser WA, Hesdorffer DC, Newton CR, et al. The burden of premature mortality of epilepsy in high-income countries: a systematic review from the Mortality Task Force of the International League Against Epilepsy. *Epilepsia.* (2017) 58:17–26. doi: 10.1111/epi.13604
20. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol.* (2009) 62:e1–34. doi: 10.1016/j.jclinepi.2009.06.006
21. Tacconelli E. Systematic reviews: CRD's guidance for undertaking reviews in health care. *Lancet Infect Dis.* (2010) 10:226. doi: 10.1016/S1473-3099(10)70065-7
22. NHS. *PROSPERO: International Prospective Register of Systematic Reviews York, UK YO10 5DD.* Centre for Reviews and Dissemination, University of York (2018). Available online at: https://www.crd.york.ac.uk/prospero/ (accessed July 31, 2018).
23. Scheffer IE, Berkovic S, Capovilla G, Connolly MB, French J, Guilhoto L, et al. ILAE classification of the epilepsies: position paper of the ILAE Commission for Classification and Terminology. *Epilepsia.* (2017) 58:512–21. doi: 10.1111/epi.13709
24. APA. *Diagnostic and Statistical Manual of Mental Health Disorders.* Washington, DC: American Psychological Association (2013).
25. Wehmeyer ML, Shogren K, Angel Verdugo M, Nota L, Soresi S, Lee S-H, et al. Cognitive impairment and intellectual disability. Special education international perspectives: biopsychosocial, cultural, and disability aspects. *Adv Spec Educ.* (2014) 27:55–89. doi: 10.1108/S0270-4013_2014_0000027002

26. WHO. *WHO Global Disability action Plan 2014-2021: Better Health for All People With Disability*. Geneva: World Health Organization, 2015 (2015).
27. Palisano R, Rosenbaum P, Walter S, Russell D, Wood E, Galuppi B. Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Dev Med Child Neurol.* (1997) 39:214–23. doi: 10.1111/j.1469-8749.1997.tb07414.x
28. WHO. *Deafness and Hearing Loss*. (2020). Available online at: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss
29. WHO. *Blindness and Vision Impairment*. (2019). Available online at: https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment (accessed September 24, 2019).
30. Sandeep Moola ZM, Tufanaru C, Aromataris E, Sears K, Sfetc R, Currie M, et al. Chapter 7: systematic reviews of etiology and risk. In: Aromataris EMZ, editor. *Joanna Briggs Institute Reviewer's Manual.* The Joanna Briggs Institute (2017). Available online at: https://wiki.jbi.global/display/MANUAL/Chapter+7%3A+Systematic+reviews+of+etiology+and+risk
31. Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLoS Med.* (2007) 4:e297. doi: 10.1371/journal.pmed.0040297
32. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA*. (2000) 283:2008–12. doi: 10.1001/jama.283.15.2008
33. Ackers R, Besag FM, Hughes E, Squier W, Murray ML, Wong IC. Mortality rates and causes of death in children with epilepsy prescribed antiepileptic drugs: a retrospective cohort study using the UK General Practice Research Database. *Drug Saf.* (2011) 34:403–13. doi: 10.2165/11588480-000000000-00000
34. Berg AT, Shinnar S, Testa FM, Levy SR, Smith SN, Beckerman B. Mortality in childhood-onset epilepsy. *Arch Pediatr Adolesc Med.* (2004) 158:1147–52. doi: 10.1001/archpedi.158.12.1147
35. Callenbach PM, Westendorp RG, Geerts AT, Arts WF, Peeters EA, van Donselaar CA, et al. Mortality risk in children with epilepsy: the Dutch study of epilepsy in childhood. *Pediatrics.* (2001) 107:1259–63. doi: 10.1542/peds.107.6.1259
36. Autry AR, Trevathan E, Van Naarden Braun K, Yeargin-Allsopp M. Increased risk of death among children with Lennox-Gastaut syndrome and infantile spasms. *J Child Neurol.* (2010) 25:441–7. doi: 10.1177/0883073809348355
37. Camfield CS, Camfield PR, Veugelers PJ. Death in children with epilepsy: a population-based study. *Lancet.* (2002) 359:1891–5. doi: 10.1016/S0140-6736(02)08779-2
38. Christensen J, Pedersen CB, Sidenius P, Olsen J, Vestergaard M. Long-term mortality in children and young adults with epilepsy–a population-based cohort study. *Epilepsy Res.* (2015) 114:81–8. doi: 10.1016/j.eplepsyres.2015.05.001
39. Nickels KC, Grossardt BR, Wirrell EC. Epilepsy-related mortality is low in children: a 30-year population-based study in Olmsted County, MN. *Epilepsia.* (2012) 53:2164–71. doi: 10.1111/j.1528-1167.2012.03661.x
40. Selassie AW, Wilson DA, Wagner JL, Smith G, Wannamaker BB. Population-based comparative analysis of risk of death in children and adolescents with epilepsy and migraine. *Epilepsia.* (2015) 56:1957–65. doi: 10.1111/epi.13219
41. Sillanpaa M, Shinnar S. Long-term mortality in childhood-onset epilepsy. *N Engl J Med.* (2010) 363:2522–9. doi: 10.1056/NEJMoa0911610
42. Arvio M, Salokivi T, Tiitinen A, Haataja L. Mortality in individuals with intellectual disabilities in Finland. *Brain Behav.* (2016) 6:e00431. doi: 10.1002/brb3.431
43. Bourke J, Nembhard WN, Wong K, Leonard H. Twenty-five year survival of children with intellectual disability in Western Australia. *J Pediatr.* (2017) 188:232–9.e2. doi: 10.1016/j.jpeds.2017.06.008
44. Florio T, Trollor J. Mortality among a cohort of persons with an intellectual disability in New South Wales, Australia. *J Appl Res Intellect Disabil.* (2015) 28:383–93. doi: 10.1111/jar.12190
45. Forsgren L, Edvinsson SO, Nystrom L, Blomquist HK. Influence of epilepsy on mortality in mental retardation: an epidemiologic study. *Epilepsia.* (1996) 37:956–63. doi: 10.1111/j.1528-1157.1996.tb00533.x
46. Lauer E, McCallion P. Mortality of people with intellectual and developmental disabilities from select US state disability service systems and medical claims data. *J Appl Res Intellect Disabil.* (2015) 28:394–405. doi: 10.1111/jar.12191
47. McCarron M, Carroll R, Kelly C, McCallion P. Mortality rates in the general irish population compared to those with an intellectual disability from 2003 to 2012. *J Appl Res Intellect Disabil.* (2015) 28:406–13. doi: 10.1111/jar.12194
48. Tyrer F, Smith LK, McGrother CW. Mortality in adults with moderate to profound intellectual disability: a population-based study. *J Intellect Disabil Res.* (2007) 51(Pt. 7):520–7. doi: 10.1111/j.1365-2788.2006.00918.x
49. Shavelle RM, Sweeney LH, Brooks JC. Comparative mortality of persons with intellectual disability in California: an update (2000-2010). *J Insurance Med.* (2014) 44:158–63.
50. Smith GS, Fleming M, Kinnear D, Henderson A, Pell JP, Melville C, et al. Rates and causes of mortality among children and young people with and without intellectual disabilities in Scotland: a record linkage cohort study of 796 190 school children. *BMJ Open.* (2020) 10:e034077. doi: 10.1136/bmjopen-2019-034077
51. Cooper SA, Allan L, Greenlaw N, McSkimming P, Jasilek A, Henderson A, et al. Rates, causes, place and predictors of mortality in adults with intellectual disabilities with and without Down syndrome: cohort study with record linkage. *BMJ Open.* (2020) 10:e036465. doi: 10.1136/bmjopen-2019-036465
52. Blair E, Watson L, Badawi N, Stanley FJ. Life expectancy among people with cerebral palsy in Western Australia. *Dev Med Child Neurol.* (2001) 43:508–15. doi: 10.1017/S0012162201000949
53. Jahan I, Karim T, Das MC, Muhit M, McIntyre S, Smithers-Sheedy H, et al. Mortality in children with cerebral palsy in rural Bangladesh: a population-based surveillance study. *Dev Med Child Neurol.* (2019) 61:1336–43. doi: 10.1111/dmcn.14256
54. Reid SM, Carlin JB, Reddihough DS. Survival of individuals with cerebral palsy born in Victoria, Australia, between 1970 and (2004). *Dev Med Child Neurol.* (2012) 54:353–60. doi: 10.1111/j.1469-8749.2012.04218.x
55. Blohme J, Tornqvist K. Visually impaired Swedish children. The 1980 cohort study–aspects on mortality. *Acta Ophthalmol Scand.* (2000) 78:560–5. doi: 10.1034/j.1600-0420.2000.078005560.x
56. Appleton RE. Mortality in paediatric epilepsy. *Arch Dis Child.* (2003) 88:1091–4. doi: 10.1136/adc.88.12.1091
57. Mbizvo GK, Bennett K, Simpson CR, Duncan SE, Chin RFM. Epilepsy-related and other causes of mortality in people with epilepsy: a systematic review of systematic reviews. *Epilepsy Res.* (2019) 157:106192. doi: 10.1016/j.eplepsyres.2019.106192
58. Moseley BD, Wirrell EC, Wong-Kisiel LC, Nickels K. Early onset epilepsy is associated with increased mortality: a population-based study. *Epilepsy Res.* (2013) 105:410–4. doi: 10.1016/j.eplepsyres.2013.03.002
59. Neligan A, Bell GS, Shorvon SD, Sander JW. Temporal trends in the mortality of people with epilepsy: a review. *Epilepsia.* (2010) 51:2241–6. doi: 10.1111/j.1528-1167.2010.02711.x
60. Berg AT, Nickels K, Wirrell EC, Geerts AT, Callenbach PM, Arts WF, et al. Mortality risks in new-onset childhood epilepsy. *Pediatrics.* (2013) 132:124–31. doi: 10.1542/peds.2012-3998
61. Wilmshurst JM, Kakooza-Mwesige A, Newton CR. The challenges of managing children with epilepsy in Africa. *Semin Pediatr Neurol.* (2014) 21:36–41. doi: 10.1016/j.spen.2014.01.005
62. Tomson T, Walczak T, Sillanpaa M, Sander JW. Sudden unexpected death in epilepsy: a review of incidence and risk factors. *Epilepsia.* (2005) 46(Suppl. 11):54–61. doi: 10.1111/j.1528-1167.2005.00411.x
63. Whitney R, Donner EJ. Risk factors for Sudden Unexpected Death in Epilepsy (SUDEP) and their mitigation. *Curr Treat Opt Neurol.* (2019) 21:7. doi: 10.1007/s11940-019-0547-4
64. Kariuki SM, Kakooza-Mwesige A, Wagner RG, Chengo E, White S, Kamuyu G, et al. Prevalence and factors associated with convulsive status epilepticus in Africans with epilepsy. *Neurology.* (2015) 84:1838–45. doi: 10.1212/WNL.0000000000001542
65. Sadarangani M, Seaton C, Scott JA, Ogutu B, Edwards T, Prins A, et al. Incidence and outcome of convulsive status epilepticus in Kenyan children: a cohort study. *Lancet Neurol.* (2008) 7:145–50. doi: 10.1016/S1474-4422(07)70331-9

66. Kariuki SM, Matuja W, Akpalu A, Kakooza-Mwesige A, Chabi M, Wagner RG, et al. Clinical features, proximate causes, and consequences of active convulsive epilepsy in Africa. *Epilepsia.* (2014) 55:76–85. doi: 10.1111/epi.12392
67. Newton CR, Kariuki SM. Status epilepticus in sub-Saharan Africa: new findings. *Epilepsia.* (2013) 54(Suppl. 6):50–3. doi: 10.1111/epi.12277
68. Maulik PK, Mascarenhas MN, Mathers CD, Dua T, Saxena S. Prevalence of intellectual disability: a meta-analysis of population-based studies. *Res Dev Disabil.* (2011) 32:419–36. doi: 10.1016/j.ridd.2010.12.018
69. Vena JE, Sultz HA, Carlo GL, Fiedler RC, Barnes RE. Sources of bias in retrospective cohort mortality studies: a note on treatment of subjects lost to follow-up. *J Occup Med.* (1987) 29:256–61.
70. Boyle CA, Decoufle P, Holmgreen P. Contribution of developmental disabilities to childhood mortality in the United States: a multiple-cause-of-death analysis. *Paediatr Perinatal Epidemiol.* (1994) 8:411–22. doi: 10.1111/j.1365-3016.1994.tb00480.x
71. Decoufle P, Autry A. Increased mortality in children and adolescents with developmental disabilities. *Paediatr Perinatal Epidemiol.* (2002) 16:375–82. doi: 10.1046/j.1365-3016.2002.00430.x
72. Landes SD. The intellectual disability mortality disadvantage: diminishing with age? *Am J Intellect Dev Disabil.* (2017) 122:192–207. doi: 10.1352/1944-7558-122.2.192
73. Einfeld SL, Ellis LA, Emerson E. Comorbidity of intellectual disability and mental disorder in children and adolescents: a systematic review. *J Intellect Dev Disabil.* (2011) 36:137–43. doi: 10.1080/13668250.2011.572548
74. Kinnear D, Morrison J, Allan L, Henderson A, Smiley E, Cooper SA. Prevalence of physical conditions and multimorbidity in a cohort of adults with intellectual disabilities with and without Down syndrome: cross-sectional study. *BMJ Open.* (2018) 8:e018292. doi: 10.1136/bmjopen-2017-018292
75. Matson JL, Cervantes PE. Comorbidity among persons with intellectual disabilities. *Res Autism Spect Disord.* (2013) 7:1318–22. doi: 10.1016/j.rasd.2013.07.018
76. Ross EJ, Graham DL, Money KM, Stanwood GD. Developmental consequences of fetal exposure to drugs: what we know and what we still must learn. *Neuropsychopharmacology.* (2015) 40:61–87. doi: 10.1038/npp.2014.147
77. Crichton JU, Mackinnon M, White CP. The life-expectancy of persons with cerebral palsy. *Dev Med Child Neurol.* (1995) 37:567–76. doi: 10.1111/j.1469-8749.1995.tb12045.x
78. Hutton JL, Cooke T, Pharoah PO. Life expectancy in children with cerebral palsy. *BMJ.* (1994) 309:431–5. doi: 10.1136/bmj.309.6952.431
79. Strauss DJ, Shavelle RM, Anderson TW. Life expectancy of children with cerebral palsy. *Pediatr Neurol.* (1998) 18:143–9. doi: 10.1016/S0887-8994(97)00172-0
80. Hutton JL, Colver AF, Mackie PC. Effect of severity of disability on survival in north east England cerebral palsy cohort. *Arch Dis Child.* (2000) 83:468–74. doi: 10.1136/adc.83.6.468
81. Hutton JL, Pharoah PO. Effects of cognitive, motor, and sensory disabilities on survival in cerebral palsy. *Arch Dis Child.* (2002) 86:84–9. doi: 10.1136/adc.86.2.84
82. Molnar GE. Rehabilitation in cerebral palsy. *Western J Med.* (1991) 154:569–72.
83. Novak I, McIntyre S, Morgan C, Campbell L, Dark L, Morton N, et al. A systematic review of interventions for children with cerebral palsy: state of the evidence. *Dev Med Child Neurol.* (2013) 55:885–910. doi: 10.1111/dmcn.12246

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurology

ORIGINAL RESEARCH
published: 20 May 2021
doi: 10.3389/fneur.2021.675244

Prevalence, Severity, and Clinical Management of Brain Incidental Findings in Healthy Young Adults: MRi-Share Cross-Sectional Study

Aïcha Soumaré[1], Naka Beguedou[2,3,4], Alexandre Laurent[2,3,4], Bruno Brochet[5], Constance Bordes[1], Sandy Mournet[1], Emmanuel Mellet[2,3,4], Edwige Pereira[6], Clothilde Pollet[6], Morgane Lachaize[1], Marie Mougin[6], Ami Tsuchida[2,3,4], Hugues Loiseau[7], Thomas Tourdias[8,9], Christophe Tzourio[6,10†], Bernard Mazoyer[2,3,4,8†] and Stéphanie Debette[1,5†]*

[1] UMR1219 Bordeaux Population Health Center (Team VINTAGE), INSERM-University of Bordeaux, Bordeaux, France, [2] Neurofunctional Imaging Group, Institute of Neurodegenerative Disease-UMR5293, University of Bordeaux, Bordeaux, France, [3] Neurofunctional Imaging Group, Institute of Neurodegenerative Disease-UMR5293, CNRS, Bordeaux, France, [4] Neurofunctional Imaging Group, Institute of Neurodegenerative Disease-UMR5293, CEA, Bordeaux, France, [5] Department of Neurology, University Hospital Centre Bordeaux Pellegrin Hospital Group, Bordeaux, France, [6] UMR1219 Bordeaux Population Health Center (Team HEALTHY), INSERM-University of Bordeaux, Bordeaux, France, [7] Department of Neurosurgery, University Hospital Centre Bordeaux Pellegrin Hospital Group, Bordeaux, France, [8] Department of Diagnostic and Therapeutic Radiology and Neuroimaging, University Hospital Centre Bordeaux Pellegrin Hospital Group, Bordeaux, France, [9] Magendie Neurocenter INSERM-U1215, University of Bordeaux, Bordeaux, France, [10] Department of Medical Information, University Hospital Centre Bordeaux Pellegrin Hospital Group, Bordeaux, France

OPEN ACCESS

Edited by:
Alessandra Solari,
Fondazione IRCCS Istituto Neurologio Carlo Besta, Italy

Reviewed by:
Jaume Sastre-Garriga,
Centre d'Esclerosi Múltiple de Catalunya (Cemcat), Spain
Paolo Preziosa,
Vita-Salute San Raffaele University, Italy

***Correspondence:**
Stéphanie Debette
stephanie.debette@u-bordeaux.fr

[†] These authors have contributed equally to this work and share senior authorship

Specialty section:
This article was submitted to Neuroepidemiology, a section of the journal Frontiers in Neurology

Received: 02 March 2021
Accepted: 19 April 2021
Published: 20 May 2021

Citation:
Soumaré A, Beguedou N, Laurent A, Brochet B, Bordes C, Mournet S, Mellet E, Pereira E, Pollet C, Lachaize M, Mougin M, Tsuchida A, Loiseau H, Tourdias T, Tzourio C, Mazoyer B and Debette S (2021) Prevalence, Severity, and Clinical Management of Brain Incidental Findings in Healthy Young Adults: MRi-Share Cross-Sectional Study. Front. Neurol. 12:675244. doi: 10.3389/fneur.2021.675244

Background and Objectives: Young adults represent an increasingly large proportion of healthy volunteers in brain imaging research, but descriptions of incidental findings (IFs) in this age group are scarce. We aimed to assess the prevalence and severity of IFs on brain MRIs of healthy young research participants aged 18–35 years, and to describe the protocol implemented to handle them.

Methods: The study population comprised 1,867 participants aged 22.1 ± 2.3 years (72% women) from MRi-Share, the cross-sectional brain MRI substudy of the i-Share student cohort. IFs were flagged during the MRI quality control. We estimated the proportion of participants with IFs [any, requiring medical referral, potentially serious (PSIFs) as defined in the UK biobank]: overall, by type and severity of the final diagnosis, as well as the number of IFs.

Results: 78/1,867 participants had at least one IF [4.2%, 95% Confidence Interval (CI) 3.4–5.2%]. IFs requiring medical referral ($n = 38$) were observed in 36/1,867 participants (1.9%, 1.4–2.7%), and represented 47.5% of the 80 IFs initially flagged. Referred IFs were retrospectively classified as PSIFs in 25/1,867 participants (1.3%, 0.9–2.0%), accounting for 68.4% of anomalies referred (26/38). The most common final diagnosis

was cysts or ventricular abnormalities in all participants (9/1,867; 0.5%, 0.2–0.9%) and in those with referred IFs (9/36; 25.0%, 13.6–41.3%), while it was multiple sclerosis or radiologically isolated syndrome in participants with PSIFs (5/19; 26.3%, 11.5–49.1%) who represented 0.1% (0.0–0.4%) and 0.2% (0.03–0.5%) of all participants, respectively. Final diagnoses were considered serious in 11/1,867 participants (0.6%, 0.3–1.1%). Among participants with referred IFs, 13.9% (5/36) required active intervention, while 50.0% (18/36) were put on clinical surveillance.

Conclusions: In a large brain imaging study of young healthy adults participating in research we observed a non-negligible frequency of IFs. The etiological pattern differed from what has been described in older adults.

Keywords: incidental findings, brain MRI, prevalence, young adults, epidemiology, multiple sclerosis, radiologically isolated syndrome

INTRODUCTION

The widespread use of advanced brain MRI techniques in clinical research entails increasingly frequent detection of incidental findings (IFs). An IF is "a finding concerning an individual research participant that has potential health or reproductive importance and is discovered in the course of conducting research, but is beyond the aims of the study" (1). While most IFs are not clinically meaningful, some can reflect underlying diseases amenable to treatment or can, in rare instances, be life-threatening (e.g., aneurysms, neoplasms). In such cases, early detection and treatment could be of clinical benefit for the participant. Conversely, detecting and disclosing IFs can entail increased clinical workload and costs, cause psychological or financial distress to participants, and expose them to potentially harmful interventions (1–4). However, specific guidelines on the identification and management of brain IFs in a research setting are currently lacking (5–7).

Young adults (18–35 years) participating in biomedical research represent a particularly challenging group regarding IF discovery. First, although individuals in this age group represent a large proportion of "healthy volunteers" in brain imaging research, descriptions of IFs in this age range are scarce. The very few available studies were based on small or selected samples with a broad definition of IFs, low image resolution, and no assessment of their clinical severity (8, 9). Second, characteristics and severity of IFs could differ by age (10, 11), with young asymptomatic adults experiencing greater consequences with longer term effects than older ones. In this context, a systematic description of the prevalence and clinical relevance of IFs in young adults could provide valuable information to notify research participants in this age range of the likelihood and consequences of IFs.

We sought to assess the prevalence of IFs and their severity in 1,867 young adults aged 18–35 years, participating in the MRi-Share brain imaging substudy of the i-Share student cohort, and to describe the standardized protocol implemented to manage them.

Abbreviations: IFs, Incidental Findings; PSIFs, Potentially Serious Incidental Findings; UKB, UK biobank.

MATERIALS AND METHODS

Study Design and Population

The i-Share (internet-based Students' Health Research Enterprise) project[1] is an ongoing population-based cohort study of French-speaking students that was launched in 2013 (i) to assess the frequency and impact of various diseases or conditions affecting young adults, and (ii) to explore the pathophysiology and early mechanisms underlying common chronic disorders, including diseases occurring at a later age. To be eligible, students had to be officially registered at a University or another higher education institution (HEI), be at least 18 years of age, and be able to read and understand French (12). Students were informed about the objectives of the study through promotion campaigns (flyers, information booths on admission days, lectures, social media, and newsletters). Overall, the study was conducted in >420 universities or HEIs (96% in France), the largest recruitment coming from the universities of Bordeaux, Versailles-Saint-Quentin-en-Yvelines, and Nice-Sophia-Antipolis.

The MRi-Share ancillary study is a brain imaging study embedded within the i-Share cohort, which entails a brain MRI and a battery of cognitive tests (13). The objectives of MRi-Share are (i) to characterize morphological and functional variability of the brain in young adults by building a database of morphological and functional MRI images, (ii) to describe the anatomical and functional brain architecture in this population, and (iii) to characterize brain connectivity and its relation with cognitive skills. To participate in MRi-Share, i-Share participants had to be aged between 18 and 35 years, to be registered in a University or HEI in the Bordeaux area, to have completed the i-Share baseline self-administered online questionnaire, and signed an informed consent. Participants with a contraindication to brain MRI (e.g., claustrophobia, pacemaker, and other implanted electronic or metal devices), pregnancy or nursing were not eligible. Between October 2015 and June 2017, i-Share participants were invited to take part in MRi-Share. Those interested were then invited for

[1]i-Share est la plus grande étude scientifique jamais réalisée sur la santé des jeunes [Internet]. Université de Bordeaux. Available online at: https://www.i-share.fr (accessed November 24, 2020).

a visit during which they received detailed information on this ancillary study and were given a virtual tour of the MRI facility. Contraindications for the brain MRI were verified and a referent physician answered any questions they had before signing a written informed consent. MRi-Share participants received a compensation of 40 euros.

MRI Protocol

The MRI acquisition protocol for MRi-Share (13) was designed to emulate that of the UK biobank (UKB) MR imaging study (14) as much as possible, to enable combined analyses of the two databases, since early adulthood is currently not covered by the UKB design. MRIs were performed on a Siemens 3T Prisma scanner (Erlangen, Germany) with a 64-channel head coil (gradients: 80 mT/m−200 T/m/s), between November 2015 and November 2017 at the Bordeaux Institute of BIOimaging. The acquisition lasted about 45 min and included the following five sequences: T1-weighted (T1w) structural imaging (3D MPRAGE, sagittal acquisition, TR/TE/TI = 2,000/2.0/880 ms, repeat × 2, 1 mm^3 isotropic, 192 × 256 × 256); T2-weighted (T2w) FLAIR structural imaging (3D SPACE, sagittal acquisition, TR/TE/TI = 5,000/394/1800 ms, repeat ×2, 1 mm^3 isotropic, 192 × 256 × 256); Diffusion Weighted Imaging [DWI, axial acquisition, echoplanar imaging, TR/TE = 3,540/75.0 ms, multiband ×3, 100 directions, multishell b = 0 s/mm^2 (8 + 8 phase-encoding reversed), b = 300 s/mm^2 (8 directions), b = 1,000 s/mm^2 (32 directions), b = 2,000 s/mm^2 (60 directions), 1.75 mm^3 isotropic, 118 × 118 × 84]; Susceptibility-Weighted Structural Imaging (SWI, axial acquisition, TR/TE1 = 24.0/9.42 ms, 0.8 × 0.8 × 3 mm^3 anisotropic, 252 × 288 × 48); and Resting-state functional MRI (2D $T2^*$-BOLD resting state, axial acquisition, echoplanar imaging, TR/TE = 850/35.0 ms, multiband × 6, 2.4 mm^3 isotropic, 88 × 88 × 66). A detailed summary of acquisition parameters for each modality is presented in **Supplementary Table 1**. Following the MRI scan, each participant had to complete a questionnaire about their thoughts while undergoing the functional MRI and to perform two cognitive tests during 20 min.

Protocol for Assessment and Management of Incidental Findings

Definitions and Assessment of IFs

This study is focused on IFs identified on structural brain MRI exclusively. Within days following the MRI acquisition, T1w and T2w FLAIR images were systematically checked visually for quality by one of two MD investigators trained in brain imaging with >30 years' experience [EM, BM (also professor of neuroradiology at Bordeaux University Hospital)]. If an IF (defined as proposed previously) (1) was detected during this quality control and considered to be potentially harmful for the participant's health, it was shown to a specialized clinical neuroradiologist at Bordeaux University Hospital (TT, professor of neuroradiology with >12 years' experience) who checked the clinical relevance of this IF to decide whether it required medical referral. DWI and/or SWI images were used to better characterize IFs detected on T1w and T2w FLAIR. However, raw DWI and/or SWI images did not undergo visual quality control because those modalities are prone to artifacts induced by eddy currents and/or susceptibility effects; efficient quality control of these acquisitions must be performed after some pre-processing as described previously (13). Of note, the following IFs were not reported: (i) T2-hyperintensities that were isolated or in small numbers (<5), and without any features suggestive of an underlying inflammatory condition such as ovoid shape and periventricular, juxtacortical, or posterior fossa location; (ii) small pineal cysts. For the latter, in the absence of recommendations the threshold was initially set at 10 mm (until May 2016), and subsequently at 15 mm, as the 10 mm threshold generated too many cases and a size >15 mm was described to be potentially associated with neurologic symptoms attributable to mass effect on adjacent structures or hydrocephalus through the compression of the cerebral aqueduct (15). All IFs flagged as requiring referral were reported to a referent neurologist at Bordeaux University Hospital (SD), and categorized as requiring immediate (e.g., acute stroke, encephalitis), urgent (within 1 week, e.g., malignant brain tumor), or routine medical referral. Prior to consenting to participate, MRi-Share participants were informed beforehand that the study might, in rare instances, entail the discovery of an IF, for which they might be contacted if considered to be potentially harmful for their health, with detailed information provided in the setting of a medical visit by a certified neurologist. i-Share participants volunteering for MRi-Share who refused to receive feedback about a potential IF were not eligible for MRi-Share. Participants were also informed that the MRI exam was not a diagnostic test and that some anomalies might not be detected.

Disclosure and Handling of Referred IFs

The referent neurologist called the participant <48 h before the next available neurological outpatient clinic slot, in order to minimize the period of anxiety and stress. No diagnosis was given by phone; participants were informed of the presence of an imaging finding on their MRI scan requiring additional investigation and an appointment in the outpatient clinic was organized. In the outpatient clinic, the neurologist explained the observed abnormality to the participant, collected information about medical history and ongoing treatments, conducted a physical examination, and informed the participant about the proposed follow-up procedures, i.e., additional imaging, blood tests and/or referral to another physician (if needed for diagnosis or management purposes, as described in the letter of information received prior to providing consent). Psychological support was also proposed at the Student Health Service center when needed. Following this interview with the referent neurologist, the participant was systematically invited to undergo a complementary "clinical" brain imaging (MRI or CT-scan, with or without contrast enhancement depending on the nature of the IF), which was interpreted by a clinical neuroradiologist in the context of clinical care. Once the final diagnosis was confirmed, the appropriate care was determined by a specialist on the basis of the type, location, severity, size, and progression of the IF, as well as clinical symptoms and medical history.

Potentially Serious Incidental Findings (PSIFs) and Severity of Final Diagnosis

In order to facilitate the comparison with published studies, we retrospectively classified IFs requiring medical referral as potentially serious incidental findings (PSIFs) using a protocol developed by the UKB that was not available at the time of our study (11, 16, 17). First, IFs were defined as PSIFs if listed as such by the UKB or if meeting their definition of PSIF, i.e., an IF "indicating the possibility of a condition which, if confirmed, would carry a real prospect of seriously threatening lifespan, or of having a substantial effect on major body functions or quality of life" (17). Finally, for participants with PSIFs who were followed-up, final diagnoses were considered as "either: serious (if they were likely to threaten lifespan, or have a substantial impact on quality of life or major body function); not serious (if this was not the case or if the diagnosis was already known); or indeterminate (if there remained insufficient data to classify a final diagnosis as serious or not)" (11). In participants with more than one PSIF, the most serious final clinical diagnosis was accounted for.

Other Measurements

Information about sociodemographic and academic characteristics, health status, personal and family medical history, and lifestyle habits were collected through the i-Share baseline self-administered online questionnaire. The following variables were considered in the analyses to describe the study sample: age at i-Share inclusion, age at MRI, sex, field of study (healthcare/health related disciplines vs. others), self-rated health (defined on a qualitative scale as very good or good vs. fair, bad or very bad), having visited a physician in the past 12 months, regular consumption of medications, history of hospitalization in the past 12 months, familial economic situation during childhood (rated as very comfortable or comfortable vs. fair, difficult or very difficult), current sources of income (familial, scholarship on social grounds, and income-generating activities during the University year), self-reported physician-diagnosed migraine, self-reported physician-diagnosed type 1 diabetes, self-reported physician-diagnosed multiple sclerosis (MS), family history of stroke or cardiovascular disease (myocardial infarction, angina pectoris), family history of cancer, being a current smoker, heavy drinking habits in the past 12 months [defined as frequent episodes of binge drinking (≥6 drinks in about 2 h on the same occasion) 2–6 times per week or every day], use of psychoactive drugs in the past 12 months (cannabis, ecstasy/3,4 methylenedioxymethamphetamine, amphetamines, nitrous oxide, inhalant, or cocaine), and use of other illicit drugs at least once in a lifetime (magic mushrooms or other hallucinogenic plants, crack/free-base, heroin, LSD, or ketamine).

Statistical Analyses

To assess whether MRi-Share participants were representative of i-Share participants at large, we compared characteristics of i-Share participants recruited at universities or other HEIs in the Bordeaux area with respect to their participation in MRi-Share. We first conducted univariate analyses (analysis of covariance for continuous variables, chi-square or Fisher exact test for categorical variables). Second, we performed multivariable logistic regression including all variables associated with MRi-Share participation in univariate analyses with a $p \leq 0.05$ in the model. Third, we compared MRi-Share participants' characteristics according to the presence or absence of IFs. Finally, we extracted data from radiological and medical reports to give descriptive statistics: number and proportion of participants with IFs (any, referred, PSIFs) with their corresponding 95% Confidence Intervals (CI) when appropriate: overall, by type of diagnosis and severity; number of IFs. We also presented the management of IFs. All analyses were performed using SAS software version 9.4 (SAS Institute Inc., Cary, NC, USA), and a two-tailed $p \leq 0.05$ was considered statistically significant.

RESULTS

Of the 14,836 students included in i-Share with a completed baseline questionnaire at the end of the inclusion period for the MRi-Share ancillary study, 8,798 were recruited in universities or other HEIs in Bordeaux or surroundings. Of these, 2,000

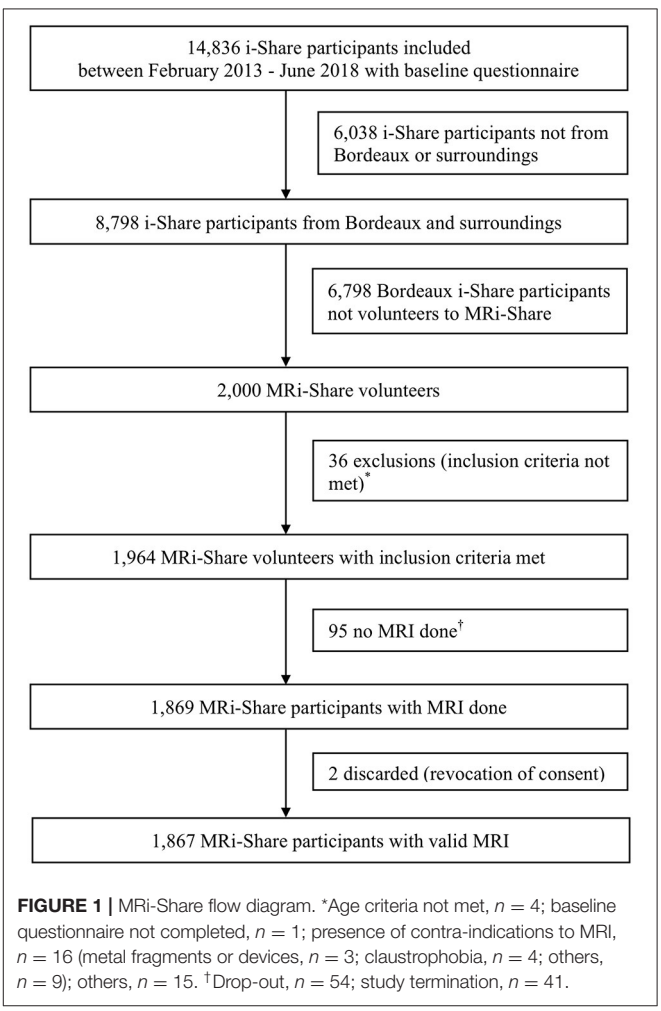

FIGURE 1 | MRi-Share flow diagram. *Age criteria not met, $n = 4$; baseline questionnaire not completed, $n = 1$; presence of contra-indications to MRI, $n = 16$ (metal fragments or devices, $n = 3$; claustrophobia, $n = 4$; others, $n = 9$); others, $n = 15$. †Drop-out, $n = 54$; study termination, $n = 41$.

students responded to the MRi-Share invitation with 1,964 meeting inclusion criteria. After excluding 95 participants for whom the MRI could not be performed (drop out or termination of study participation) and two participants who rescinded their consent, our final MRi-Share study sample comprised 1,867 participants (**Figure 1**). Baseline characteristics of the 1,867 students included in MRi-Share are presented in **Table 1**. Our study population comprised 72% of women with a mean age (SD) at i-Share inclusion and at MRI of 21.2 (2.3) years and 22.1 (2.3) years, respectively. Among i-Share participants, participation in the MRi-Share study was associated with older age, healthcare or health related studies, having had a comfortable or very comfortable familial economic situation in childhood, having an income originating from the family or having a scholarship on social grounds, and use of psychoactive drugs. Migraine and current smoking were less common in MRi-Share participants than in i-Share participants who did not take part in the brain imaging study (multivariable logistic regression analysis, **Table 1**).

IFs were detected in 78 of the 1,867 participants (4.2%, 95% CI: 3.4–5.2%). IFs requiring medical referral ($n = 38$) were found in 36 of those 1,867 participants (1.9%, 1.4–2.7%), two participants having two IFs. Referral was deemed urgent for one participant (0.05%, 0.0–0.3% of the study sample), and routine for the others. Twenty six (68.4%) of the 38 IFs referred were retrospectively characterized as PSIFs [in 25/1,867 individuals

TABLE 1 | Comparison of MRi-Share participants with other participants of i-Share Bordeaux.

	MRi-Share participants with MRI* (*N* = 1,867)	Other Bordeaux i-Share participants (*N* = 6,798)	*P*†	OR (95% CI)‡	*P*‡
Age at MRI, years, mean (SD)	22.1 (2.3)				
Age at i-Share recruitment, years, mean (SD)	21.2 (2.3)	20.7 (2.8)	<0.001	1.07 (1.05–1.09)	<0.001
Male Gender	27.8 (519)	24.2 (1,648)	0.002	1.10 (0.97–1.24)	0.14
Field of study					
Other	46.2 (861)	71.3 (4,845)	<0.001	1.00	<0.001
Healthcare	53.8 (1,002)	28.7 (1,953)		2.68 (2.41–3.00)	
Self-rated health					
Fair, bad or very bad	17.5 (327)	20.2 (1,371)	0.01	1.00	0.47
Very good or good	82.5 (1,539)	79.8 (5,427)		1.05 (0.92–1.21)	
PCP visits 12 months before i-Share inclusion	87.1 (1,626)	86.1 (5,850)	0.23		
Regular medication	23.9 (446)	23.1 (1,567)	0.44		
Hospitalization 12 months before i-Share inclusion	16.7 (312)	17.5 (1,193)	0.40		
Familial economic situation during childhood					
Fair, difficult or very difficult	42.0 (783)	49.6 (3,371)	<0.001	1.00	0.002
Very comfortable or comfortable	58.0 (1,083)	50.4 (3,427)		1.21 (1.07–1.36)	
Source of income					
Family	81.8 (1,528)	79.6 (5,412)	0.03	1.20 (1.03–1.39)	0.02
Scholarship on social grounds	37.9 (707)	42.6 (2,899)	0.001	1.16 (1.02–1.31)	0.02
Activities during University year	41.0 (766)	37.0 (2,517)	0.002	1.11 (0.99–1.24)	0.08
Self-reported physician-diagnosed migraine	18.3 (342)	22.9 (1,560)	<0.001	0.85 (0.75–0.98)	0.02
Self-reported physician-diagnosed type 1 diabetes	0.4 (8)	0.4 (24)	0.63		
Self-reported physician-diagnosed MS	0.1 (1)	0.1 (7)	0.99		
Family history of CVD or stroke	14.5 (248)	13.8 (843)	0.45		
Family history of cancer	11.5 (203)	11.4 (726)	0.87		
Current smoker	29.1 (543)	32.1 (2,181)	0.01	0.81 (0.72–0.93)	0.002
Heavy drinking 12 months before i-Share inclusion	5.1 (92)	4.1 (268)	0.07		
Use of psychoactive drugs 12 months before i-Share inclusion§	47.4 (873)	39.4 (2,623)	<0.001	1.44 (1.27–1.62)	<0.001
Use of other illicit drugs at least once in a lifetime¶	3.9 (73)	4.3 (289)	0.49		

Data were collected through the baseline i-share questionnaire before the MRI was done. Values are percentages (number) unless stated otherwise. PCP, Primary care physician; MS, Multiple Sclerosis; CVD, cardiovascular disease.

Heavy drinking in the past 12 months is defined as frequent episodes of binge drinking (≥six drinks in about 2 h on the same occasion) 2–6 times per week or every day.

**Bordeaux i-Share participants included in MRi-Share ancillary study (fulfilling inclusion criteria, with a signed consent), and a valid MRI.*

† Analysis of covariance for continuous variables; chi-square or Fisher exact test for categorical variables.

‡ Odds ratios (OR) and p-values obtained from Multivariable logistic regression models including all variables significantly associated with the participation in MRi-Share in univariate analyses (with a $p \leq 0.05$ in analyses).

§ Cannabis, ecstasy/3, 4-Methylenedioxymethamphetamine, amphetamines, nitrous oxide, inhalant, or cocaine.

¶ Magic mushrooms or other hallucinogenic plants, crack/free-base, heroin, LSD, or ketamine.

(1.3%, 0.9–2.0%)]. Indeed, some lesions that in our opinion required medical referral were not considered as such in the UKB; moreover, for some lesions that were most likely not serious, we preferred to formally rule out differential diagnoses or rare complications by conducting follow-up investigations (e.g., brain MRI with contrast enhancement, electroencephalography). Participants with IFs (any, requiring medical referral, or PSIFs) did not significantly differ from those without in terms of baseline characteristics (**Table 2**).

The procedure for detection and management of IFs is outlined in **Figure 2**. Of the 36 participants with IFs requiring medical referral, 35 (97.2%) were seen by the referent neurologist, and 33 (91.7%) underwent the recommended clinical brain imaging (MRI or CT). One participant remained unreachable after failing to attend the scheduled appointment with the referent neurologist and another participant refused the recommended clinical brain imaging. In both cases, the primary care physician was informed about the IF, with participants' prior consent. In one participant, the IF was retrospectively found to have already been diagnosed on a previous clinical brain MRI prior to MRi-Share participation. In total, 30 out of the 33 (90.9%) participants with a complementary

TABLE 2 | Baseline characteristics of MRi-Share participants with MRI according to incidental findings (IFs) status*.

	Participants without IFs (*N* = 1,789)	Participants with any IFs (*N* = 78)	Participants with IFs referred (*N* = 36)	Participants with PSIFs† (*N* = 25)	*P*‡	*P*§	*P*¶
Age at MRI, years, mean (SD)	22.1 (2.3)	22.2 (2.5)	22.4 (2.6)	22.5 (2.2)	0.79	0.40	0.33
Age at i-Share recruitment, years, mean (SD)	21.2 (2.3)	21.3 (2.7)	21.5 (2.8)	21.8 (2.3)	0.75	0.43	0.23
Complementary brain imaging, median (min, max)			1.0 (0.0–4.0)	1.0 (0.0–4.0)			
Visits with medical specialists, median (min, max)			1.0 (0.0–6.0)	2.0 (0.0–6.0)			
Male Gender	27.6 (493)	33.3 (26)	22.2 (8)	20.0 (5)	0.27	0.48	0.40
Field of study							
Other	46.3 (827)	43.6 (34)	50.0 (18)	44.0 (11)	0.63	0.66	0.82
Healthcare	53.7 (958)	56.4 (44)	50.0 (18)	56.0 (14)			
Self-rated health							
Fair, bad or very bad	17.5 (313)	17.9 (14)	27.8 (10)	28.0 (7)	0.92	0.11	0.18
Very good or good	82.5 (1,475)	82.1 (64)	72.2 (26)	72.0 (18)			
PCP visits 12 months before i-Share inclusion	87.1 (1,558)	87.2 (68)	88.9 (32)	88.0 (22)	0.99	0.76	0.99
Regular medication	24.1 (431)	19.2 (15)	30.6 (11)	36.0 (9)	0.32	0.37	0.17
Hospitalization 12 months before i-Share inclusion	16.5 (295)	21.8 (17)	27.8 (10)	28.0 (7)	0.22	0.07	0.17
Familial economic situation during childhood							
Fair, difficult or very difficult	41.7 (746)	47.4 (37)	50.0 (18)	48.0 (12)	0.32	0.32	0.53
Very comfortable or comfortable	58.3 (1,042)	52.6 (41)	50.0 (18)	52.0 (13)			
Source of income							
Family	81.8 (1,464)	82.1 (64)	77.8 (28)	76.0 (19)	0.96	0.53	0.44
Scholarship on social ground	37.6 (672)	44.9 (35)	47.2 (17)	56.0 (14)	0.19	0.24	0.06
Activities during University year	41.3 (739)	34.6 (27)	38.9 (14)	36.0 (9)	0.24	0.77	0.59
Self-reported physician-diagnosed migraine	18.2 (325)	21.8 (17)	25.0 (9)	28.0 (7)	0.42	0.29	0.20
Self-reported physician-diagnosed type 1 diabetes	0.4 (7)	1.3 (1)	2.8 (1)	4.0 (1)	0.29	0.15	0.11
Self-reported physician-diagnosed MS	0.1 (1)	0.0 (0)	0.0 (0)	0.0 (0)	0.99	0.99	0.91
Family history of CVD or stroke	14.8 (243)	7.2 (5)	9.7 (3)	15.0 (3)	0.08	0.61	0.99
Family history of cancer	11.3 (191)	16.7 (12)	15.2 (5)	22.7 (5)	0.17	0.42	0.10
Current smoker	29.5 (527)	20.5 (16)	22.2 (8)	16.0 (4)	0.09	0.34	0.14
Heavy drinking 12 months before i-Share inclusion	5.1 (88)	5.6 (4)	9.1 (3)	4.3 (1)	0.78	0.24	0.99
Use of psychoactive drugs 12 months before i-Share inclusion‖	47.7 (841)	41.6 (32)	42.9 (15)	37.5 (9)	0.29	0.57	0.32
Use of other illicit drugs at least once in a lifetime**	4.0 (72)	1.3 (1)	2.8 (1)	4.0 (1)	0.37	0.99	0.99

Data were collected through the baseline i-share questionnaire before the MRI was done. Values are percentages (number) unless stated otherwise. PCP, Primary care physician; MS, Multiple Sclerosis; CVD, cardiovascular disease.

Heavy drinking in the past 12 months is defined as frequent episodes of binge drinking (≥six drinks in about 2 h on the same occasion) 2–6 times per week or every day.

**Bordeaux i-Share participants included in MRi-Share ancillary study (fulfilling inclusion criteria, with a signed consent) and a valid MRI (N = 1,867).*

† PSIFs (Potentially Serious Incidental Findings) are IFs referred that were retrospectively identified according to the list of PSIFs developed by the UK biobank, or its definition of PSIFs. Analysis of covariance for continuous variables or chi-square/Fisher exact test for categorical variables are used to compare: ‡participants without IFs vs. those with IFs; §participants without IFs vs. those with IFs referred; ¶participants without IFs vs. those with PSIFs.

‖Cannabis, ecstasy/3, 4-Methylenedioxymethamphetamine, amphetamines, nitrous oxide, inhalant, or cocaine.

***Magic mushrooms or other hallucinogenic plants, crack/free-base, heroin, LSD, or ketamine.*

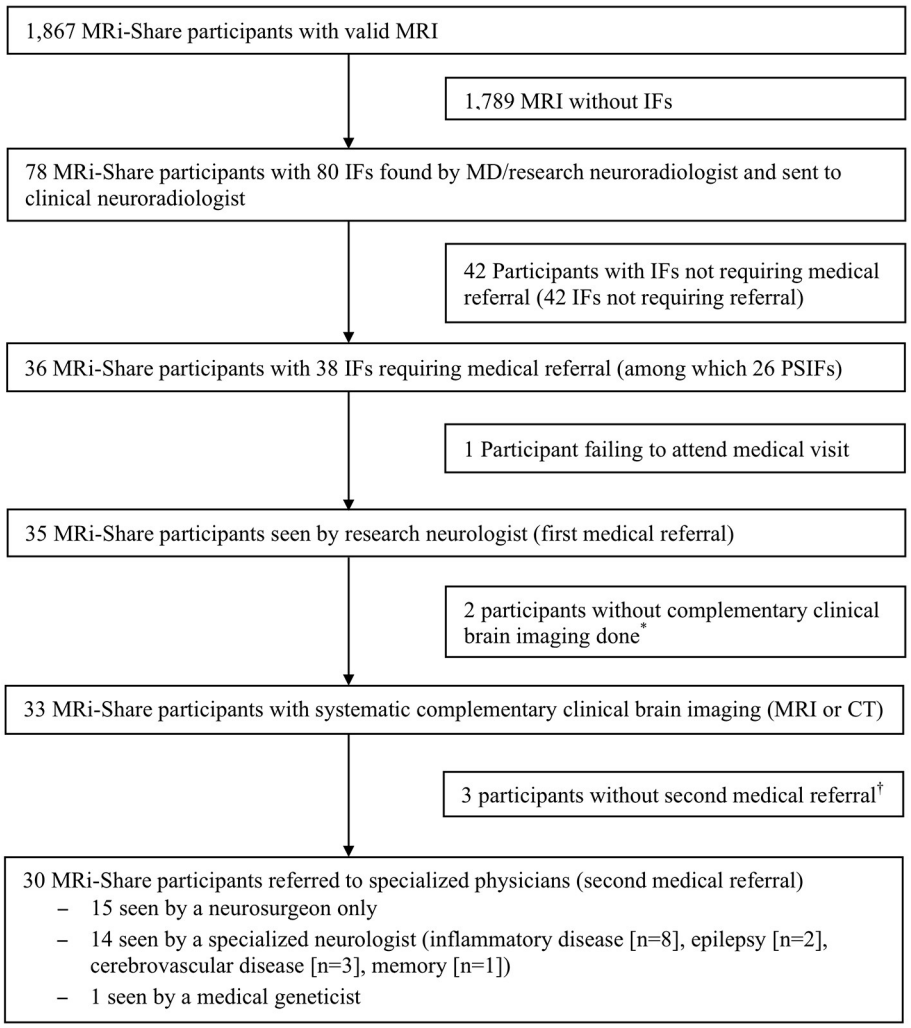

FIGURE 2 | Procedure for detection and management of incidental findings (IFs). MD, medical doctor; PSIFs, potentially serious incidental findings. *Refusal by the participant, $n = 1$; not necessary (previous reassuring MRI performed prior to i-Share research MRI), $n = 1$. †Refusal of follow-up visit, $n = 2$; not necessary, $n = 1$.

brain imaging were seen by at least one additional medical specialist.

Table 3 presents the initial putative diagnosis for the 38 IFs of the 36 participants requiring medical referral. After additional medical referral and ancillary examinations, the suspected diagnosis changed for five of the 36 participants referred (13.8%, **Table 3**). No conclusion could be drawn on the final diagnosis for three participants, i.e., the one who failed to attend the visit with the referent neurologist, the one who refused the complementary clinical imaging, and one who underwent the complementary brain MRI but did not attend additional visits and ancillary examinations requested for an accurate diagnosis. All participants with an IF seen by the referent neurologist underwent a thorough clinical interview to rule out any prior symptoms that could retrospectively be attributed to the IF. Some were prescribed additional ancillary investigations, such as an electroencephalography if the IF was deemed to be a potential source of epilepsy. In participants with T2 abnormalities suggestive of demyelination based on their location, size, and morphology, a brain and spinal cord MRI with gadolinium injection and a lumbar puncture were performed.

Cysts/ventricular abnormalities were the most frequent referred IFs in MRi-Share, detected in nine of 1,867 participants (0.5%, 95% CI: 0.2–0.9% of the study sample), comprising pineal cysts, arachnoid cysts, and hydrocephalus (**Table 3**, **Supplementary Table 2**). Vascular abnormalities, composed primarily of cavernomas, were the second most common referred IFs in the study sample, observed in six participants (0.3%, 0.1–0.7%). They were followed by white matter hyperintensities suggestive of inflammatory disease observed in five participants (0.3%, 0.09–0.7%), of whom two were diagnosed with MS (0.1%, 0.0–0.4%) according to the McDonald criteria (18), and three with Radiologically Isolated Syndrome (RIS) (0.2%, 0.03–0.5%) based on the Okuda and DIS-Barkhof criteria (19, 20). One of the two MS patients had initially been diagnosed with a RIS.

TABLE 3 | Diagnosis' classification of the 38 incidental findings (IFs) of the 36 participants requiring medical referral according to etiology.

Initial diagnosis on research MRI (*N* = 36 participants)		***N***	**Final diagnosis after clinical brain imaging and assessment (*N* = 33 participants)**		***N***
Cysts/ventricular abn. (*n* = 9)	Pineal cyst	5	**Cysts/ventricular abn.** (*n* = 9)	Pineal cyst	5
	Ventriculomegaly with stenosis of aqueduct of Sylvius[‡]	1		**Passive hydrocephalus with stenosis of aqueduct of Sylvius[‡§]**	1
	Arachnoid cyst[†]	3		**Arachnoid cyst[†§]**	2
				Arachnoid cyst[†]	1
Vascular anomaly (*n* = 7)	**Cavernoma[‡]**	4	**Vascular anomaly** (*n* = 6)	**Cavernoma[‡§]**	4
	DVA	2		DVA	2
	dPVS*[†]	1			
Inflammatory WMH (*n* = 12)	**RIS or MS*[‡]**	11	**Inflammatory WMH** (*n* = 5)	**MS[‡§]**	2
	Inflammatory leukoencephalopathy*[‡]	1		**RIS[‡¶]**	3
Tumors (*n* = 3)	**Ependymoma[†]**	1	**Tumors** (*n* = 4)	**Ependymoma[†§]**	1
	MVNT[†]	2		**MVNT[†¶]**	2
				Ganglioglioma[†§]	1
Cortical malformations (*n* = 3)	Neuronal migration disorder	2	**Cortical malformations** (*n* = 3)	Neuronal migration disorder	2
	Focal cortical dysplasia[‡]	1		**Focal cortical dysplasia[‡¶]**	1
Other, neurological (*n* = 1)	**Basal ganglia calcifications[‡]**	1	**Other, neurological** (*n* = 5)	**Fahr's syndrome[‡¶]**	1
				WMH without underlying inflammatory disease	3
				Undetermined leukoencephalopathy	1
Other, non-neurological (*n* = 3)	Hypertrophy lymphoid tissue in cavum	1	Other, non-neurological (*n* = 3)	Hypertrophy lymphoid tissue in cavum	1
	Bone lesion	1		Benign bone lesion	1
	Lesion or retention in sinus	1		Cyst in sinus	1

abn., abnormalities; DVA, Developmental Venous Anomaly; MVNT, Multinodular and Vacuolating Neuronal Tumor; dPVS, dilated Virchow Robin Space; WMH, White Matter Hyperintensities; RIS, Radiologically Isolated Syndrome; MS, Multiple Sclerosis. Thirty-four participants referred have one IF; two participants have two IFs.

Initial diagnosis could not be confirmed for three inflammatory WMH suggestive of RIS or MS.

**Initial diagnosis modified after additional medical referral and ancillary examinations: one dPVS (vascular anomaly) turned out to be a ganglioglioma (tumor), three inflammatory WMH suggestive of RIS or MS turned out to be WMH without underlying inflammatory disease, one inflammatory WMH suggestive of inflammatory leukoencephalopathy turned out to be an undetermined leukoencephalopathy.*

In bold, Potentially Serious Incidental Findings (PSIFs) are IFs referred that were retrospectively identified according to: [†]the list of PSIFs developed by the UK biobank, or [‡]the UK biobank definition of PSIFs.

Final diagnosis classified as [§]serious, or [¶]indeterminate, according to the UK biobank definition of diagnosis severity in participants with PSIFs followed-up.

Other IFs included cortical malformations, non-neurological IFs, and Fahr's syndrome (**Table 3**, **Supplementary Table 2**).

Among participants with PSIFs according to the UKB list or definition and follow-up available for a final diagnosis ($n = 19$), RIS or MS were the most common IFs (26.3%, **Supplementary Table 2**). Serious diagnoses, as defined in the UKB (11), occurred in 11/1,867 participants (0.6%, 0.3–1.1% of total study sample), representing 57.9% (36.2–76.9%) of participants with PSIFs and a final diagnosis (11/19) (**Table 3**). Conversely, non-serious and indeterminate diagnoses occurred in one (0.05%, 0.0–0.3% of the study sample) and seven (0.4%, 0.2–0.8% of the study sample) participants, respectively.

Regarding the management of identified IFs, active intervention was required for five participants with referred IFs (also defined as PSIFs) (0.3%, 0.09–0.7% of the study sample; 13.9%, 5.6–29.1% of participants with referred IFs), and comprised surgery, medical treatment, or both. Clinical surveillance with or without follow-up brain imaging was prescribed for 18 participants with referred IFs (1.0%, 0.6–1.5% of the study sample; 50.0%, 34.5–65.5% of participants with referred IFs).

DISCUSSION

In 1,867 young students (aged 18–35 years) who underwent 3T brain MRI as part of their participation in the MRi-Share research project, IFs were detected overall in 4.2% (3.4–5.2%) of participants, and IFs requiring medical referral in 1.9% (1.4–2.7%) of participants. The frequency of PSIFs according to the UKB list or definition was 1.3% (0.9–2.0%), while final diagnoses were considered serious in 0.6% (0.3–1.1%) of the participants. The leading final diagnosis was cysts or ventricular abnormalities in participants with referred IFs (25.0%), while it was MS or RIS among those with PSIFs followed-up (26.3%). In this young student population, the prevalence of MS and RIS was, respectively, 0.1% (0.0–0.4%) and 0.2% (0.03–0.5%).

Comparison With Other Studies and Implications of Findings

To our knowledge, only two studies have focused on brain IFs specifically in healthy young adults (8, 9), describing a prevalence of IFs between 5.8 and 9.4%. In one study (N = 2,536 men, mean age: 20.5 years) (8), IFs were defined as abnormal findings based on the sole judgment of one general radiologist. In the other study (N = 203; mean age: 21.9 years) (9), scans were read by a neuroradiologist, but IFs could also broadly include variations of the norm. A subsequent meta-analysis suggested that the rate of PSIFs in these two studies was in fact lower than in MRi-Share, between 0.5 and 0.7% (17) vs. 1.3%. Potential explanations for these differences with MRi-Share findings are several fold, including differences in sample characteristics, MRI methodology (1–1.5T vs. 3T MRI and advanced sequences such as SWI in MRi-Share), and IF definitions. Differences in detection protocols [single reader—in one case non-specialty (8) vs. two consecutive subspecialty readers with >12–30 years' experience in brain imaging in MRi-Share] can also be highlighted. Indeed, lesion detection in neuroradiology settings relies on the level of expertise and experience of the reader, with non-specialty readers (non-neuroradiologists) performing at lower accuracy compared with subspecialty ones (neuroradiologists) (21); moreover, blind double interpretation is supposed to reduce diagnostic errors in radiology particularly with the added value of a specialist neuroradiology second opinion vs. a general radiologist (22). Recently, the UKB assessed the prevalence, type, and final clinical diagnosis of PSIFs in 7,334 middle-aged and older research participants (40–69 years; median age: 63 years) (11). MRI machines (Siemens Prisma) and acquisition parameters were the same as in MRi-Share by design (13). In the UKB, brain PSIFs were detected in 58 participants (0.8%, 0.6–1.0%) using two protocols: a systematic review by a radiologist for the first 1,000 scans (protocol 1, 2.3%), and a radiographer flagging for confirmation by a radiologist for the subsequent 6,334 scans (protocol 2, 0.6%). Serious diagnoses occurred in 0.2% (0.1–0.4%) of the sample (protocol 1, 0.4%; protocol 2, 0.2%) representing 29.3% of those with PSIFs. The slightly higher prevalence of PSIFs and serious diagnoses in MRi-Share could be due to differences in the age of participants (although more IFs would be expected with increasing age) (23), selection bias, and systematic reading by a neuroradiologist or MD investigators highly trained in brain imaging and subsequently by a highly trained clinical neuroradiologist in MRi-Share. Finally, a systematic review and meta-analysis of 16 neuroimaging studies (N = 19,559 individuals) (24) and an umbrella review of two systematic reviews (N = 27,316 individuals) (25) reported IF discovery rates of 2.7 and 22%, respectively. However, these two studies did not provide any additional results for young adults beyond the aforementioned studies (8, 9).

While cysts and vascular anomalies were the most common brain IFs overall in MRi-Share (a quarter of referred IFs), consistent with prior research in young and pediatric populations (8, 9, 26), PSIFs and serious IFs were both dominated by MS or RIS (nearly a quarter of cases). MS and RIS were not described in middle-aged and older adults from UKB, where tumors were the dominating PSIFs, mostly of a different type (meningioma, pituitary tumor, vestibular schwannoma) than the few tumors seen in MRi-Share (ependymoma, multinodular and vacuolating neuronal tumor). A similar variation of tumor histological subtypes by age was recently reported by the Central Nervous System tumor registry of the Bordeaux (Gironde) region in France (27).

The relatively high frequency of MS and RIS in this sample of healthy young adults (0.3% in total, of which 0.2% for RIS) has important implications. MS is a potentially disabling neurological disease with a considerable impact on quality of life (28). There is converging evidence that patients with MS and early initiation of disease modifying therapy (DMT) have a more favorable outcome with a lower frequency of clinical attacks. RIS is a syndrome described for the first time in 2009 and defined by incident MRI findings typical of MS in persons without a clinical history of neurological symptoms suggestive of central nervous system demyelination (20). Despite growing research and clinical interest in RIS, its epidemiology remains unclear. Data on its diagnosis in various settings and populations, its natural course, and predictors are sorely needed. Over half of RIS patients were recently shown to develop MS over 10 years of follow-up in the largest international series (29). Whether to treat persons with RIS using MS DMT is debated and currently assessed in clinical trials (NCT02739542, NCT03122652), but clinical follow-up is strongly recommended to initiate treatment early if clinical symptoms arise. In terms of frequency, a systematic review based on autopsy studies and clinical registries of MRI data found a cumulative incidence of RIS of ∼0.1% (30). Hospital-based studies estimated a prevalence of 0.05% in the broad age group of 0–90 years (0.15–0.7% in 15–40 years) (31, 32). Another study that collected data from all imaging centers in a region of Sweden, over a year, thus reflecting a population-based catchment area, reported a prevalence of 0.1% among 1,907 individuals aged 0–91 years (33). Because it has been suggested that estimated prevalence rates of an abnormality within the general population cannot serve as a meaningful standard against which to interpret rates of corresponding serious IF in a given sample (4), the need for data in young healthy adults in a research setting is crucial.

Strengths and Limitations

Strengths of our study include: the large sample size in an understudied age group; high resolution 3T brain MRI; systematic screening of all images by a professor of neuroradiology or MD highly trained in brain imaging studies followed by a second review by a highly trained clinical neuroradiologist (in case of IF discovery), all blinded to the participants' clinical status, in order to have the best accuracy in terms of lesion detection while reducing diagnostic errors; extensive follow-up investigations enabling more accurate characterization of IFs. The management procedure of IFs was optimized through consultations with the ethics advisory board of the i-Share study and additionally reviewed by an independent ethics adviser.

We acknowledge limitations. First, our study sample is not representative of all young adults aged 18–35 years in

the Bordeaux area. Only 45% of young adults pursue higher education. Moreover, i-Share participants are not representative of all students, as is the case in all prospective population-based studies, regardless of the sampling method used, with a broader participation of women and most likely a selection bias toward students with an interest in health research. Finally, MRi-Share participants reported significantly better socioeconomic conditions than other i-Share participants in the Bordeaux area, were more often students in the health sector, and older than i-Share participants not taking part in MRi-Share. A selection bias of MRi-Share participants toward students with a history of neurological symptoms cannot be excluded, although compared to other i-Share participants they tended to more often self-rate their health as good or very good and less frequently present a history of migraine. Second, we did not evaluate non-medical consequences of reporting IFs to study participants. However, we carefully designed the IF disclosure process to minimize anxiety, offered psychological support to participants when needed, and made arrangements with medical specialists to reduce the waiting time for follow-up medical visits. The financial impact of IF disclosure was minimized by the fact that the French public national health insurance system carries the main financial burden of medical follow-up. Furthermore, to make up for costs not covered by the aforementioned system, students could access complementary private health insurances covered either by their parents' health insurance plans or through the purchase of their own insurance plan at reduced cost. Third, due to its focus on young adults, MRi-Share lacks participants of middle and older ages; this prevented us from exploring whether the prevalence of IFs differed across the adult lifespan through formal statistical comparisons. Fourth, the protocol we implemented to detect and manage brain IFs should be considered within the specific research context of MRi-Share and might not be well-suited in other settings, e.g., in countries where private healthcare predominates, or for studies based on much larger samples given the resources this would require. Guidelines on the management of IFs are thoroughly needed. The experience described here, complementing prior studies, will be informative for scientific societies or expert groups devising such guidelines in the future.

What This Study Adds and Future Directions

To our knowledge this is the largest study on IFs in young adults in a research setting and the first that used a standardized protocol, optimized through consultations with ethics advisors, with two independent radiological readings of IFs. It is also the first to report on the management and severity of IFs in young adults. Moreover, although our results remain primarily descriptive, as in most of the literature on incidental findings, the fact that we used the same type of MRI scanner, the same image acquisition protocol, and the same IF definitions as in the UK biobank dataset, allows a qualitative comparison of findings.

Our results provide novel insight into the frequency and severity of precisely defined IFs in young adults, and also shed new light on the nature of these IFs, which appears to differ notably from that in older adults. We found that, with MS, RIS was the most common PSIF observed in young healthy research participants, thereby highlighting the importance of ongoing therapeutic trials on the management of RIS. We also observed that incidentally discovered brain tumors in young adults participating in research appeared to differ histologically from those identified in older adults, although no formal statistical comparisons could be performed.

In the future, our study could be complemented by a more extensive exploration of risk factors associated with IFs and their severity in young adults, requiring much larger samples, and by a formal assessment of differences in prevalence and etiology between age groups. Guidelines on the detection and management of IFs would be highly valuable in order to optimize the way IFs are handled and also reduce inconsistencies between studies reporting them. In analogy with recommendations on the management of incidental genetic findings emerging from sequencing studies in research (34), scientific societies or expert groups could propose a list of actionable brain MRI IFs requiring medical referral, ideally with specific recommendations by age group, considering the different patterns observed across the adult lifespan and age-specific clinical implications.

CONCLUSION

Our study provides some guidance on the expected frequency and severity of IFs on brain MRI in young healthy adults participating in research, an understudied group, and shows that the etiological pattern of these IFs is distinct from patterns described in older adults. White matter lesions revealing MS or RIS were the most common potentially serious IFs detected in our study. Altogether, our data may inform IF detection and management protocols in future research studies involving brain MRIs of young adults. Given the growing frequency of brain imaging research, with increasingly large samples and high resolution, our findings also highlight the need for expert guidelines on brain MRI IFs management.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: legal and ethical restrictions prohibit public sharing of the individual raw data related to IFs, because such data contain potentially identifiable and sensitive/confidential information on participants. Requests to access these datasets should be directed to Stéphanie Debette, stephanie.debette@u-bordeaux.fr.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Commission Nationale de l'Informatique et des Libertés (CNIL) (DR-2013-019) (for the i-Share project), the regional Ethics Committee (Comité de Protection des Personnes Sud-Ouest et Outre-Mer III) on 24 August 2015 (CPP2015-A00850-49) (for the MRi-Share protocol). IFs' assessment and management protocol was discussed in detail with the

independent ethics advisory board of the i-Share study. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

SD, BM, and CT joint supervisors. CT conceived and designed the i-Share study (principal investigator), and obtained funding for the i-Share cohort. BM conceived and designed the MRi-Share study (principal investigator). SD was involved in the design and conception of the i-share study (co-investigator) and obtained funding for analyst's time. AS and SD conceptualized and developed the question of the study's described in this manuscript. AS cleaned and prepared the IFs dataset. AS and CB cleaned and prepared the portion of the i-Share dataset used in this work, analyzed and conducted statistical analyses. AS, SM, and SD interpreted the study's results. AS drafted the manuscript (with tables/figures). NB, EM, and BM were involved in MRi-Share data acquisition. AL, AT, EM, and BM contributed to MRi-Share imaging data processing (software development for AL and AT). EM, TT, and BM were responsible for MRi-Share imaging data interpretation including flagging and/or confirmation of IFs. AL, AT, and NB contributed to MRi-Share data management. BB, HL, and SD were involved in clinical data acquisition. EP contributed to i-Share planning and data management. CP and MM contributed to i-Share planning and recruitment. ML contributed to data monitoring. All authors critically revised the manuscript and approved the submitted version.

FUNDING

The i-Share study has received funding by the French National Research Agency (Agence Nationale de la Recherche, ANR) *via* the Investissements d'Avenir (Investment for the Future) program (Grand Number ANR-10-COHO-05) and *via* the Initiative d'Excellence (IdEX) de l'Université de Bordeaux program (Grant Number ANR No.-10-IDEX-03-02). Supplementary funding was received from the Conseil Régional of Nouvelle-Aquitaine, reference 4370420. The MRi-Share cohort has been supported by the French National Research Agency (Agence Nationale de la Recherche, ANR)-LabEx TRAIL (Grant Number ANR-10-LABX-57) and by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement N° 640643). SD is also supported by a grant overseen by the French National Research Agency (ANR) as part of the the Investissements d'Avenir (Investment for the Future) Program (Grant Number ANR-18-RHUS-0002, the EU Joint Programme - Neurodegenerative Disease Research (JPND) and funding from the European Union's Horizon 2020 research and innovation program under Grant Agreements Nos. 643417, 640643, 667375, and 754517. The funders had no role in the study design, in the collection, analysis, interpretation of data, writing of the report, or in the decision to submit the article for publication. All authors are independent from the funders.

ACKNOWLEDGMENTS

We thank members of the i-Share ethics council (Dr. Annick Alpérovitch, Prof. Ali Benmakhlouf, Prof. Diane Purper-Ouakil, Dr. Pierre Chauvin, and Dr. France Filiatrault) and Prof. Yann Joly for their advice. The authors are indebted to the following individuals for their invaluable contribution to the MRi-Share project: Nathalie Tzourio-Mazoyer, Sabrina Schilling, Serge Anandra, Amandine André, Christophe Bernard, Aurore Capelli, Claire Cardona, Arnaud Chaussé, Christophe Delalande, Vincent Durand, Louise Knafo, Maylis Melin, Caroline Roussillon, Nicolas Vinuesa, and the i-Share relay students. The authors also thank Dr. Cécile Marchal, Dr. Guillaume Penchet, Dr. Cécile Dulau, Prof. Igor Sibon, Dr. Sabrina Debruxelles, Dr. Sophie Auriacombe for their contribution to the clinical assessment of incidental findings.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2021.675244/full#supplementary-material

REFERENCES

1. Wolf SM, Lawrenz FP, Nelson CA, Kahn JP, Cho MK, Clayton EW, et al. Managing incidental findings in human subjects research: analysis and recommendations. *J Law Med Ethics.* (2008) 36:219–48. doi: 10.1111/j.1748-720X.2008.00266.x
2. Grossman RI, Bernat JL. Incidental research imaging findings: Pandora's costly box. *Neurology.* (2004) 62:849–50. doi: 10.1212/01.WNL.0000118214.02495.41
3. Jagadeesh H, Bernstein M. Patients' anxiety around incidental brain tumors: a qualitative study. *Acta Neurochir.* (2014) 156:375–81. doi: 10.1007/s00701-013-1935-2
4. Royal JM, Peterson BS. The risks and benefits of searching for incidental findings in MRI research scans. *J Law Med Ethics.* (2008) 36:305–14. doi: 10.1111/j.1748-720X.2008.00274.x
5. Borgelt E, Anderson JA, Illes J. Managing incidental findings: lessons from neuroimaging. *Am J Bioeth.* (2013) 13:46–7. doi: 10.1080/15265161.2012.754069
6. Wardlaw JM, Davies H, Booth TC, Laurie G, Compston A, Freeman C, et al. Acting on incidental findings in research imaging. *BMJ.* (2015) 351:h5190. doi: 10.1136/bmj.h5190
7. Bunnik EM, van Bodegom L, Pinxten W, de Beaufort ID, Vernooij MW. Ethical framework for the detection, management and communication of incidental findings in imaging studies, building on an interview study of researchers' practices and perspectives. *BMC Med Ethics.* (2017) 18:10. doi: 10.1186/s12910-017-0168-y
8. Weber F, Knopf H. Incidental findings in magnetic resonance imaging of the brains of healthy young men. *J Neurol Sci.* (2006) 240:81–4. doi: 10.1016/j.jns.2005.09.008
9. Reneman L, de Win MM, Booij J, van den Brink W, den Heeten GJ, Freling N, et al. Incidental head and neck findings on MRI in young healthy volunteers:

prevalence and clinical implications. *AJNR Am J Neuroradiol.* (2012) 33:1971–4. doi: 10.3174/ajnr.A3217
10. Illes J, Rosen AC, Huang L, Goldstein RA, Raffin TA, Swan G, et al. Ethical consideration of incidental findings on adult brain MRI in research. *Neurology.* (2004) 62:888–90. doi: 10.1212/01.WNL.0000118531.90418.89
11. Gibson LM, Nolan J, Littlejohns TJ, Mathieu E, Garratt S, Doherty N, et al. Factors associated with potentially serious incidental findings and with serious final diagnoses on multi-modal imaging in the UK Biobank Imaging Study: a prospective cohort study. *PLoS ONE.* (2019) 14:e0218267. doi: 10.1371/journal.pone.0218267
12. Montagni I, Guichard E, Kurth T. Association of screen time with self-perceived attention problems and hyperactivity levels in French students: a cross-sectional study. *BMJ Open.* (2016) 6:e009089. doi: 10.1136/bmjopen-2015-009089
13. Tsuchida A, Laurent A, Crivello F, Petit L, Joliot M, Pepe A, et al. The MRi-Share database: brain imaging in a cross-sectional cohort of 1,870 University students. *bioRxiv [Preprint].* (2020). doi: 10.1101/2020.06.17.154666
14. Alfaro-Almagro F, Jenkinson M, Bangerter NK, Andersson JLR, Griffanti L, Douaud G, et al. Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. *Neuroimage.* (2018) 166:400–24. doi: 10.1016/j.neuroimage.2017.10.034
15. Gokce E, Beyhan M. Evaluation of pineal cysts with magnetic resonance imaging. *World J Radiol.* (2018) 10:65–77. doi: 10.4329/wjr.v10.i7.65
16. Gibson LM, Littlejohns TJ, Adamska L, Garratt S, Doherty N, Wardlaw JM, et al. Impact of detecting potentially serious incidental findings during multi-modal imaging. *Wellcome Open Res.* (2017) 2:114. doi: 10.12688/wellcomeopenres.13181.1
17. Gibson LM, Paul L, Chappell FM, Macleod M, Whiteley WN, Al-Shahi Salman R, et al. Potentially serious incidental findings on brain and body magnetic resonance imaging of apparently asymptomatic adults: systematic review and meta-analysis. *BMJ.* (2018) 363:k4577. doi: 10.1136/bmj.k4577
18. Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann Neurol.* (2011) 69:292–302. doi: 10.1002/ana.22366
19. Barkhof F, Filippi M, Miller DH, Scheltens P, Campi A, Polman CH, et al. Comparison of MRI criteria at first presentation to predict conversion to clinically definite multiple sclerosis. *Brain.* (1997) 120 (Pt 11):2059–69. doi: 10.1093/brain/120.11.2059
20. Okuda DT, Mowry EM, Beheshtian A, Waubant E, Baranzini SE, Goodin DS, et al. Incidental MRI anomalies suggestive of multiple sclerosis: the radiologically isolated syndrome. *Neurology.* (2009) 72:800–5. doi: 10.1212/01.wnl.0000335764.14513.1a
21. Wang W, van Heerden J, Tacey MA, Gaillard F. Neuroradiologists compared with non-neuroradiologists in the detection of new multiple sclerosis plaques. *AJNR Am J Neuroradiol.* (2017) 38:1323–7. doi: 10.3174/ajnr.A5185
22. Briggs GM, Flynn PA, Worthington M, Rennie I, McKinstry CS. The role of specialist neuroradiology second opinion reporting: is there added value? *Clin Radiol.* (2008) 63:791–5. doi: 10.1016/j.crad.2007.12.002
23. Vernooij MW, Ikram MA, Tanghe HL, Vincent AJ, Hofman A, Krestin GP, et al. Incidental findings on brain MRI in the general population. *N Engl J Med.* (2007) 357:1821–8. doi: 10.1056/NEJMoa070972
24. Morris Z, Whiteley WN, Longstreth WT Jr, Weber F, Lee YC, Tsushima Y, et al. Incidental findings on brain magnetic resonance imaging: systematic review and meta-analysis. *BMJ.* (2009) 339:b3016. doi: 10.1136/bmj.b3016
25. O'Sullivan JW, Muntinga T, Grigg S, Ioannidis JPA. Prevalence and outcomes of incidental imaging findings: umbrella review. *BMJ.* (2018) 361:k2387. doi: 10.1136/bmj.k2387
26. Dangouloff-Ros V, Roux CJ, Boulouis G, Levy R, Nicolas N, Lozach C, et al. Incidental brain MRI findings in children: a systematic review and meta-analysis. *AJNR Am J Neuroradiol.* (2019) 40:1818–23. doi: 10.3174/ajnr.A6281
27. Pouchieu C, Gruber A, Berteaud E, Ménégon P, Monteil P, Huchet A, et al. Increasing incidence of central nervous system (CNS) tumors (2000-2012): findings from a population based registry in Gironde (France). *BMC Cancer.* (2018) 18:653. doi: 10.1186/s12885-018-4545-9
28. Kobelt G, Thompson A, Berg J, Gannedahl M, Eriksson J. New insights into the burden and costs of multiple sclerosis in Europe. *Mult Scler.* (2017) 23:1123–36. doi: 10.1177/1352458517694432
29. Lebrun-Frenay C, Kantarci O, Siva A, Sormani MP, Pelletier D, Okuda DT. Radiologically isolated syndrome: 10-year risk estimate of a clinical event. *Ann Neurol.* (2020) 88:407–17. doi: 10.1002/ana.25799
30. Granberg T, Martola J, Kristoffersen-Wiberg M, Aspelin P, Fredrikson S. Radiologically isolated syndrome–incidental magnetic resonance imaging findings suggestive of multiple sclerosis, a systematic review. *Mult Scler.* (2013) 19:271–80. doi: 10.1177/1352458512451943
31. Wasay M, Rizvi F, Azeemuddin M, Yousuf A, Fredrikson S. Incidental MRI lesions suggestive of multiple sclerosis in asymptomatic patients in Karachi, Pakistan. *J Neurol Neurosurg Psychiatry.* (2011) 82:83–5. doi: 10.1136/jnnp.2009.180000
32. Granberg T, Martola J, Aspelin P, Kristoffersen-Wiberg M, Fredrikson S. Radiologically isolated syndrome: an uncommon finding at a University clinic in a high-prevalence region for multiple sclerosis. *BMJ Open.* (2013) 3:e003531. doi: 10.1136/bmjopen-2013-003531
33. Forslin Y, Granberg T, Jumah AA, Shams S, Aspelin P, Kristoffersen-Wiberg M, et al. Incidence of radiologically isolated syndrome: a population-based study. *AJNR Am J Neuroradiol.* (2016) 37:1017–22. doi: 10.3174/ajnr.A4660
34. Kalia SS, Adelman K, Bale SJ, Chung WK, Eng C, Evans JP, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. *Genet Med.* (2017) 19:249–55. doi: 10.1038/gim.2016.190

Conflict of Interest: SD is currently a guest editor for the Research Topic this work will be submitted to, and has collaborated/is collaborating on research projects or publications with the other two editors of this topic. Outside the submitted work, BB served on the scientific advisory board of Biogen, BMS, Novartis, Roche, Merck, and Jansen, received speaker honoraria from Biogen, BMS, and Roche, served as an editorial board member of Multiple Sclerosis International (2019–2020), received research support from Roche, Biogen, Bayer, Sanofi, and Caridian.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Systems Neuroscience

REVIEW
published: 06 July 2021
doi: 10.3389/fnsys.2021.670404

Prediction Along a Developmental Perspective in Psychiatry: How Far Might We Go?

Frauke Nees[1,2], Lorenz Deserno[3,4,5], Nathalie E. Holz[2], Marcel Romanos[3] and Tobias Banaschewski[2]*

[1] Institute of Medical Psychology and Medical Sociology, University Medical Center Schleswig Holstein, University of Kiel, Kiel, Germany, [2] Department of Child and Adolescent Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany, [3] Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital Würzburg, Würzburg, Germany, [4] Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, [5] Department of Psychiatry and Psychotherapy, Technische Universität Dresden, Dresden, Germany

Most mental disorders originate in childhood, and once symptoms present, a variety of psychosocial and cognitive maladjustments may arise. Although early childhood problems are generally associated with later mental health impairments and psychopathology, pluripotent transdiagnostic trajectories may manifest. Possible predictors range from behavioral and neurobiological mechanisms, genetic predispositions, environmental and social factors, and psychopathological comorbidity. They may manifest in altered neurodevelopmental trajectories and need to be validated capitalizing on large-scale multi-modal epidemiological longitudinal cohorts. Moreover, clinical and etiological variability between patients with the same disorders represents a major obstacle to develop effective treatments. Hence, in order to achieve stratification of patient samples opening the avenue of adapting and optimizing treatment for the individual, there is a need to integrate data from multi-dimensionally phenotyped clinical cohorts and cross-validate them with epidemiological cohort data. In the present review, we discuss these aspects in the context of externalizing and internalizing disorders summarizing the current state of knowledge, obstacles, and pitfalls. Although a large number of studies have already increased our understanding on neuropsychobiological mechanisms of mental disorders, it became also clear that this knowledge might only be the tip of the Eisberg and that a large proportion still remains unknown. We discuss prediction strategies and how the integration of different factors and methods may provide useful contributions to research and at the same time may inform prevention and intervention.

Keywords: life span, prediction, modeling, developmental psychiatry, neurobiology, biomarker

OPEN ACCESS

Edited by:
Sudha Seshadri, Boston University, United States

Reviewed by:
Giovanni Salum, Federal University of Rio Grande do Sul, Brazil
Hirofumi Morishita, Icahn School of Medicine at Mount Sinai, United States

****Correspondence:***
Frauke Nees
nees@med-psych.uni-kiel.de

Received: *21 February 2021*
Accepted: *15 June 2021*
Published: *06 July 2021*

Citation:
Nees F, Deserno L, Holz NE, Romanos M and Banaschewski T (2021) Prediction Along a Developmental Perspective in Psychiatry: How Far Might We Go? Front. Syst. Neurosci. 15:670404. doi: 10.3389/fnsys.2021.670404

INTRODUCTION

Mental disorders are the leading cause of years lived with disability (Whiteford et al., 2013), with most of them having their origin early in life during childhood and adolescence (e.g., Kessler et al., 2005, 2007; Thapar and Riglin, 2020). Despite this clear figure of disability and burden of mental disorders along with staggering economic cost (Gustavsson et al., 2011), research in developmental

psychiatry and into the neurobiology of mental disorders has so far not yet led to a prolific integration of neuro-psycho-bio-social insight into the diagnostic conceptualizations, treatment, and prevention of mental disorders. We need to enhance our understanding of the underlying disease mechanisms, and to validate clinical, (neuro-)biological, and multivariate predictors of psychopathology and therapy outcome. This also includes the specific integration of individual needs and social contexts and translates treatments beyond clinical settings to individuals' daily life.

In the present review, we will describe challenges and unmet needs in the field of developmental psychiatry commenting on (1) the current status on markers and mechanism of psychopathological pathways, (2) the lack of biological validity of diagnostic categories, and proposed approaches of symptom dimensions, (3) clinical heterogeneity, (4) the role of environmental influences and psychosocial risk and resilience factors for psychopathology, and (5) methodological advances in the field of neuroscience.

Markers and Mechanisms of Psychopathological Pathways

A major focus in developmental psychiatry has been to understand neurobiological and psychosocial changes to identify targets for prevention and prediction strategies (Rosenberg et al., 2018). For example, for attention-deficit hyperactivity disorder (ADHD), a common mental disorder characterized by core symptoms of impulsivity, hyperactivity, and inattention (Polanczyk and Rohde, 2007), changes in executive functions, in particular, in inhibition and working memory have been reported (Faraone et al., 2015). This might be accompanied with frontostriatal and parietal hypoactivation found during inhibition (van Rooij et al., 2015) and spatial working memory in ADHD, with ADHD showing persistent difficulties with working memory operations (Martinussen et al., 2005).

Thus, obviously a wide range and diverse, rather unspecific, pattern of neuropsychobiological responses have been observed already for one type of mental disorder. It is therefore not surprising that sensitivity for only one disorder in such mechanistic links is still limited. Common mechanisms may contribute to different forms of psychopathology and their associated symptomatology. For example, in anxiety disorders, shared key features may not only range along the anxiety spectrum, but there may also be subgroups within anxiety disorders that share common mechanisms (McTeague et al., 2009, 2010, 2012; Flor and Nees, 2014). Mental health seems therefore to span along construct dimensions, reflected in neurobiological changes (Dias et al., 2015; Kotov et al., 2017). Such aspects have been proposed by the Research Domain Criteria (RDoC) Project, an initiative of the National Institute of Mental Health (Cuthbert and Insel, 2010, 2013). The RDoC project suggests to base the classification of mental disorders on ***dimensions of observable behavior and neurobiological measures*** related to these functions rather than on symptom-based descriptive categorical diagnoses. Such dimensions represent the loadings onto symptoms, where each individual receives a dimensional score. It builds a basis for parsing heterogeneity to be predicated by abnormalities in multiple distinct system- and circuit-based psychobiological changes that might be present across versus within disorders, and manifest along dimensions (Marquand et al., 2016). This is also important given that individual differences can already be observed at the subclinical stage. We have shown this for conduct problems and clinically relevant brain changes to negative affective processing in a healthy adolescent sample: regression analyses revealed a significant linear increase of left orbitofrontal cortex (OFC) activity with increasing conduct problems up to the clinical range, while in the high conduct problems group, a significant inverted u-shaped effect indicated that left OFC responses decreased again in individuals with high conduct problems (Böttinger et al., 2021).

Several studies have, however, also acknowledged a ***general, and thus one, dimension of psychopathology***, the so-called p factor, underlying multiple disorders (Caspi et al., 2014; Caspi and Moffitt, 2018). This proposition of such a single p factor for common mechanisms can be seen in line RDoC when relying on a transdiagnostic understanding, i.e., that the same mechanism underlies different diagnoses, yet it is different from RDoC when we applying the heterogeneity assumption. Using data from the longitudinal New Zealand Dunedin Multidisciplinary Health and Development Study (Caspi et al., 2020), one of the most recent child cohort studies examining children longitudinally into adulthood, it was shown that the p factor significantly improved the model fit when externalizing, internalizing, and thought disorder subfactors were integrated and allowed to inter-correlate (Caspi and Moffitt, 2018). This can be used as an evidence of common a common mechanism underlying these disorders, but, on the other hand, ignores potential individual differences within such a mechanism, for example, on subdimensions of this mechanism.

The p factor has also been shown to substantially overlap with the p factor in childhood and risk for mental disorders in adulthood (Allegrini et al., 2020), and the stability of the p factor across childhood was highly driven by genetic influences (Neumann et al., 2016). This may be further be seen as an indication of a rather static marker of psychopathology, representing an important *snap-shot*, but potentially being not such valid when describing changes and developments over time. An impact of genetic constitutions has further been unraveled through studies on the so-called polygenic risk scores (PRS) that reflects an individual's inherited susceptibility to a disease. Higher levels of PRS for internalizing problems, for example, determined adolescents' co-occurring internalizing/externalizing problems, indicating common genetic components for externalizing and internalizing disorders, and by lower levels of the aggression PRS through greater early childhood behavioral inhibition (Wang et al., 2020). An ADHD PRS was shown to distinguish between individuals with a persistently high level of ADHD symptoms (through to adolescence) from those individuals whose symptoms declined or remitted by adolescence (childhood-limited) (Riglin et al., 2016).

While these markers may be useful for the routine clinical practice, not all are real indicators for underlying mechanisms

and useful to capture an individual disease process, having been identified through rather pragmatic pipelines. This is not an invalid procedure *per se*, but needs to be carefully considered within the respective context, i.e., whether they are used as diagnostic, prognostic, or predictive target; otherwise, their effectiveness might be strongly be limited or even misleading.

HOW TO DEAL WITH THE CURRENT KNOWLEDGE ON MARKERS AND MECHANISMS OF PSYCHOPATHOLOGICAL PATHWAYS?

Falk et al. (2013) have already raised several important questions in this respect. Among those questions, one covered brain–behavior mechanisms and interactions: "What would a 'representative group of brains' tell us about the generalizability of current samples and current findings regarding brain–behavior mechanisms?" This question nicely implied what we are referring to the heterogeneity *problem*, and illustrates that thinking in a one-dimensional domain, grouping individuals in comprehensive fashion, does not always also mean that we end up with more concrete and specific conclusions.

And even if we have arrived at the individual level, a following question would come up, namely, "How do individual differences in brain structure and function affect cognitive, affective, and behavioral outcomes and how do social situations and broader environmental contexts interact with these processes?" For example, several previous studies on intelligence suggested a high heritability of the intelligence quotient; however, more recent work indicates that for the whole population the heritability is not as high as previously assumed (Turkheimer et al., 2003). While the earlier studies had used samples primarily with individuals showing a high socioeconomic status (SES), more recent studies stem from more representative samples. Since the SES is a significant moderator of genetic heritability, the heritability was higher in high SES, above 70%, and in contrast with only 10% in low SES individuals (Turkheimer et al., 2003; Henrich et al., 2010), which would lead to inconsistent estimates of outcomes.

It is thus crucial to disentangle ***more specifically inter-relationships among the factors of interest and map them among developmental trajectories.*** With regard to externalizing symptoms, some individuals have constantly high symptoms, while intermediate groups of individuals shift up or down slowly or rapidly. Similar patterns were observed for internalizing symptoms. For adolescence, the internalizing trajectory was also found being independent of high externalizing trajectories, and persisting externalizing problem scores were associated with decreasing internalizing scores, and early environmental risk factors and sex predicted externalizing trajectories to a larger extent than internalizing (e.g., Nivard et al., 2017). In this respect, we need to specify the relevant mechanisms of change for such shifts, and this also includes the evaluation of how the brain develops across the life span in individuals with and without mental disorders, that warrants further investigation (e.g., Baltes et al., 1977; Paus, 2010; Zelazo and Paus, 2010). Moreover, as addressed above, ***symptoms are highly heterogeneous*** and ***overlapping*** (Moffitt et al., 2007; Caspi et al., 2020). In the Dunedin Study (Caspi et al., 2020), for example, less than 15% of participating individuals diagnosed with externalizing or internalizing disorders showed a homotypic symptomatology (Caspi et al., 2020).

Strategies for understanding the etiology of mental disorders in this respect need to capitalize on data from ***longitudinal, cohort studies*** like the Mannheimer Risikokinderstudie (MARS; Laucht et al., 2000b), the Adolescent Brain Cognitive Development (ABCD) study (Karcher and Barch, 2020), the Kinder- und Jugendgesundheitssurvey (KIGGS; Mauz et al., 2019), the Saguenay Youth Study (Pausova et al., 2017), or the Imaging Genetics (IMAGEN) study (Schumann et al., 2010). Those examine the psychosocial and neurobiological etiology, prevalence, and developmental trajectories of (sub-)clinical symptoms indicative of vulnerability for future psychopathology in children, adolescents, and adults. However, so far, brain development has been investigated mainly during middle childhood (e.g., ABCD), adolescence (e.g., IMAGEN), or early adulthood (e.g., MARS). Up to date only few studies, such as IMAGEN, Generation R (Jaddoe et al., 2006), Saguenay Youth Study, or NCANDA (Rohlfing et al., 2014), pursued a longitudinal neuroimaging design. This aggravates conclusions about the cause and effect relationships, as brain changes can be both a consequence of behaviors or experiences and a causing factor.

To map and model changes over time in a non-linear fashion, study designs need to integrate more than two assessment time points, which is so far realized, for example, in IMAGEN starting during adolescence and follow-ups around every 2 years. Given the lack of neurodevelopmental studies early in life, it is still unclear whether especially early developmental curves of the brain serve as possible predictors for psychopathology, and whether and how this depends on different risk constellations and interactions between neurobiology, genes, and behavior.

It is therefore important that we obtain data very early in life along such cohort studies and longitudinal approaches starting already in birth, and apply a precise and comprehensive analysis of phenotypic abnormalities (***deep phenotyping***), integrating the individual components of the phenotype and thus being able to more specifically consider age of symptom onset and underlying psychobiosocial mechanisms (e.g., Holz et al., 2020; Thapar and Riglin, 2020). For example, about 45% of individuals who develop a mental disorder early in life reported aversive experiences during childhood (Green et al., 2010). This may depend on region-specific sensitive time windows of brain development (Callaghan and Tottenham, 2016), during which functionally neurobiological changes may occur as consequences of early negative experiences (Nelson and Gabard-Durnam, 2020). Moreover, previous etiological studies showed an earlier and more frequent manifestation at every age as well as greater stability for externalizing compared to internalizing disorders (e.g., Laucht et al., 2000a). They were related to very early psychosocial risk factors, which became apparent in early childhood and include mother–child interactions (Laucht et al., 2000a). Moreover, age of onset was discussed as one of the

crucial factors for the course of symptom severity in behavioral disorders: an earlier development of symptoms is associated with stronger burden and chronification (Perra et al., 2020). Moreover, early risk factors also explain a substantial proportion of variance in behavioral and neurobiological mechanisms later in life (e.g., Holz et al., 2014, 2015) and might be important for the persistence of symptoms from child- into adulthood, as shown, for example, for ADHD or anxiety disorders (e.g., Thapar and Riglin, 2020). Thus, only the investigation of risk and resilience mechanisms already early in life enables a proper integration of this knowledge into prevention and early intervention programs.

Environmental Influences and Neurobiological Measures: Focus on Psychosocial Risk and Resilience

Under the premise to identify neurobehavioral markers as targets for early intervention and prevention, malleability of these markers is critical. Twin studies suggest that genetic factors indeed contribute prominently to continuity in mental health problems, but that environmental influences are a major contributor to dynamic change. Further, in children younger than 5 years of age, the influence of environmental factors is even more important than genetic factors. In short, the younger the child, the more dependent and vulnerable it seems to be in relation to the surrounding environment. Importantly, children differ from adults in their unique physiological and behavioral characteristics and the potential exposure to risks in the environment (e.g., Rutter et al., 2001; Ronald, 2011; Hannigan et al., 2017).

Along the proposed ***deep phenotyping*** approaches of psychopathology, those should therefore not only address the neurobiological, but particularly also ***environmental and psychosocial domains and their dynamic interplay*** to identify the complex etiology of mental disorders (Holz et al., 2020). We need to extent existing knowledge in neuroscience research to broader relevant populations, and the ways that macrolevel structures (e.g., social structure, neighborhood safety, school quality, and media exposure) influence neural processes (Paus, 2010), thus integrating social and neurobiological perspectives (see **Figure 1**).

For the social domain, it has been demonstrated that socially well-connected individuals live longer compared to those with weaker social bonds (Holt-Lunstad et al., 2010). Caregivers, including their parenting styles, are important factors for developmental processes, for example, influencing emotion regulation and emotional reactivity (Bernier et al., 2016; Holz et al., 2018). Distress of the infant can be effectively reduced by responsive caregivers decreasing the development of fear over time (Leerkes et al., 2009), also in association with brain phenotypes (Ellis et al., 2011). Interestingly, an individual's social ties have not only beneficial health effects, but social connections also determine the way individuals perceive their surrounding environment. Loneliness and social exclusion bias perception of the social world to be more threatening or vice versa. This is related to stronger activity in the visual cortex in response to social stimuli, and thus greater attention to negative social information (Cacioppo et al., 2009). A lifestyle with infrequent and/or negative social contacts might therefore negatively affect the processing of socioemotional and reward stimuli (Etkin et al., 2006; Adolphs, 2010; Rademacher et al., 2010).

On the other hand, social support may be beneficial for well-being being an important mediator in health and disease. Specifically, receiving social support affects vmPFC brain regions that are relevant in inhibiting activity of regions, like the dorsal anterior cingulate cortex and anterior insula, associated with threat and stress processing (Eisenberger et al., 2011). Moreover, perceived social support has been shown to moderate the well-known relationships between trait anxiety and amygdala reactivity (Hyde et al., 2011). Adolescents with a negative family history and relatively severe life stress showed increased amygdala reactivity to threat (Swartz et al., 2015). Models of environmental sensitivity (Pluess, 2015) include hypotheses on sensitivity to processing sensory information, biological susceptibility to context, and differential susceptibility (Aron and Aron, 1997; Belsky et al., 2007; Boyce, 2016). It is suggested that individuals who are more sensitive show not only a higher risk for consequences of adverse environmental conditions (e.g., Monroe and Simons, 1991), but are also more responsive to positive characteristics of the environment (Pluess and Belsky, 2013). This framework may inform initiatives that aims to identify individuals who are most affected by adverse environmental influences (Meaney, 2018), and in turn, who benefit most from treatment strategies (de Villiers et al., 2018), and can also inform prevention.

With respect to associated brain changes, the amygdala and the prefrontal cortex (PFC) have been suggested to play a key role. Although the volume of the amygdala rapidly enlarges within the first years of life, structural changes process until 4 years of age in girls and 18 years in boys (Callaghan and Tottenham, 2016). Early life exposures might therefore result in alterations of brain regions like the amygdala in childhood, with an influence also of later stress specifically in boys. Aside prefrontal cortical gray matter increases are observed until adolescence, which indicates higher sensitivity to environmental influences during childhood and adolescence (Callaghan and Tottenham, 2016).

However, the research on neurobiological variability and individual differences in environmental sensitivity is scarce, and has mostly been conducted in adult and adolescent populations. Early childhood neurobiological phenotypes are still lacking. In this respect, environmental variables can act as moderators of interest and might have also contributed to the neurobiological phenotype and thus the sensitivity marker of interest. Moreover, environmental factors can also function as predictors. It also becomes clear that we would benefit from studies on brain phenotypes characterized at or shortly after birth, and thus have been only minimally influenced by environmental experiences (Nolvi et al., 2020). Connectivity between the amygdala, insula, and ventral medial PFC (vmPFC), which have been implicated in individual differences in the processing of fear and in the risk to the development of mental disorders, has also been identified as predictor of higher fear and sadness already in the newborn from 6 to 24 months of age. Specifically, amygdala–insula connectivity and amygdala–vmPFC connectivity were relevant for fear and sadness trajectories, respectively

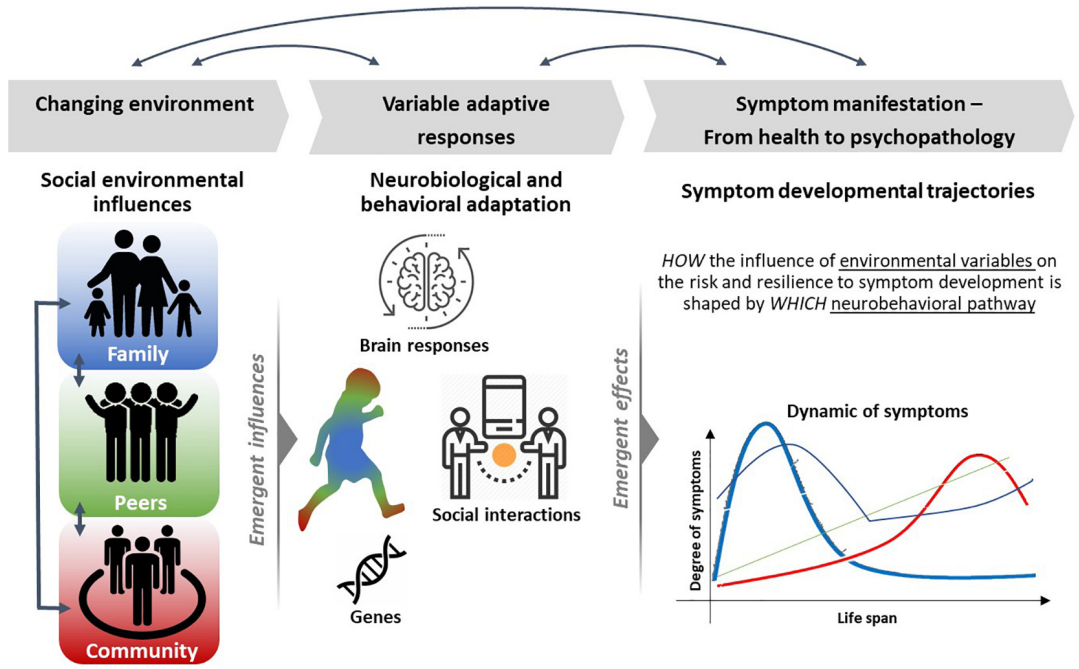

FIGURE 1 | Overview on an integrative and multimodal approach to study how environmental influences dynamically shape risk and resilience on the level of neurobehavioral adaptation at different developmental stages. There are several influences through the social environment (*left*) and individuals differ in their degree to which they are exposed to environmental variation. Such variations may be reflected in neurobiopsychological and behavioral reactivity, for example, through brain–behavior relationships or gene–environmental interactions (*middle*). These variations underlie life-time-dependent changes over time and together represent trajectories of risk into and resilience for maladaptive behavior and psychopathological symptomatologies (*right*).

(Thomas et al., 2019), and amygdala–vmPFC connectivity at birth predicted cognitive development at 6 months of age (Nigg, 2006; Degnan and Fox, 2007; Gartstein et al., 2012).

At these very early times of life, it also becomes evident that biological influences during pregnancy are rapidly influencing developing fetal brain systems, particularly those that are altered in mental disorders. Maternal cortisol levels during pregnancy, indicating elevated levels of psychosocial stress, were significantly related to stronger connectivity between the amygdala and brain regions relevant for the processing and integration of sensory information as well as the default mode network in females and reduced amygdala connectivity to these regions in males (Graham et al., 2019). Further, in females, this connectivity mediated the association between maternal cortisol and higher internalizing symptoms (Graham et al., 2019).

Advanced Methodological Issues to Achieve a Better Clinical Impact

Data from population-based cohort studies might help to increase the transfer into the clinical system selecting an adequate and most beneficial treatment, and particularly prevention strategy. However, the associated notion that symptom observations and reports are highly correlated within and also common across disorders (Hahn et al., 2017) does, however, not mean to completely abandon categorical approaches when parsing heterogeneity.

To pursue a promising avenue toward neuro-psycho-biological research that can have practical impact on psychiatric healthcare, we need to shift many of the traditional approaches and tools, and also need to combine approaches to build on each other. This shift becomes even more important when thinking about the still existing translation gap into clinically useful prevention and intervention strategies. A dimensional approach would result in different, individually tailored, treatment strategies depending on the occurrence of specific symptoms or additional modulators such as comorbidity of cognitive status, which could improve the treatment outcome. Methods applying categorical approaches can be added providing information on clusters or subtypes of individuals who share common features, for example, in the case where groups with ***extreme phenotypes of psychopathology*** still show alterations in multiple mechanisms.

To explore moderators of the neurobiological prediction of behavior and model clinical variance, we summarize some of the available analysis methods, which we think are reliable in this respect. This includes methodological issues like switching, e.g., from *group-level statistics* that compare clinical with non-clinical, healthy control groups, to ***approaches that characterize individual heterogeneity and developmental processes along subclinical and clinical symptom continua*** (Becht and Mills, 2020). Moreover, the use of multivariate and computational modeling and analyses to break down high-dimensional data from these large and representative samples regarding brain–behavior mechanisms is required

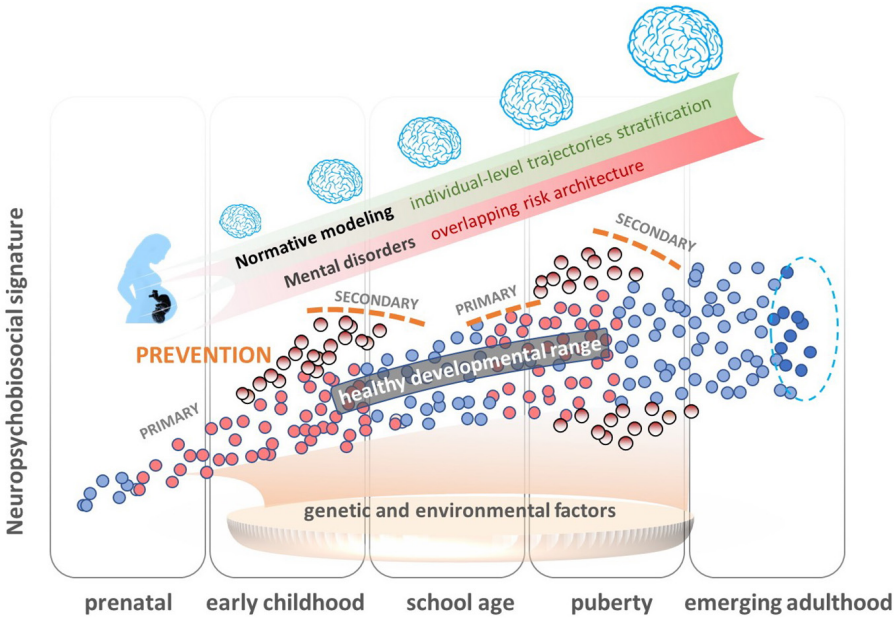

FIGURE 2 | Overview of the normative modeling approach as an innovative analytical framework for parsing underlying biological heterogeneity within epidemiological cohorts without dichotomizing into cases and controls. This approach uses probabilistic regression methods, machine learning, and artificial intelligence (AI) methods to characterize variation across the population, estimate normative models of development (e.g., brain development) during critical phases of vulnerability, and detect individual differences in risk signatures in clinical cohorts. These models therefore enable predictions at an individual subject level within the population, to explore how these signatures predispose individuals to somatic and mental health outcomes, such as failure to thrive, motor, and language delay, behavioral and emotional disorders, and inferences on how deviation patterns map onto biological underpinnings informing primary and secondary prevention (early intervention) approaches. Dots represent individual neuropsychobiosocial signatures. Blue dots represent individuals who stay in the normal range throughout development, and red are those individuals who are in the normal range at a specific developmental time, but move then out of this range (pale red dots).

(Falk et al., 2013; Jollans and Whelan, 2018). In this context, the propagation of a ***prediction or risk calculation*** pipelines might be a promising avenue. Information on mechanisms underlying symptom dimensions and trajectories may be implemented into risk assessment and prognosis procedures to supplement clinical decision-making.

Cross-Lagged Modeling

To provide support for developmental models, cross-lagged panel models (CLPMs) and autoregressive latent trajectory models with structured residuals (ALT-SR) have been used. With CLPM, as a type of discrete time structural equation modeling, panel data with two or more variables, measured at two or more time points, are analyzed. In this way, any directional effects of one variable on another variable at different time points can be estimated (Kuiper and Ryan, 2018). The ALT-SR is an extension with a crosslagged (or other) structure that is fit to the time-specific residuals from a parallel process latent growth curve model. The validity of CLPM and ALT-SR has, for example, been tested for cascades of externalizing and internalizing disorders, which have strong tendency to co-occur from childhood (Rhee et al., 2015; Martel et al., 2017). Such analyses are important for understanding the cause and nature of their co-occurrence, e.g., whether there is a directional or reciprocal causal relation and how this is mediated, and can have implications for treatment. Although analyses using CPLM and ALT-SR were consistent, the use of ALT-SR resulted in a better fit than CLPMs. Moreover, there is evidence for effects only apparent when applying the ALT-SR. This includes a negative effect of externalizing on internalizing problems in adolescence, while effects of internalizing on externalizing problems were found for both ALT-SR and CLPM (Murray et al., 2020). With typically utilized CLPMs, between- and within-person processes cannot be disaggregated, and thus their parameters reflect a difficult-to-interpret blend of the two. This disadvantage can be solved using ALT-SR (Curran et al., 2014). With ALT-SR, effects of unmeasured between-person confounds are partialed out (Berry and Willoughby, 2017).

Normative Modeling

Normative modeling (Marquand et al., 2016, 2019; **Figure 2**) provides an innovative analytical framework for parsing underlying biological heterogeneity within epidemiological and clinical cohorts, providing inferences beyond the level of mean group differences. This approach uses Bayesian regression methods, such as Gaussian process regression, to characterize variation across the population as a function of clinical predictor variables, such as age, while taking predictive uncertainty into account. These models do not require deviations to overlap across individuals (e.g., in the same brain regions) and enable statistical inferences at the level of the individual participant in order to quantify deviations from the expected normative pattern from a reference cohort. In the context

of psychiatry, normative modeling has been used to explain neurodevelopmental deviation patterns in clinical samples. As such, it has been applied to study developmental variability in cortical thickness (Zabihi et al., 2019) and brain asymmetry (Floris et al., 2020) in autism spectrum disorders, or gray and white matter in schizophrenia spectrum disorders (Wolfers et al., 2018). These studies have demonstrated that age-related deviations are predictive of clinical symptom scores and provided evidence that clinically relevant deviation patterns have minimal inter-subject consistency with neurobiological effects that overlap only in a few patients. In addition, spatial deviations were often different from classical case-control findings, thus corroborating the need to consider large variation across subjects at the individual level. Despite methodological challenges that may be encountered in large population-based cohorts with missing data and variance attributable to study site, normative modeling provides a promising method to estimate development (e.g., brain development) during critical phases of vulnerability and, thereby, derive individual risk and resilience signatures.

However, we also need to note that normative modeling usually requires very large sample sizes and they largely depend on the information content of the included variables. Variables representing only unprecise measures of the mechanisms of interest lower any added value at the clinical or individual level. Other machine learning and computational modeling approaches rather focus on the identification of processes and mechanisms underlying observable data, e.g., in reinforcement learning or dynamical systems. Parameters derived from these models can increase the predictive value of normative modeling as well as of other classification or clustering approaches (e.g., Brodersen et al., 2013).

Prediction and Risk Calculation

Aside normative modeling to perform prediction analyses, recent studies also acknowledge the development of a so-called risk calculator (Caye et al., 2020). This might be helpful for personalized medicine to predict, for example, adult ADHD from childhood characteristics, based on the representative population cohort ALSPAC-UK with 5,113 participants, followed from birth to age 17 (Caye et al., 2020). So far, the course of ADHD could not have been correctly predicted in the clinical setting based on assessments in children nor could have been prevention adequately performed for those at risk. Caye et al. (2020) therefore aimed to combine knowledge about risk factors into a multivariable risk score, similar to frameworks in the context of cardiovascular diseases, instead of using information from a single risk factor like the presence of subthreshold symptoms or of a first-degree relative diagnosed with ADHD (Brent et al., 2015; Taylor et al., 2015; Buntrock et al., 2016). They also validated their risk tool using further cohorts including the 1993 Pelotas Birth Cohort (Brazil, 3,911 participants, birth to age 18), the MTA clinical sample (United States, 476 children with ADHD and 241 controls followed for 16 years from a minimum of 8 and a maximum of 26 years old), and the E-Risk cohort (United Kingdom, 2,040 participants, birth to age 18). An add on by the knowledge and results from neurobiological studies might provide an important link to biomarker identifications.

OUTLOOK

A developmental perspective in psychiatry is helpful. In order to detect mental health outcomes, studying individual differences in brain development is a key aspect. However, so far most longitudinal neuroimaging studies tested effects on a group level, for example, to identify those individuals who had lower brain volume at a baseline level or also those who show accelerated brain volume at follow-up testing compared with individuals who had higher brain volumes at baseline.

To overcome this constraint, such individual differences should then also be used for prediction analyses to address heterogeneity in developmental trajectories. Future research needs to treat age not as a confound, but rather as the primary effector of interest. We need to incorporate processes of brain maturation in youth, and thus the perspective of developmental psychiatry, when determining risk and resilience factor for mental disorders. We need longitudinal large-scale studies and data, starting already early in life. Moreover, it is vital to apply a translational, transnosological, and multi-disciplinary systems approach and go beyond on brain development, which most of the available studies focus on. There is still a lack of integration of molecular, immunological, endocrinological, environmental, social, physiological, cognitive, and brain imaging readouts. Resources should include large-scale clinical (i.e., patient-based), at-risk, and epidemiological (i.e., population-based) *cohorts* that underwent comprehensive longitudinal deep phenotyping protocols, including state-of-the-art digital health technologies. Applying cutting-edge bioinformatics, we can overcome transdisciplinary research gaps and clinical service boundaries by providing multimodal predictive signatures of risk and protection. The combination with unique assembly of developmental tools, models, and human cohort readouts will allow for a deepened mechanistic understanding why children develop trajectories resulting in disease or recovery. Insight into the developmental mechanisms can then inform extensive translational intervention capacities and networks.

AUTHOR CONTRIBUTIONS

FN wrote the manuscript. LD, NH, MR, and TB reviewed the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the research consortium on ADHD, ESCA-Life, funded by the German Federal Ministry of Education and Research (FKZ 01EE1408E). Bundesministerium für Bildung und Forschung (Forschungsnetz IMAC-Mind 01GL1745B, TP1/TP2).

REFERENCES

Adolphs, R. (2010). What does the amygdala contribute to social cognition? *N.Y. Acad. Sci.* 1191, 42–61. doi: 10.1111/j.1749-6632.2010.05445.x

Allegrini, A. G., Cheesman, R., Rimfeld, K., Selzam, S., Pingault, J. B., Eley, T. C., et al. (2020). The p factor: genetic analyses support a general dimension of psychopathology in childhood and adolescence. *J. Child Psychol. Psychiatry* 61, 30–39. doi: 10.1111/jcpp.13113

Aron, E. N., and Aron, A. (1997). Sensory-processing sensitivity and its relation to introversion and emotionality. *J. Pers. Soc. Psychol.* 73, 345–368.

Baltes, P. B., Reese, H. W., and Nesselroade, J. R. (1977). *Life-Span Developmental Psychology: Introduction to Research Methods*. Hillsdale, NJ: Erlbaum.

Becht, A. I., and Mills, K. L. (2020). Modeling individual differences in brain development. *Biol. Psychiatry* 88, 63–69. doi: 10.1016/j.biopsych.2020.01.027

Belsky, J., Bakermans-Kranenburg, M. J., and Van Ijzendoorn, M. H. (2007). For better and for worse: differential susceptibility to environmental influences. *Curr. Dir. Psychol. Sci.* 16, 300–304. doi: 10.1111/j.1467-8721.2007.00525.x

Bernier, A., Carlson, S. M., and Whipple, N. (2016). From external regulation to self-regulation: early parenting precursors of Young children's executive functioning. *Child Dev.* 81, 326–339. doi: 10.1111/j.1467-8624.2009.01397.x

Berry, D., and Willoughby, M. T. (2017). On the practical interpretability of cross-lagged panel models: rethinking a developmental workhorse. *Child Dev.* 88, 1186–1206.

Böttinger, B. W., Baumeister, S., Millenet, S., Barker, G. J., Bokde, A. L. W., Büchel, C., et al. (2021). Orbitofrontal control of conduct problems? Evidence from healthy adolescents processing negative facial affect. *Eur. J. Child Adolesc. Psychiatry* doi: 10.1007/s00787-021-01770-1 [Epub ahead of print].

Boyce, W. T. (2016). Differential susceptibility of the developing brain to contextual adversity and stress. *Neuropsychopharmacol. Rep.* 41, 142–162. doi: 10.1038/npp.2015.294

Brent, D. A., Brunwasser, S. M., Hollon, S. D., Weersing, V. R., Clarke, G. N., Dickerson, J. F., et al. (2015). Effect of a cognitive-behavioral prevention program on depression 6 years after implementation among at-risk adolescents: a randomized clinical trial. *Arch. Gen. Psychiatry.* 72, 1110–1118.

Brodersen, K. H., Desero, L., Schlangenhauf, F., Lin, Z., Penny, W. D., Buhmann, J. M., et al. (2013). Dissecting psychiatric spectrum disorders by generative embedding. *Neuroimage Clin.* 16, 98–111. doi: 10.1016/j.nicl.2013.11.002

Buntrock, C., Ebert, D. D., Lehr, D., Smit, F., Riper, H., Berking, M., et al. (2016). Effect of a web-based guided self-help intervention for prevention of major depression in adults with subthreshold depression: a randomized clinical trial. *JAMA* 315, 1854–1863. doi: 10.1001/jama.2016.4326

Cacioppo, J. T., Norris, C. J., Decety, J., Monteleone, G., and Nusbaum, H. (2009). In the eye of the beholder: individual differences in perceived social isolation predict regional brain activation to social stimuli. *J. Cogn. Neurosci.* 21, 83–92. doi: 10.1162/jocn.2009.21007

Callaghan, B. L., and Tottenham, N. (2016). The neuro-environmental loop of plasticity: a cross-species analysis of parental effects on emotion circuitry development following typical and adverse caregiving. *Neuropsychopharmacol. Rep.* 41, 163–176.

Caspi, A., Houts, R. M., Ambler, A., Danese, A., Elliott, M. L., Hariri, A., et al. (2020). Longitudinal assessment of mental health disorders and comorbidities across 4 decades among participants in the dunedin birth cohort study. *JAMA Netw. Open* 3:e203221. doi: 10.1001/jamanetworkopen.2020.3221

Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., et al. (2014). The p factor. *Clin. Psychol. Sci.* 2, 119–137. doi: 10.1177/2167702613497473

Caspi, A., and Moffitt, T. E. (2018). All for one and one for all: mental disorders in one dimension. *Am. J. Psychiatry.* 175, 831–844. doi: 10.1176/appi.ajp.2018.17121383

Caye, A., Agnnew-Blais, J., Arseneault, L., Goncalves, H., Kieling, C., Langley, K., et al. (2020). A risk calculator to predict adult attention-deficit/ hyperactivity disorder: generation and external validation in three birth cohorts and one clinical sample. *Epidemiol. Psychiatr. Sci.* 29:e37.

Curran, P. J., Howard, A. L., Bainter, S. A., Lane, S. T., and McGinley, J. S. (2014). The separation of between-person and within-person components of individual change over time: a latent curve model with structured residuals. *J. Consult. Clin. Psychol.* 82, 879–894. doi: 10.1037/a0035297

Cuthbert, B. N., and Insel, T. R. (2010). Toward new approaches to psychotic disorders: the NIMH research. *Schizophr. Bull.* 36, 1061–1062. doi: 10.1093/schbul/sbq108

Cuthbert, B. N., and Insel, T. R. (2013). Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Med.* 11:126. doi: 10.1186/1741-7015-11-126

de Villiers, B., Lionetti, F., and Pluess, M. (2018). Vantage sensitivity: a framework for individual differences in response to psychological intervention. *Soc. Psychiatry Psychiatr. Epidemiol.* 53, 545–554. doi: 10.1007/s00127-017-1471-0

Degnan, K. A., and Fox, N. (2007). Behavioral inhibition and anxiety disorders: multiple levels of a resilience process. *Dev. Psychopathol.* 19, 729–746.

Dias, T. G.C., Iyer, S. P., Carpenter, S. D., Cary, R. P., Wilson, V. B., Mitchel, S. H., et al. (2015). Characterizing heterogeneity in children with and without ADHD based on reward system connectivity. *Dev. Cogn. Neurosci.* 11, 155–174. doi: 10.1016/j.dcn.2014.12.005

Eisenberger, N. I., Master, S. L., Inagaki, T. K., Taylor, S. E., Shirinyan, D., Lieberman, M. D., et al. (2011). Attachment figures activate a safety signal-related neural region and reduce pain experience. *Proc. Natl. Acad. U. S. A.* 108, 11721–11726. doi: 10.1073/pnas.1108239108

Ellis, B. J., Boyce, W. T., Belsky, J., Bakermans-Kranenburg, M. J., and van Ijzendoorn, M. H. (2011). Differential susceptibility to the environment: an evolutionary—neurodevelopmental theory. *Dev. Psychopathol.* 23, 7–28.

Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R., and Hirsch, J. (2006). Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron* 51, 871–882. doi: 10.1016/j.neuron.2006.07.029

Falk, E. B., Hyded, L. W., Mitchelle, C., Faule, J., Gonzalez, R., Heitzegi, M. M., et al. (2013). What is a representative brain? Neuroscience meets population science. *Proc. Natl. Acad. Sci. U. S. A.* 110, 17615–17622. doi: 10.1073/pnas.1310134110

Faraone, S. V., Asherson, P., Banaschewski, T., Biederman, J., Buitelaar, J. K., Ramos-Quiroga, J. A., et al. (2015). Attention-deficit/hyperactivity disorder. *Nat. Rev. Dis. Primers* 1:15020.

Flor, H., and Nees, F. (2014). Learning, memory and brain plasticity in posttraumatic stress disorder: context matters. *Restor. Neurol. Neurosci.* 32, 95–102. doi: 10.3233/RNN-139013

Floris, D. L., Wolfers, T., Zabihi, M., Holz, N. E., Zwiers, M. P., Charman, T., et al. (2020). Atypical brain asymmetry in Autism-A candidate for clinically meaningful stratification. *Biol. Psychiatry. Cogn. Neurosci. Neuroimaging* 25, 2451–9022.

Gartstein, M. A., Putnam, S. P., and Rothbart, M. K. (2012). Etiology of preschool behavior problems: contributions of temperament attributes in early childhood. *Infant. Mental Health J.* 33, 197–211.

Graham, A. M., Rasmussen, J. M., Entringerb, S., Ben Ward, E., Rudolph, M. D., Gilmore, J. H., et al. (2019). Maternal cortisol concentrations during pregnancy and sex specific associations with neonatal amygdala connectivity and emerging internalizing behaviors. *Biol. Psychiatry* 85, 172–181. doi: 10.1016/j.biopsych.2018.06.023

Green, J. G., Mclaughlin, K. A., Berglund, P. A., Gruber, M. J., Sampson, N. A., Zaskavsky, A. M., et al. (2010). Childhood adversities and adult psychiatric disorders in the national comorbidity survey replication I: associations with first onset of DSM-IV disorders. *Arch. Gen. Psychiatry* 67, 113–123. doi: 10.1001/archgenpsychiatry.2009.186

Gustavsson, A., Svensson, M., Jacobi, F., Allgulander, C., Alonso, J., Beghi, E., et al. (2011). Cost of disorders of the brain in Europe 2010. *Eur. Neuropsychopharmacol.* 21, 718–779. doi: 10.1016/j.euroneuro.2011.08.008

Hahn, T., Nierenberg, A. A., and Whitfield-Gabrieli, S. (2017). Predictive analytics in mental health: applications, guidelines, challenges and perspectives. *Mol. Psychiatry* 22, 37–43. doi: 10.1038/mp.2016.201

Hannigan, L. J., McAdams, T. A., Plomin, R., and Eleym, T. C. (2017). Parent- and child-driven effects during the transition to adolescence: a longitudinal, genetic analysis of the home environment. *Dev. Sci.* 20:e12432. doi: 10.1111/desc.12432 Epub 2016 Jun 19.

Henrich, J., Heine, S. J., and Norenzayan, A. (2010). The weirdest people in the world? *Behav. Brain Sci.* 33, 61–83.

Holt-Lunstad, J., Smith, T. B., and Layton, J. B. (2010). Social relationships and mortality risk: a meta-analytic review. *PLoS Med.* 7:e1000316. doi: 10.1371/journal.pmed.1000316

Holz, N. E., Boecker, R., Baumeister, S., Hohm, E., Zohsel, K., Buchmann, A. F., et al. (2014). Effect of prenatal exposure to tobacco smoke on inhibitory control:

neuroimaging results from a 25-year prospective study. *JAMA Psychiatry* 71, 786–796. doi: 10.1001/jamapsychiatry.2014.343

Holz, N. E., Boecker, R., Hohm, E., Zohsel, K., Buchmann, A. F., Blomeyerm, D., et al. (2015). The long-term impact of early life poverty on orbitofrontal cortex volume in adulthood: results from a prospective study over 25 years. *Neuropsychopharmacol. Rep.* 40, 996–1004. doi: 10.1038/npp.2014.277

Holz, N. E., Boecker-Schlier, R., Jennen-Steinmetz, C., Hohm, E., Buchmann, A. F., Blomeyer, D., et al. (2018). Early maternal care may counteract familial liability for psychopathology in the reward circuitry. *Soc. Cogn. Affect. Neurosci.* 13, 1191–1201. doi: 10.1093/scan/nsy087

Holz, N. E., Nees, F., Meyer-Lindenberg, A., Tost, H., Hölling, H., Keil, T., et al. (2020). [Cohort studies in child and adolescent psychiatry]. *Nervenarzt* 92, 208–218. doi: 10.1007/s00115-020-01018-4

Hyde, L. W., Gorka, A., Manuck, S. B., and Hariri, A. R. (2011). Perceived social support moderates the link between threat-related amygdala reactivity and trait anxiety. *Neuropsychologia* 49, 651–656. doi: 10.1016/j.neuropsychologia.2010.08.025

Jaddoe, V. W. V., Mackenbach, J. P., Moll, H. A., Steegers, E. A. P., Tiemeier, H., Verhulst, F. C., et al. (2006). The generation R study: deasign and cohort profile. *Eur. J. Epidemiol.* 21, 475–484.

Jollans, L., and Whelan, R. (2018). Neuromarkers for mental disorders: harnessing population. *Neurosci. Front. Psychiatry* 9:242. doi: 10.3389/fpsyt.2018.00242

Karcher, N. R., and Barch, D. M. (2020). The ABCD study: understanding the development of risk for mental and physical health outcomes. *Neuropsychopharmacology* 46, 131–142.

Kessler, R. C., Amminger, G. P., Aguilar-Gaxiola, S., Alonso, J., Lee, S., and Ustun, T. B. (2007). Age of onset of mental disorders: a review of recent literature. *Curr. Opin. Psychiatry* 20, 359–364. doi: 10.1097/YCO.0b013e32816ebc8c

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., and Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch. Gen. Psychiatry* 62, 593–602. doi: 10.1001/archpsyc.62.6.593

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., et al. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): a dimensional alternative to traditional nosologies. *J. Abnorm. Psychol.* 126, 454–477. doi: 10.1037/abn0000258

Kuiper, R. M., and Ryan, O. (2018). Drawing conclusions from cross-lagged relationships: re-considering the role of the time-interval. *Stuct. Equ. Modeling* 25, 809–823. doi: 10.1080/10705511.2018.1431046

Laucht, M., Esser, G., and Schmidt, M. H. (2000a). Externalisierende und internalisierende störungen in der kindheit: untersuchungen zur entwicklungspsychopathologie. *Z. Klin. Psychol. Psoychother.* 29, 284–292. doi: 10.1026//0084-5345.29.4.284

Laucht, M., Esser, G., and Schmidt, M. H. (2000b). Längsschnittforschung zur entwicklungsepidemiologie psychischer störungen: zielsetzung, konzeption und zentrale befunde der mannheimer risikokinderstudie. *Z. Klein. Psychol. Psychother.* 29, 246–262.

Leerkes, E. M., Blankson, N. A., and O'Brien, M. (2009). Differential effects of maternal sensitivity to infant distress and Non-distress on social-emotional functioning. *Child Dev.* 80, 762–775. doi: 10.1111/j.1467-8624.2009.01296.x

Marquand, A. F., Kia, S. M., Zabihi, M., Wolfers, T., Buitelaar, J. K., and Beckman, C. F. (2019). Conceptualizing mental disorders as deviations from normative functioning. *Mol. Psychiatry* 24, 1415–1424. doi: 10.1038/s41380-019-0441-1

Marquand, A. F., Rezek, I., Buitelaar, J., and Beckmann, C. F. (2016). Understanding heterogeneity in clinical cohorts using normative models: beyond case-control studies. *Biol. Psychiatry* 80, 552–561. doi: 10.1016/j.biopsych.2015.12.023

Martel, M. M., Pan, P. M., Hoffmann, M. S., Gadelha, A., do Rosario, M. C., Mari, J. J., et al. (2017). A general psychopathology factor (P factor) in children: Structural model analysis and external validation through familial risk and child global executive function. *J. Abnorm. Psychol.* 126, 137–148. doi: 10.1037/abn0000205

Martinussen, R., Hayden, J., Hogg-Johnson, S., and Tannock, R. (2005). A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *J. Am. Acad. Child Adolesc. Psychiatry* 44, 377–384.

Mauz, E., Lange, M., Houben, R., Hoffmann, R., Allen, J., Gößwald, A., et al. (2019). Cohort profile: KiGGS cohort longitudinal study on the health of children, adolescents and young adults in Germany. *Int. J. Epidemiol.* 49, 375k–375k.

McTeague, L. M., Lang, P. J., Laplante, M. C., Cuthbert, B. N., Shumen, J. R., and Bradley, M. M. (2010). Aversive imagery in posttraumatic stress disorder: trauma recurrence, comorbidity, and physiological reactivity. *Biol. Psychiatry* 67, 346–356. doi: 10.1016/j.biopsych.2009.08.023

McTeague, L. M., Lang, P. J., Laplante, M. C., Cuthbert, B. N., Strauss, C. C., and Bradley, M. M. (2009). Fearful imagery in social phobia: generalization, comorbidity, and physiological reactivity. *Biol. Psychiatry* 65, 374–382. doi: 10.1016/j.biopsych.2008.09.023

McTeague, L. M., Lang, P. J., Wangelin, B. C., Laplante, M. C., and Bradley, M. M. (2012). Defensive mobilization in specific phobia: fear specificity, negative affectivity, and diagnostic prominence. *Biol. Psychiatry* 72, 8–18. doi: 10.1016/j.biopsych.2012.01.012

Meaney, M. J. (2018). Perinatal maternal depressive symptoms as an issue for population health. *Am. J. Psychiatry* 175, 1084–1093.

Moffitt, T. E., Harrington, H., Caspi, A., Kim-Cohen, J., Goldberg, D., Gregory, A. M., et al. (2007). Depression and generalized anxiety disorder: cumulative and sequential comorbidity in a birth cohort followed prospectively to age 32 years. *Arch. Gen. Psychiatry* 64, 651–660. doi: 10.1001/archpsyc.64.6.651

Monroe, S. M., and Simons, A. D. (1991). Diathesis-stress theories in the context of life stress research: implications for the depressive disorders. *Psychol. Bull.* 110, 406–425. doi: 10.1037/0033-2909.110.3.406

Murray, A. L., Eisner, M., and Ribeaud, D. (2020). Within-person analysis of developmental cascades between externalising and internalising problems. *J. Child Psychol. Psychiatry* 61, 681–688. doi: 10.1111/jcpp.13150

Nelson, C. A., and Gabard-Durnam, L. J. (2020). Early adversity and critical periods: neurodevelopmental consequences of violating the expectable environment. *Trends Neurosci.* 43, 133–143. doi: 10.1016/j.tins.2020.01.002

Neumann, A., Pappa, I., Lahey, B. B., Verhulst, F. C., Medina-Gomez, C., Jaddoe, V. W., et al. (2016). Single nucleotide polymorphism heritability of a p factor in children. *J. Am. Acad. Child Adolesc. Psychiatry* 55, 1038–1045. doi: 10.1016/j.jaac.2016.09.498

Nigg, J. T. (2006). Temperament and developmental psychopathology. *J. Child Psychol. Psychiatry* 47, 395–422. doi: 10.1111/j.1469-7610.2006.01612.x

Nolvi, S., Rasmussen, J. M., Graham, A. M., Gilmore, J. H., Styner, M., Fair, D. A., et al. (2020). Neonatal brain volume as a marker of differential susceptibility to parenting quality and its association with neurodevelopment across early childhood. *Dev. Cogn. Neurosci.* 45:100826. doi: 10.1016/j.dcn.2020.100826

Nivard, M. G., Lubke, G. H., Dolan, C. V., Evans, D. M., St Porcain, B., Munafo, M. R., et al. (2017). Joint developmental trajectories of internalizing and externalizing disorders between childhood and adolescence. *Dev. Psychopathol.* 29, 919–928.

Paus, T. (2010). Population neuroscience: why and how. *Hum. Brain Mapp.* 31, 891–903. doi: 10.1002/hbm.21069

Pausova, Z., Paus, T., Abrahamowicz, M., Bernard, M., Gaudet, M., Leonard, G., et al. (2017). Cohotr profile: the Saguenay Youth Study (SYS). *Int. J. Epidemiol.* 46:e19.

Perra, O., Paine, A. L., and Hay, D. F. (2020). Continuity and change in anger and aggressiveness from infancy to childhood: The protective effects of positive parenting. *Dev. Psychopathol.* 1–20. doi: 10.1017/S0954579420000243

Pluess, M. (2015). Individual differences in environmental sensitivity. *Child Dev. Perspect.* 9, 138–143. doi: 10.1111/cdep.12120

Pluess, M., and Belsky, J. (2013). Vantage sensitivity: individual differences in response to positive experiences. *Psychol. Bull.* 139, 901–916.

Polanczyk, G., and Rohde, L. A. (2007). Epidemiology of attention-deficit/hyperactivity disorder across the lifespan. *Curr. Opin. Psychiatry* 20, 386–392.

Rademacher, L., Krach, S., Kohls, G., Irmak, A., Grunder, G., and Spreckelmeyer, K. N. (2010). Dissociation of neural networks for anticipation and consumption of monetary and social rewards. *Neuroimage* 49, 3276–3285. doi: 10.1016/j.neuroimage.2009.10.089

Rhee, S. H., Lahey, B. B., and Waldman, I. D. (2015). Comorbidity among dimensions of childhood psychopathology: converging evidence from behavior genetics. *Child Dev. Perspect.* 9, 26–31. doi: 10.1111/cdep.12102

Riglin, L., Collishaw, S., Thapar, A. K., Dalsgaard, S., Langley, K., Smith, G. D., et al. (2016). Association of genetic risk variants with attention-deficit/hyperactivity disorder trajectories in the general population. *JAMA Psychiatry* 73, 1285–1292.

Rohlfing, T., Cummins, K., Henthorn, T., Chu, W., and Nichols, B. N. (2014). N-CANDA data integration: anatomy of an asynchronous infrastructure for multi-site, multi-instrument longitudinal data capture. *J. Am. Med. Inform. Assoc.* 21, 758–762.

Ronald, A. (2011). Is the child 'father of the man'? evaluating the stability of genetic influences across development. *Dev. Sci.* 14, 1471–1478. doi: 10.1111/j.1467-7687.2011.01114.x

Rosenberg, M. D., Casey, B. J., and Holmes, A. J. (2018). Prediction complements explanation in understanding the developing brain. *Nat. Commun.* 9:589. doi: 10.1038/s41467-018-02887-9

Rutter, M., Pickles, A., Murray, R., and Eaves, L. (2001). Testing hypotheses on specific environmental causal effects on behavior. *Psychol. Bull.* 127, 291–324. doi: 10.1037/0033-2909.127.3.291

Schumann, G., Loth, E., Banaschewski, T., Barbot, A., Barker, G., Büchel, C., et al. (2010). The IMAGEN study: reinforcement-related behaviour in normal brain function and psychopathology. *Mol. Psychiatry* 15, 1128–1139.

Swartz, J. R., Williamson, D. E., and Hariri, A. R. (2015). Developmental change in amygdala reactivity during adolescence: effects of family history of depression and stressful life events. *Am. J. Psychiatry* 172, 276–283. doi: 10.1176/appi.ajp.2014.14020195

Taylor, J. A., Valentine, A. Z., Sellman, E., Bransby-Adams, K., Daley, D., and Sayal, K. (2015). A qualitative process evaluation of a randomised controlled trial of a parenting intervention in community (school) settings for children at risk of attention deficit hyperactivity disorder (ADHD). *BMC Psychiatry* 15:290. doi: 10.1186/s12888-015-0670-z

Thapar, A., and Riglin, L. (2020). The importance of a developmental perspective in Psychiatry: what do recent genetic-epidemiological findings show? *Mol. Psychiatry* 25, 1631–1639. doi: 10.1038/s41380-020-0648-1

Thomas, E., Buss, C., Rasmussen, J. M., Entringer, S., Ramireza, J. S. B., Marra, M., et al. (2019). Newborn amygdala connectivity and early emerging fear. *Dev. Cogn. Neurosci.* 37:100604.

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., and Gottesman, I. I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychol. Sci.* 14, 623–628. doi: 10.1046/j.0956-7976.2003.psci_1475.x

van Rooij, D., Hoekstra, P. J., Mennes, M., von Rhein, D., Thissen, A. J. A. M., Heslenfeld, D., et al. (2015). Distinguishing adolescents with ADHD from their unaffected siblings and healthy comparison subjects by neural activation patterns during response inhibition. *Am. J. Psychiatry* 172, 6 74–683.

Wang, F. L., Galán, C. A., Lemery-Chalfant, K., Wilson, M. N., and Shaw, D. S. (2020). Evidence for two genetically distinct pathways to co-occurring internalizing and externalizing problems in adolescence characterized by negative affectivity or behavioral inhibition. *J. Abnorm. Psychol.* 129, 633–645. doi: 10.1037/abn0000525

Whiteford, H. A., Degenhardt, L., Rehm, J., Baxter, A. J., Ferrari, A. J., Erskine, H. E., et al. (2013). Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet* 382, 1575–1586. doi: 10.1016/S0140-6736(13)61611-6

Wolfers, T., Doan, N. T., Kaufmann, T., Alnaes, D., Mobergt, T., Agartz, I., et al. (2018). Mapping the heterogeneous phenotype of schizophrenia and bipolar disorder using normative models. *JAMA Psychiatry* 75, 1146–1155. doi: 10.1001/jamapsychiatry.2018.2467

Zabihi, M., Oldehinkel, M., Wolfers, T., Frouin, V., Govard, D., Loth, E., et al. (2019). Dissecting the heterogeneous cortical anatomy of autism spectrum disorder using normative models. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 4, 567–578. doi: 10.1016/j.bpsc.2018.11.013

Zelazo, P. D., and Paus, T. (2010). Developmental social neuroscience: an introduction. *Soc. Neurosci.* 5, 417–421. doi: 10.1080/17470919.2010.510002

Conflict of Interest: TB served in an advisory or consultancy role for ADHS digital, Infectopharm, Lundbeck, Medice, Neurim Pharmaceuticals, Oberberg GmbH, Roche, and Takeda, received conference support or speaker's fee by Medice and Takeda, and received royalties from Hogrefe, Kohlhammer, CIP-Medien, and Oxford University Press. The present work is unrelated to the above grants and relationships.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Systems Neuroscience

REVIEW
published: 21 July 2021
doi: 10.3389/fnsys.2021.685957

Differences in Alzheimer's Disease and Related Dementias Pathology Among African American and Hispanic Women: A Qualitative Literature Review of Biomarker Studies

*Sarah K. Royse[1], Ann D. Cohen[2], Beth E. Snitz[3] and Caterina Rosano[1]**

[1] Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, United States, [2] Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA, United States, [3] Department of Neurology, University of Pittsburgh School of Medicine, Pittsburgh, PA, United States

Introduction: The population of older adults with Alzheimer's disease and Related Dementias (ADRD) is growing larger and more diverse. Prevalence of ADRD is higher in African American (AA) and Hispanic populations relative to non-Hispanic whites (nHW), with larger differences for women compared to men of the same race. Given the public health importance of this issue, we sought to determine if AA and Hispanic women exhibit worse ADRD pathology compared to men of the same race and nHW women. We hypothesized that such differences may explain the discrepancy in ADRD prevalence.

Methods: We evaluated 932 articles that measured at least one of the following biomarkers of ADRD pathology *in vivo* and/or post-mortem: beta-amyloid (Aß), tau, neurodegeneration, and cerebral small vessel disease (cSVD). Criteria for inclusion were: (1) mean age of participants >65 years; (2) inclusion of nHW participants and either AA or Hispanics or both; (3) direct comparison of ADRD pathology between racial groups.

Results: We included 26 articles (Aß = 9, tau = 6, neurodegeneration = 16, cSVD = 18), with seven including sex-by-race comparisons. Studies differed by sampling source (e.g., clinic or population), multivariable analytical approach (e.g., adjusted for risk factors for AD), and cognitive status of participants. Aß burden did not differ by race or sex. Tau differed by race (AA < nHW), and by sex (women > men). Both severity of neurodegeneration and cSVD differed by race (AA > nHW; Hispanics < nHW) and sex (women < men). Among the studies that tested sex-by-race interactions, results were not significant.

Conclusion: Few studies have examined the burden of ADRD pathology by both race and sex. The higher prevalence of ADRD in women compared to men of the same race may be due to both higher tau load and more vulnerability to cognitive decline in the presence of similar Aß and cSVD burden. AA women may also exhibit more neurodegeneration and cSVD relative to nHW populations. Studies suggest that

OPEN ACCESS

Edited by:
Tomáš Paus,
University of Toronto, Canada

Reviewed by:
Mihaly Hajos,
Yale University, United States
Michelle M. Mielke,
Mayo Clinic, United States

***Correspondence:**
Caterina Rosano
rosanoc@edc.pitt.edu

Received: *26 March 2021*
Accepted: *28 June 2021*
Published: *21 July 2021*

Citation:
Royse SK, Cohen AD, Snitz BE and Rosano C (2021) Differences in Alzheimer's Disease and Related Dementias Pathology Among African American and Hispanic Women: A Qualitative Literature Review of Biomarker Studies. Front. Syst. Neurosci. 15:685957. doi: 10.3389/fnsys.2021.685957

between-group differences in ADRD pathology are complex, but they are too sparse to completely explain why minority women have the highest ADRD prevalence. Future work should recruit diverse cohorts, compare ADRD biomarkers by both race and sex, and collect relevant risk factor and cognitive data.

Keywords: Alzheimer's disease, African American, Hispanic, sex differences, tau, neurodegeneration, cerebral small vessel disease, amyloid

INTRODUCTION

Alzheimer's disease and Related Dementias (ADRD) are progressive neurodegenerative illnesses clinically characterized by difficulties with memory and other cognitive abilities such as language and executive function (Matthews et al., 2019). Presently, 47 million people worldwide suffer from ADRD (Babulal et al., 2019). As the proportion of individuals older than 65 grows, this figure is estimated to increase by roughly 10 million new cases each year (Babulal et al., 2019). In addition to growing in number, the population of those older than 65 is expanding to include more racial minorities. In the United States, African American (AA) and Hispanic populations are projected to see steep increases in the number of people diagnosed with ADRD over the next 40 years (Matthews et al., 2019). Currently, compared to non-Hispanic white (nHW) populations, AA are 2 times more likely and Hispanics 1.5 times more likely to be clinically diagnosed with ADRD (Alzheimer's Association, 2019). When stratified by sex, women of these racial groups are at even higher risk of being diagnosed with ADRD compared to men (Matthews et al., 2019). In fact, there is emerging evidence that relative to all men and nHW women aged 65 and older, AA and Hispanic women have the first and second highest prevalence of ADRD, respectively (Matthews et al., 2019).

There are several race-associated factors that may explain these differences. AA and Hispanic populations are more likely than nHW to develop cardiometabolic diseases such as hypertension, diabetes, and obesity (Alzheimer's Association, 2019). Such risk factors are linked to higher risk of AD largely through their effects on cerebral vascular integrity. Exposure to vascular risk factors increases the likelihood of cerebral small vessel disease, which is associated with both onset and progression of ADRD (Babulal et al., 2019). Additionally, minorities are less likely to have access to high quality food or exercise-friendly neighborhoods (Zlokovic et al., 2020); both good nutrition and exercise are vital for maintaining brain health throughout the lifespan and may help curb effects of AD pathology on cognition (Zlokovic et al., 2020).

Because prevalence of AD is greater among all women compared to men independent of race (Matthews et al., 2019), any race-related factor may additionally interact with those that are specific to sex. One proposed biological mechanism of greater ADRD risk among women is menopausal-related changes in endogenous estrogen production. Estrogen is thought to be neuroprotective against pernicious effects of ADRD pathology, so its rapid decrease during menopause may precipitate the development of AD (Fisher et al., 2018). Others have hypothesized that the ε4 allele of the APOE gene interacts with sex to confer greater risk of AD among women relative to men (Nebel et al., 2018), but the mechanisms of such effect modification are poorly understood. In addition to biological risk factors, it should also be noted that higher prevalence rates of ADRD in women relative to men may be due in part to survivor bias. That is, women have longer life expectancies than men of their same race (Arias, 2021) and age is the greatest risk factor for AD (Jack et al., 2015). Thus, compared to men, women may spend more time at high risk of being diagnosed with ADRD coupled with longer time spent with the disease (Andersen et al., 1999).

Many risk factors for ADRD overlap in women and minorities. Compared to men and nHW populations, both women and minorities, respectively, are at greater risk of psychosocial risk factors such as depressive symptoms (Barnes and Bennett, 2014; Nebel et al., 2018). Importantly, those who experience depressive symptoms in midlife are more likely to develop AD, potentially due to shared neural substrates and pathways for memory and stress hormone dysregulation, respectively (Nebel et al., 2018). Further, both post-menopausal women and minorities are affected by poorer sleep quality, which emerging evidence suggests is related to ADRD pathology clearance (Grandner et al., 2016; Nebel et al., 2018; Irwin and Vitiello, 2019).

Given these race- and sex-related discrepancies for risk factors of ADRD, it follows that we would expect to also find race- and sex- related discrepancies in ADRD pathology.

The hallmark pathological features of AD are extracellular plaques made of beta-amyloid (Aß) and intracellular neurofibrillary tangles comprised of hyperphosphorylated tau (Jack et al., 2013). While the biological cascade of events is complex, it is generally accepted that these proteins begin to aggregate decades prior to symptomatology and that Aß precedes tau which precedes neurodegeneration (Jagust, 2018). In the beginning stages of AD, Aß and tau aggregate in the parietal cortex and medial temporal lobe, respectively (Jagust, 2018). Progressively, these abnormal proteins begin to deposit in other areas of the brain, ultimately yielding greater atrophy and resultant worsening clinical symptoms.

The advent of *in vivo* biomarkers has been integral in uncovering the underlying mechanisms of AD. Brain Aß load can be estimated through positive correlation with amyloid positron emission tomography (PET) radiotracer uptake or negative correlation with cerebral spinal fluid (CSF) Aß-42 markers. Tau burden in the brain is positively correlated with both PET radiotracer retention and CSF-derived total- and phosphorylated-tau. In addition to these, cerebral Aß and tau pathology can be measured through blood-based markers (e.g., Aß42/Aß40 ratio, total-tau, and phosphorylated-tau), but efforts

to improve the sensitivity of these biomarkers are ongoing (Zetterberg, 2019). Neurodegeneration is commonly estimated through magnetic resonance imaging (MRI). However, it can also be measured by CSF- or blood-derived neurofibrillary light (NfL) chain, which is correlated positively with axonal injury (Petzold, 2005; Olsson et al., 2016; Zetterberg, 2019). Of note, neuroimaging biomarkers like PET and MRI provide information regarding the severity of AD pathology as well as its topographical distribution. In contrast, CSF- and blood-based biomarkers reflect the severity of pathology, but do not provide insight into its topography.

Racially diverse and representative studies that examine ADRD pathology differences between racial groups are fairly limited and as a result, consensus surrounding race-related differences in Aß, tau, and neurodegeneration have not yet been reached (Babulal et al., 2019). In contrast, work examining sex-related differences in primarily nHW cohorts is more common. It has been reported that compared to men, women show similar levels of Aß (Ferretti et al., 2018), elevated tau (Buckley et al., 2019; Ossenkoppele et al., 2020), and less neurodegeneration (Jack et al., 2015; Sundermann et al., 2016; Ossenkoppele et al., 2020).

The overarching goal of this review was to examine sex-by-race differences in biomarkers of ADRD pathology. To do this, we conducted an Ovid MEDLINE search for studies that compare relevant AD biomarkers by race and sex-by-race. More specifically, we summarized race- and sex-by-race-related comparisons of Aß, tau, and neurodegeneration; this classification system is congruent with the National Institute on Aging and the Alzheimer's Association's proposed AT(N) framework (Jack et al., 2018), which categorizes research participants as neuropathologically normal or abnormal for A (Aß), T (tau), and N (neurodegeneration). In addition to these measures, we also included work related to cerebral small vessel disease (cSVD) as recent evidence suggests that vascular damage may affect other biological changes related to AD, including Aß clearance (Shaaban et al., 2017); further, cSVD is related to both race and sex, and thus may be important to understanding differing ADRD risk profiles among minority women (Jorgensen et al., 2018). In summary, based on existing literature, it appears that compared to nHW, minorities exhibit worse cSVD; compared to men, women present more tau. Thus, we hypothesized that a combination of more severe risk profiles for both cSVD and tau could explain why minority women have the highest prevalence of ADRD compared to the rest of the older adult population. Alternatively, it may be that pathology is similar in AA and Hispanic women compared to men of their same race and nHW populations, but that vulnerability to cognitive decline differs between groups.

MATERIALS AND METHODS

We used Ovid MEDLINE to retrieve articles for the narrative literature review through October 2020; line-by-line search terms are outlined in **Supplementary Table 1**.

Once we completed the search and removed duplicates, we screened articles for eligibility using titles and abstracts. During screening, we excluded articles for the following reasons: (1) the study was conducted in animals; (2) the study was conducted in a sample with a mean age less than 65; (3) the study did not include African Americans or Hispanics in the sample; (4) the study only conducted analyses adjusting for race; (5) the study sample included only one race; (6) the study did not include relevant biomarkers for AD; or (7) the study sample was comprised of individuals with other psychiatric or major illnesses or injuries.

Subsequent to title and abstract screening, we assessed full-text articles. We excluded articles if the study did not compare an AD biomarker by race ($n = 55$), if only one race was included in the sample ($n = 1$), or if the mean age of the sample was younger than 65 ($n = 2$).

A flow diagram outlining the search process is outlined in **Figure 1**, following PRISMA guidelines (Moher et al., 2009).

For each full-text article that we assessed, we noted the characteristics that are most likely to influence results: sampling source, multivariable analytical approach, and cognitive status of the participants.

Differing sampling sources may yield inconsistent results between two studies, particularly if one set of participants is recruited from an Alzheimer's Disease Research Center (ADRC) and the other from the community. Compared to the rest of the population, ADRC-recruited participants are often younger, more educated, and more likely to have at least one APOE-ε4 allele (Farrer et al., 1997; Snitz et al., 2018). In other words, participants recruited from an ADRC are not necessarily representative of the population; it follows that results identified in such recruitment samples may not be identical to those from a population-based sample.

Additionally, the multivariable analytical approach can affect results if investigators do not account for factors that influence ADRD pathology. For example, older age is a risk factor for neurodegeneration (Jack et al., 2015). Thus, an imbalance in chronological age between comparison groups may lead to spurious between-group differences of neurodegeneration. Further, other health conditions or diseases that are outside of the nervous system, but can affect the brain, like diabetes and hypertension (Alzheimer's Association, 2019), also necessitate consideration for statistical adjustment; more frequent occurrence in one of the comparison groups can again yield an apparent group difference in ADRD pathology.

Cognitive status of the participant is a proxy for ADRD pathology and as such, may lead to varying results between cross-sectional studies if the samples differ in this regard. That is, a sample of participants who are cognitively normal would likely only capture those with a low burden of ADRD pathology while that comprised of clinically-diagnosed AD patients would exhibit advanced stages of pathology. As such, these two cohorts, which represent different parts of the natural history of AD, are not comparable, and will likely not yield identical results.

RESULTS

In total, 26 articles were included for review. Of these, nine measured Aß, 6 measured tau, 16 measured neurodegeneration, and 18 measured cSVD. Seven articles additionally assessed

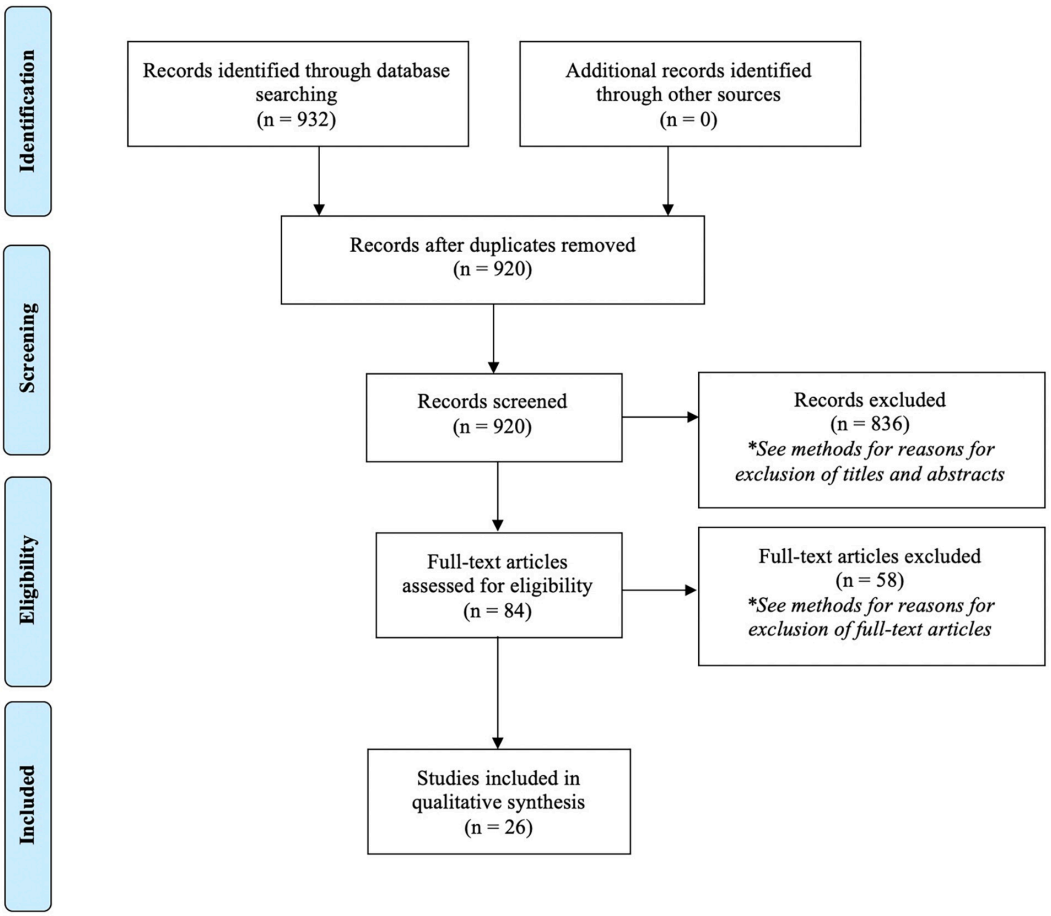

FIGURE 1 | PRISMA flow diagram delineating the search for literature completed through October 2020.

sex-by-race differences. In all studies, biomarker comparisons were made cross-sectionally. A summary of the main findings surrounding race differences in ADRD biomarkers is shown in **Figure 2**; race- and sex-by-race-specific findings are outlined in **Supplementary Table 2**.

Amyloid Biomarkers

Among nine studies examining amyloid biomarkers, three used CSF Aß-42 measurements (Howell et al., 2017; Garrett et al., 2019; Morris et al., 2019), six used PET imaging (Gu et al., 2015; Gottesman et al., 2016; Duara et al., 2019; Morris et al., 2019; Amariglio et al., 2020; Han et al., 2020), and one examined post-mortem tissue (Riudavets et al., 2006). Sources of recruitment included clinical and community settings; notable population-based samples included the Atherosclerosis Risk in Communities (ARIC) study, the Harvard Aging Brain Study (HABS), and the Washington Heights Inwood Columbia Aging Project (WHICAP). Participant cognitive status spanned normal cognition, mild cognitive impairment (MCI), and AD. Sample sizes ranged from 135 to 1255. Average ages ranged from 65 to 84 years. Two of these studies examined sex-by-race related differences, the findings of which are described in the section "Sex Differences in Amyloid Biomarkers by Race." In the following section, we outline race-related differences for both male and female participants combined.

Racial Differences in Amyloid Biomarkers

All (Riudavets et al., 2006; Gu et al., 2015; Howell et al., 2017; Duara et al., 2019; Garrett et al., 2019; Morris et al., 2019; Amariglio et al., 2020) but two (Gottesman et al., 2016; Han et al., 2020) studies reported no significant racial differences in Aß burden.

Of those that measured CSF Aß-42 concentrations, one community-based study of 1255 adults ranging from normal cognition, MCI, and AD reported no significant differences between AA and nHW after adjusting for age, sex, APOE-ε4 allele status, education, clinical status, body mass index (BMI), family history of AD, and CSF drift variables (Morris et al., 2019). Two other studies (Howell et al., 2017; Garrett et al., 2019), both of which recruited from ADRC and community, similarly reported no significant CSF Aß-42 differences between AA and nHW. One of these studies included participants with cognitive status ranging from normal cognition to AD (Howell et al., 2017) and the other including only those with normal

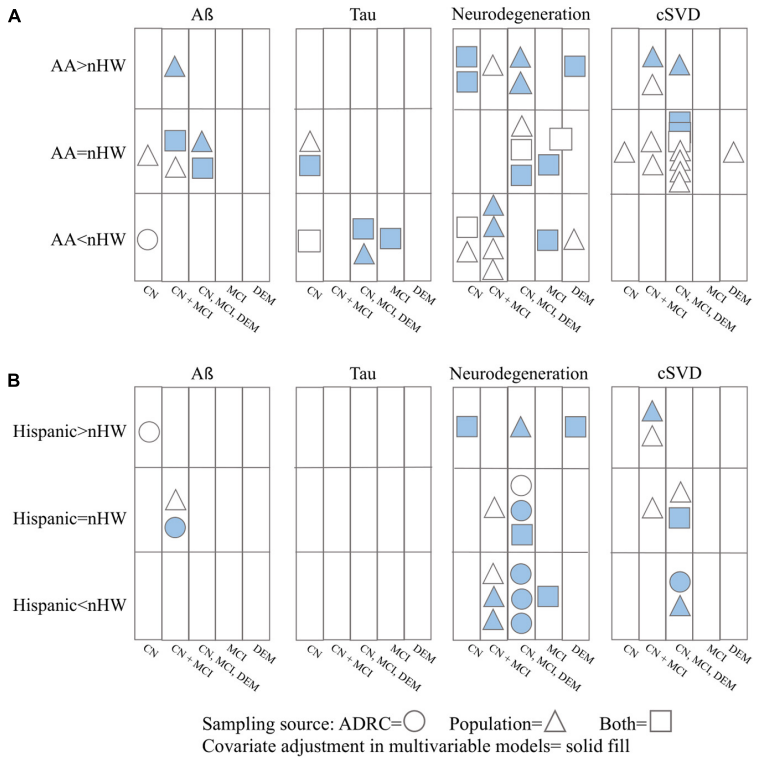

FIGURE 2 | Summary of results for ADRD pathology differences between nHW and **(A)** African Americans and **(B)** Hispanics. CN = cognitively normal; MCI = mild cognitive impairment; DEM = dementia. Studies that did not report cognitive status or recruitment source of their participants are not included (*n* = 6).

cognition and MCI (Garrett et al., 2019); results from both studies remained non-significant after stratifying by cognitive status and/or adjusting for covariates.

Brain PET studies also largely reported no significant racial differences in Aß radiotracer retention for either AA or Hispanic populations relative to nHW. Morris et al. (2019) reported global $^{[11-C]}$Pittsburgh Compound-B (PiB) standardized uptake value ratios (SUVR) that were not statistically different in AA compared to nHW in participants ranging in cognitive status (normal cognition, MCI, AD) and recruited from the community; this comparison was made after adjustment for age, sex, APOE-ε4 allele status, education, clinical status, BMI, and family history of AD. Similarly, work from the community-based HABS revealed that among 296 cognitively normal participants, AA did not statistically differ in average cortical PiB SUVR from nHW (Amariglio et al., 2020). One study that included 116 dementia-free AA, Hispanic, and nHW participants from WHICAP reported similarly null racial differences such that there were no significant differences in the proportion of participants categorized as visually $^{[18-F]}$Florbetaben (FBB) SUVR positive by race (Gu et al., 2015); 32% of AA, 31% of Hispanics, and 40% of nHW were FBB positive. In a study that recruited 159 Hispanic and nHW participants ranging in cognitive status (cognitively normal, MCI, dementia) from an ADRC, no significant racial differences were detected in FBB SUVR (Duara et al., 2019). Race additionally did not significantly predict FBB SUVR in multivariable linear regression models adjusted for age, Mini Mental State Examination (MMSE) score, and APOE-ε4 allele status.

One post-mortem study that did not report clinical status of participants detected no racial differences in Aß burden among 100 participants. Using this community-based sample that included AA and nHW brains, Riudavets et al. (2006) reported that in each racial sample, about two-thirds of participants exhibited Aß plaques. Further, logistic regression analyses revealed that race did not significantly predict Aß plaque or Aß angiopathy presence.

In contrast to these findings, the ARIC study found that among 329 participants without dementia, AA had 2-fold greater odds of global $^{[18-F]}$Florbetapir (FBP) SUVR positivity relative to those who were nHW (Gottesman et al., 2016). This relationship was not explained by age, sex, education, APOE-ε4 allele status, hypertension, diabetes, cognitive status, or white matter hyperintensity (WMH) volume. Similarly, an ADRC-based study of 85 participants with normal cognition reported that compared to nHW, a higher percentage of Hispanics, but not AA, met the threshold for FBP positivity (Han et al., 2020). Within the study sample, 50% of Hispanics, 30% of AA, and 44% of nHW participants were classified as FBP positive. However, no formal statistical test assessed if these group differences were significant.

Longitudinal analyses from HABS revealed that in those with elevated PiB SUVR at baseline, AA participants showed annual composite cognitive decline that was 0.05 standard deviations faster compared to nHWs (Amariglio et al., 2020).

Although small in absolute terms, this difference was statistically significant even after adjustment for age, sex, and baseline Preclinical Alzheimer's Cognitive Composite score. Additionally, investigators from WHICAP found that higher FBB SUVR was associated with small but significant yearly decline in average cognition, language, and executive function in AA, but not nHW participants (Gu et al., 2015).

Sex Differences in Amyloid Biomarkers by Race

Within these nine studies of Aß comparisons between racial groups, two also measured sex-by-race differences; women generally exhibited higher or similar burden of amyloid independent of race. The ARIC study found that in both AA and nHW participants, women were 1.7 times more likely than men to be globally FBP positive after adjustment for age, race, education, APOE-ε4 allele status, hypertension, and diabetes (Gottesman et al., 2016). Additionally, when stratified by race and sex, Riudavets et al. (2006) found that a greater proportion of AA and nHW women displayed Aß plaques relative to men of their respective race. However, sex was not a significant predictor of Aß plaques in linear models adjusted for age. Results for amyloid angiopathy were less straightforward. Among nHW participants, a greater proportion of women were dichotomously classified as exhibiting amyloid-angiopathy. However, among AA participants, the number of participants showing such pathology was roughly equal.

Summary and Synthesis

Overall, these studies do not provide strong evidence for racial differences in Aß deposition. While the ARIC study did report that AA had significantly greater Aß burden than nHW, this may have been the result of geographic exposures. Most of the AA participants in ARIC were from the Jackson, Mississippi site (Gottesman et al., 2016); FBP SUVRs in AA participants at the other two sites were 6% higher than all nHW participants, but 4% lower than those at the Jackson site. Thus, the true racial difference may be overestimated in this sample. These results may additionally be biased by the effects of smoothing the PET images to a common resolution. That is, amyloid was quantified using FBP, an Aß radiotracer with comparatively narrower dynamic range than others on account of its higher non-specific white matter retention (Wong et al., 2010; Landau et al., 2013). Others have reported that smoothing exacerbates the effect of white matter retention and further compresses the range of FBP (Landau et al., 2013). Assuming that AA populations exhibit worse neurodegeneration (and thus, less white matter) than nHW (see section "Neurodegeneration Biomarkers"), it follows that compared to nHW, AA participants' PET images may exhibit relatively less FBP white matter binding, positively skewing SUVRs. Similar to the ARIC study, Han et al. (2020) found greater prevalence of FBP positivity in Hispanic participants compared to nHW. However, because this difference was not formally tested for statistical significance, it should be interpreted cautiously.

Results from these studies similarly did not yield strong evidence for sex differences in Aß in either direction. The findings are in contrast to many recent studies that have largely found no differences between men and women in terms of Aß burden (Mielke et al., 2012; Altmann et al., 2014; Buckley et al., 2019). However, it should be noted that most studies that have measured sex differences in Aß have done so in largely white samples. As such, though few in number, these findings from racially diverse epidemiological cohorts indicate that sex differences in Aß among other races may differ from those identified in nHW participants.

Two articles included in this review found that compared to nHW, AA exhibit worse cognition in the presence of comparable Aß load (Gu et al., 2015; Amariglio et al., 2020). Before further discussion of these studies, it should be noted that results may be influenced by pathologies that, compared to Aß, are more strongly associated with cognitive impairment, but were not accounted for in analyses. Such pathologies include tau and Lewy bodies (Schneider et al., 2012; Hedden et al., 2013; LaPoint et al., 2017; Jansen et al., 2018; Maass et al., 2018), the latter of which may be more common in AA with AD (Barnes et al., 2015). Despite this limitation, evidence still suggests that in response to Aß, AA populations potentially possess relatively less cognitive resilience or worse cognitive adaptability to neuropathological insult (Arenaza-Urquijo and Vemuri, 2018). These studies did not report race-related differences in cognition separately for men and women, but instead combined data across sexes. However, previous work has found that compared to men, women exhibit worse cognitive resilience in the presence of similar burden of AD pathology, including Aß (Gamberger et al., 2017; Koran et al., 2017). Taken together with results from studies on racial differences, it is possible that minority women face a double burden for risk of accelerated cognitive decline in the presence of Aß.

Tau Biomarkers

A total of six studies measured tau, three of which did so with CSF (Howell et al., 2017; Garrett et al., 2019; Morris et al., 2019), one with blood (Rajan et al., 2020), one with PET imaging (Lee et al., 2018), and one with post-mortem tissue (Riudavets et al., 2006). Study samples were recruited from ADRC and population-based settings, the latter of which notably included The Chicago Health and Aging Project (CHAP) and HABS. Participant clinical status included normal cognition, MCI, and AD. Sample sizes ranged from 146 to 1327. The average age of cohorts ranged from 65 to 76 years. No studies included Hispanic participants. Sex-by-race differences were reported in only one of these studies, which is summarized in the section "Sex Differences in Tau Biomarkers by Race."

Racial Differences in Tau Biomarkers

Overall findings for racial differences in tau burden were mixed. Among the CSF-derived tau studies, one ADRC- and community-based sample of 362 participants reported racial differences in MCI, but not cognitively normal, participants (Garrett et al., 2019). Among MCI participants, AA displayed less total- and phosphorylated-tau on average relative to nHW independent of age, sex, education, family history of AD, BMI, Montreal Cognitive Assessment (MoCA), hypertension, diabetes, and income. However, in cognitively normal participants, AA and nHW did not differ in either total- or phosphorylated-tau

after covariate adjustment. In another ADRC- and community-based study comprised of 135 participants with either normal cognition, MCI, or dementia, investigators found racial differences for both total- and phosphorylated-tau such that AA exhibited less burden than nHW (Howell et al., 2017); this difference persisted after adjustment for cognitive function, age, sex, APOE-ε4 allele status, ABCA7 risk allele status, Aß-42, hypertension, diabetes, and WMH volume. A third study found racial differences in tau independent of cognitive status (Morris et al., 2019); in a cohort of 1255 cognitively normal, MCI, and AD participants, AA exhibited average total- and phosphorylated-tau concentrations that were less than nHW. This difference was not explained by age, sex, APOE-ε4 status, education, clinical status, family history of AD, BMI, or CSF drift variables.

In contrast to the above findings, two other studies reported no racial differences in tau load for either blood-based (Rajan et al., 2020) or PET-based (Lee et al., 2018) markers. CHAP, which included blood-based measurements in 1,327 participants with normal cognition, found no significant difference in total-tau between AA and nHW (Rajan et al., 2020). In a HABS subset of 146 participants whose cognitive status were not reported, investigators similarly did not find significant differences between AA and nHW participants in $^{[18-F]}$Flortaucipir (FTP) SUVR in the amygdala, entorhinal, fusiform, or inferior temporal regions-of-interest after adjustment for age, sex, education, and Aß burden (Lee et al., 2018). Of exception, AA participants showed greater FTP SUVR in the choroid plexus and hippocampus, but the authors note that this was likely due to off-target binding of the radiotracer and spill-in, respectively.

In a post-mortem study that did not report the overall difference between AA and nHW in tau lesions, AA race was not significantly associated with odds of tau lesions after adjusting for age in a sample of 100 participants (Riudavets et al., 2006). This study also did not report the cognitive status of participants.

Sex Differences in Tau Biomarkers by Race

Of the studies that measured tau differences between racial groups, one also measured and reported sex differences. In a post-mortem study that did not report the cognitive status of participants, Riudavets et al. (2006) reported greater severity of tau lesions in nHW women compared to men such that 96% of nHW women exhibited tau lesions whereas only 88% of men did so. In contrast, 96% of both AA women and men displayed tau lesions. When broken down by region, results differed slightly such that independent of race, a greater percentage of women showed more advanced tauopathy. Importantly, none of the comparisons made in this study were done so with formal statistical analysis. Thus, they should be interpreted cautiously.

Summary and Synthesis

Studies found that tau burden among AA populations was either less than or similar to that of nHW. These mixed findings between studies may be due in part to the age of participants. *In vivo* studies that reported less relative tau burden in AA included participants who were younger (average age range: 65–70 years) (Howell et al., 2017; Garrett et al., 2019; Morris et al., 2019) compared to those that reported no differences (average age range: 73–76 years) (Lee et al., 2018; Rajan et al., 2020). Thus, it may be that the analyses that found no racial differences included participants with more advanced tauopathy either due to age (Lowe et al., 2018) or AD progression. Such discrepancies between studies warrant that future work examine whether or not racial differences in tau pathology at the beginning of late adulthood and/or AD persist through the natural history of the disease. Taken together with the lack of such differences in the post-mortem study included in this review (Riudavets et al., 2006), it is possible that while younger AA populations exhibit lesser tau accumulation in relation to nHW, at some point during the course of AD, or as individuals age, the burden of tau converges to be comparable among the two racial groups. In other words, compared to nHW, in AA tau accumulation may begin later and accrue more rapidly. Findings from Howell et al. (2017) support this hypothesis; this group reported that racial differences were greatest among cognitively normal participants, but in those with MCI and AD, the gap in tau burden progressively decreased.

The sex differences by race in post-mortem tau burden in Riudavets et al. (2006) may also be influenced by age and/or disease progression. At the time of autopsy, on average, women were older compared to men of their respective races. Thus, it is reasonable that a greater proportion of women would show evidence of a more advanced disease stage. Women were much older than men particularly among the nHW participants, which may also have influenced the more pronounced sex-by-race differences in pathology among nHW participants, comparatively. These limitations notwithstanding, these sex differences are consistent with previous work conducted in nHW populations (Buckley et al., 2019; Ossenkoppele et al., 2020). Future work should apply either age stratification methods or more consistent age-adjusted analyses to further investigate sex differences in tau among AA populations.

Neurodegeneration Biomarkers

Among the 16 studies that examined neurodegeneration, 2 measured NfL (Howell et al., 2017; Rajan et al., 2020), 4 measured total brain volume (Minagar et al., 2000; Brickman et al., 2008; DeCarli et al., 2008; Aggarwal et al., 2010), 8 measured hippocampal and entorhinal volumes (DeCarli et al., 2008; Zahodne et al., 2015; Howell et al., 2017; Burke et al., 2018; Duara et al., 2019; Garrett et al., 2019; Morris et al., 2019; Arruda et al., 2020), 4 measured cortical thickness (Zahodne et al., 2015; McDonough, 2017; Rizvi et al., 2018; Arruda et al., 2020), and 4 measured lateral ventricle size (Minagar et al., 2000; Shadlen et al., 2006; Brickman et al., 2008; Burke et al., 2018). Participants in these studies ranged from cognitively normal, MCI, and AD and were recruited from both ADRC and community settings. Of those recruited from the community, notable study samples included The Cardiovascular Health Study (CHS), CHAP, HABS, and WHICAP. Study sample sizes ranged from 135 to 2786 and average ages ranged from 65 to 80 years. Among these, 2 examined sex-by-race-associated differences, which are described in the section "Sex Differences in Neurodegeneration Biomarkers by Race"; sections "Racial Differences in NfL Concentrations" through "Racial Differences in Lateral Ventricle

Size" report race-related differences for both male and female participants combined.

Racial Differences in NfL Concentrations

In an ADRC- and community-based sample comprised of 135 participants, investigators reported that CSF-derived NfL levels were lower in cognitively normal AA compared to nHW after adjustment for age, sex, APOE-ε4 allele status, ABCA7 risk allele status, Aß-42, hypertension, diabetes, and WMH volume (Howell et al., 2017). However, among those with cognitive impairment, NfL levels did not differ between races after covariate adjustment. Analyses from CHAP revealed in a sample 1,327 participants with normal cognition that blood plasma-derived concentrations of NfL were lower in AA relative to nHW across the entire cohort (Rajan et al., 2020). However, this between-group difference was not formally tested for statistical significance.

Racial Differences in Total Brain Volumes

A pattern emerged among studies that measured total brain volume such that investigators consistently reported no racial differences. Analyses from CHAP reported no significant difference in total brain volumes between AA and nHW participants in a sample of 575 participants who ranged from normal cognition to dementia (Aggarwal et al., 2010). DeCarli et al. (2008) similarly found that AA and Hispanic participants did not display significantly different average brain volumes compared to nHW in a sample of 401 participants with normal cognition, MCI, and AD; this was after adjustment for age, sex, education, and cognitive status. Finally, in a study that included 144 participants diagnosed with AD, Minagar et al. (2000) found no difference in visually-assessed cortical neurodegeneration between Hispanics and nHW after adjusting for age, gender, education, MMSE score, and disease duration.

Of exception, in a subset of WHICAP with 702 cognitively normal and MCI participants, Brickman et al. (2008) reported that AA and Hispanic participants both exhibited mean total brain volumes that were larger than that of nHW after adjusting for age, sex, and vascular disease history.

Racial Differences in Hippocampal and Entorhinal Volumes

Racial differences in hippocampal and entorhinal volumes among AA populations were mixed across six studies. One WHICAP subset that included 638 participants ranging from normal cognition to MCI found that AA displayed greater average hippocampal volume compared to nHW, with a small effect size (Cohen's *d*) of 0.23 (Zahodne et al., 2015). Additionally, DeCarli et al. (2008) reported that in an ADRC- and community-based cohort of 401 participants, among those with MCI, AA had larger hippocampal volumes after adjustment for age, sex, and education.

Three other studies reported smaller regional volumes in AA relative to nHW populations. Morris et al. (2019) reported that in a cohort of 1255 community-recruited participants with varying cognitive status (normal, MCI, and AD), AA displayed smaller average hippocampal volumes than that of nHW after covariate adjustment. However, this finding was influenced by family history of dementia. When stratified, AA participants with family history of dementia had smaller hippocampal gray matter volumes (GMV) than nHW with family history of dementia. In those without such family history, no racial differences in hippocampal GMV were detected. Similarly, DeCarli et al. (2008) reported that among those with normal cognition and dementia, AA had smaller hippocampal volumes; this could not be explained by adjustment for age, sex, and education. In another cohort of 362 participants recruited from both ADRC and community, cognitively normal AA participants' average hippocampal volumes were smaller than that of nHW independent of age, sex, educational level, BMI, family history of AD, MoCA score, hypertension, diabetes, and income (Garrett et al., 2019).

In contrast to those described above, two other studies found no significant differences in hippocampal volumes between AA and nHW participants. In one study of 135 participants ranging from cognitively normal to AD, Howell et al. (2017) found that race had no effect on hippocampal volume. Garrett et al. (2019) also reported that average hippocampal volumes in AA with MCI were not statistically different than nHW after adjusting for covariates.

Studies that included Hispanic populations also reported mixed findings. Four articles reported greater medial temporal regions in Hispanics compared to nHW. In a subset from WHICAP that only included participants with normal cognition and MCI, Hispanic participants displayed greater average hippocampal volumes relative to nHW with an effect size (Cohen's *d*) of 0.28 (Zahodne et al., 2015). Among those with MCI in a study of 401 participants that recruited from both the ADRC and community, Hispanic participants exhibited greater hippocampal volumes relative to nHW after adjusting for age, sex, and education (DeCarli et al., 2008). One ADRC-based study with 226 participants ranging from cognitively normal to AD found that compared to nHW, Hispanic participants exhibited greater average hippocampal volumes (Cohen's $d = 0.30$) and entorhinal volumes (Arruda et al., 2020). This group additionally reported that Hispanic race was associated with 0.12 mm^3 greater hippocampal volume and 0.13 mm^3 greater entorhinal volume independent of age, gender, functional activities questionnaire score, and MoCA score. In another ADRC-based study of 165 participants with normal cognition, MCI, and dementia, Hispanics showed greater hippocampal and entorhinal volumes relative to nHW after adjusting for age, depression, and Geriatric Depression Scale score (Burke et al., 2018).

Two articles reported patterns that were in contrast to the above findings. One ADRC-based study that included 159 participants with normal cognition, MCI, and AD found no significant difference in hippocampal volumes between Hispanic and nHW participants (Duara et al., 2019). Similarly, DeCarli et al. (2008) reported that among those with normal cognition or dementia, Hispanic participants displayed smaller hippocampal volumes relative to nHW after adjustment for age, sex, and education.

Racial Differences in Cortical Thickness

Findings generally suggested that compared to nHW populations, AA display smaller cortical thickness. In WHICAP, across 519 participants with normal cognition, MCI, or dementia, investigators found that after adjusting for age, education, and intracranial volume, AA exhibited smaller global cortical thickness compared to nHW participants (Rizvi et al., 2018). In another WHICAP subset that included 638 participants with normal cognition and MCI, AA displayed smaller AD-composite cortical thickness relative to nHW participants (Zahodne et al., 2015); this difference yielded an effect size (Cohen's *d*) of 0.23. An investigation of 284 HABS participants revealed a similar pattern such that after participants were propensity score-matched using age, sex, education, verbal IQ, Aß burden, and white matter hypointensity presence, bootstrap ratios found 12 of 20 AD-signature cortical thickness regions that were significantly smaller in AA participants (McDonough, 2017).

Results regarding cortical thickness in Hispanic populations relative to nHW were less consistent. Rizvi et al. (2018) reported that in a WHICAP cohort of participants ranging from normal cognition to dementia, relative to nHW, Hispanic participants displayed smaller average global cortical thickness after adjustment for age, education, and intracranial volume. However, in another WHICAP analysis that only included participants with normal cognition or MCI, AD-composite cortical thickness was almost identical in nHW and Hispanic participants (Zahodne et al., 2015). Similarly, in an ADRC-based study that included 226 cognitively normal, MCI, and AD participants, Hispanic participants did not statistically differ from nHW in average entorhinal cortex thickness (Arruda et al., 2020). This study additionally reported that in a linear regression model adjusted for age, sex, education, functional activity questionnaire score, and MoCA score, Hispanic ethnicity was not associated with entorhinal cortex thickness.

Racial Differences in Lateral Ventricle Size

Studies typically reported that compared to nHW, minorities displayed significantly smaller lateral ventricles. In a subset of 2786 participants from CHS, investigators found that in those without dementia, 8% of the AA sample and 15% of the nHW sample exhibited large ventricular volumes (Shadlen et al., 2006). This pattern was similar among those with dementia as 16% of AA and 32% of nHW exhibited large ventricular volumes. In the WHICAP cohort across 702 participants with normal cognition and MCI, compared to being nHW, AA and Hispanic race was associated with smaller lateral ventricle volumes independent of age, sex, and vascular disease history (Brickman et al., 2008). Additionally, in an ADRC-based study of 165 participants, Hispanics exhibited left and right ventricle volumes that were smaller in cognitively normal, MCI, and dementia participants, compared to nHW participants with corresponding diagnoses (Burke et al., 2018); this was after adjustment for age, education, and Geriatric Depression Scale score. Similarly, Minagar et al. (2000) reported in a study with 144 Hispanics and nHW diagnosed with AD that ventricular size was smaller in Hispanics relative to nHW after covariate adjustment.

Sex Differences in Neurodegeneration Biomarkers by Race

Of the studies that examined racial differences in neurodegeneration, two measured sex-by-race differences. These studies suggested a sex difference such that women generally displayed less neurodegeneration relative to their male counterparts of the same race. For example, in a subset of 702 cognitively normal and MCI participants in the WHICAP cohort, investigators found that when stratified by race, total brain volume was larger in women compared to men in nHW, AA, and Hispanics (Brickman et al., 2008). Despite this, a sex-by-ethnicity interaction for total brain volume was not significant. This group also reported that female sex was associated with smaller ventricles independent of race. Arruda et al. (2020) reported in an ADRC-based study among 226 Hispanic and nHW participants that female sex was associated with greater hippocampal volume independent of age, education, ethnicity, functional activities questionnaire scores, and MoCA scores. This relationship was found in both the entire cohort of cognitively normal, MCI, and AD participants as well as in a subsample limited to only those who were non-demented. However, this group also found that sex did not predict entorhinal cortex volume or thickness after covariate adjustment.

Summary and Synthesis

These studies generally reported similar amounts of global neurodegeneration in AA and Hispanic populations relative to nHW. This was consistent across studies that measured neurodegeneration using NfL concentrations (Howell et al., 2017) and total brain volumes (Minagar et al., 2000; DeCarli et al., 2008; Aggarwal et al., 2010). Still, two studies reported lesser NfL burden in AA participants relative to nHW (Howell et al., 2017; Rajan et al., 2020). However, because neither analysis adjusted for BMI, which is greater among AA populations compared to nHW (Hales et al., 2020) and negatively correlated with NfL concentration (Manouchehrinia et al., 2020), these estimates may be biased by such unmeasured confounding.

It should additionally be noted that one analysis from WHICAP found thinner global cortical thickness values in AA and Hispanic participants compared to nHW after controlling for age, education, and intracranial volume (Rizvi et al., 2018). Because this analysis was conducted in a large population-based study with higher MRI resolution and more precise segmentation methods relative to the studies that reported no racial differences in global neurodegeneration, it follows that this may be a more precise estimate of the true racial differences.

The discrepant findings between studies that measured neurodegeneration in regions characteristic of AD pathology in AA populations may be due to differences in study samples and methodology. For example, the articles that reported no differences in AD-specific regional volumes or cortical thicknesses recruited their samples at least in part from ADRCs (Howell et al., 2017; Garrett et al., 2019). Because participants recruited from ADRCs are generally healthier than the rest of the population, and thus, may exhibit better brain health, brain region size differences may have been skewed toward the null. Other investigators conducted their studies using

either lower resolution MRIs (Brickman et al., 2008; DeCarli et al., 2008) or without adjusting for confounders in their analyses (Shadlen et al., 2006; Zahodne et al., 2015). However, studies that recruited participants from the community, used high resolution MRI with precise segmentation methods, and adjusted for covariates reported smaller AD-specific brain regions among AA participants compared to nHW (McDonough, 2017; Morris et al., 2019). In consideration with results surrounding global neurodegeneration, this pattern suggests that differences in the size of brain regions between AA and nHW populations may be characteristic of those affected by AD. However, longitudinal neuroimaging studies should examine if such racial differences exist due to preclinical brain region sizes, differences in brain resilience (how the brain structurally or functionally copes with neuropathological insult (Arenaza-Urquijo and Vemuri, 2018)), or both.

In contrast to results examining regional brain sizes in AA populations, studies generally suggested that the brain regions that are affected by AD are larger or similarly-sized in Hispanics compared to nHW. These findings were consistent among the studies that did not adjust for covariates (Minagar et al., 2000; Brickman et al., 2008; Zahodne et al., 2015; Duara et al., 2019) and those that did (DeCarli et al., 2008; Burke et al., 2018; Arruda et al., 2020). Of the studies that reported confounder-adjusted results, participants ranged in cognitive status. Based on these results, it is possible that Hispanics exhibit larger preclinical medial temporal lobe sizes compared to nHW, which contribute to racial differences in patterns of neurodegeneration across the natural history of AD. However, race-specific brain resilience mechanisms should additionally be examined.

Findings related to sex differences indicated that women exhibit less neurodegeneration compared to men independent of race. This is consistent with previous results conducted in largely white samples (Jack et al., 2015; Sundermann et al., 2016; Ossenkoppele et al., 2020). This sex-specific finding was most consistent among Hispanic women. Taken together with results surrounding racial differences in neurodegeneration, it is possible that compared to their aging counterparts, Hispanic women exhibit more brain resilience to AD pathology. This may partially explain the higher prevalence of AD among Hispanic women relative to nHW. Hispanic women have higher life expectancies than nHW women (Arias et al., 2015; Arias, 2016) and the greatest risk factor for AD is advanced age (Jack et al., 2015). As such, it may be that the brains of Hispanic women initially fare better in the presence of AD-related neuropathological insults relative to nHW women, but as a result, Hispanic women live longer to experience more time at high risk of developing AD.

Cerebral Small Vessel Disease Biomarkers

Of the 18 studies that examined cerebral small vessel disease, 12 measured white matter hyperintensities (Minagar et al., 2000; Shadlen et al., 2006; Brickman et al., 2008; DeCarli et al., 2008; Aggarwal et al., 2010; Zahodne et al., 2015; Gottesman et al., 2016; Howell et al., 2017; Burke et al., 2018; Della-Morte et al., 2018; Rizvi et al., 2018; Amariglio et al., 2020) and eight measured ischemic lesions and other infarcts (Riudavets et al., 2006; DeCarli et al., 2008; Wright et al., 2008; Aggarwal et al., 2010; Wiegman et al., 2014; Zahodne et al., 2015; Morris et al., 2019). These studies recruited cognitively normal, MCI, and dementia participants from both ADRC and community. Notable population-based study samples included The Northern Manhattan Study (NOMAS), ARIC, CHAP, CHS, HABS, and WHICAP. Sample sizes ranged from 135 to 2786 and average ages of cohorts ranged from 70 to 80 years. Two of these studies examined sex-related differences, and are reported in the section "Sex Differences in cSVD Biomarkers by Race;" sections "Racial Differences in White Matter Hyperintensities" through "Racial Differences in Ischemic Lesions and Other Infarcts" report race-related differences for both male and female participants combined.

Racial Differences in White Matter Hyperintensities

These studies generally reported that relative to nHW, AA populations display significantly greater WMH burden whereas Hispanic populations display significantly less WMH burden.

Of the studies included, four reported significantly higher WMH volumes in AA relative to nHW populations, three of which were conducted in subsets of WHICAP. Among one such WHICAP subset of 519 participants who ranged from cognitively normal to AD, racial differences were detected between AA and nHW participants such that relative to nHW and Hispanic participants, AA showed greater average WMH volumes throughout the brain after controlling for age, education, and intracranial volume (Rizvi et al., 2018). AA additionally exhibited relatively greater average WMH volumes in the frontal, temporal, parietal, and occipital regions after covariate adjustment. In another WHICAP analysis that included 638 participants with either normal cognition or MCI, AA displayed mean WMH volume that was larger than that of nHW (Zahodne et al., 2015); the effect size (Cohen's *d*) of this differences was 0.49. Finally, in the third WHICAP subset, among 702 participants with normal cognition and MCI, relative to nHW race, being AA was associated with greater WMH volumes independent of age, sex, and vascular disease (Brickman et al., 2008). In addition to these findings from WHICAP, another epidemiological study, NOMAS, reported that compared to nHW, mean WMH volume in AA participants was larger after adjustment for age, sex, health behaviors, BMI, and vascular risk factors (Della-Morte et al., 2018). This cohort was comprised of 1229 older adults whose cognitive status were not reported.

In contrast, six other studies reported no racial differences in WMH between AA and nHW populations. A study from ARIC found that in a subset of 329 dementia-free participants, AA exhibited average WMH volumes that were not statistically different compared to nHW (Gottesman et al., 2016). Additionally, DeCarli et al. (2008) found in an ADRC- and community-based sample of 401 cognitively normal, MCI, and dementia participants that AA or Hispanic race did not significantly predict WMH volume in a model that also included age, gender, education, and cognitive status. Similarly, in a subsample of 296 cognitively normal HABS participants, AA had average WMH volume that was not statistically different from

nHW (Amariglio et al., 2020). Another ADRC- and community-based study of 135 cognitively normal, MCI, an dementia participants found no statistically significant differences between AA and nHW WMH volumes (Howell et al., 2017). Among 575 CHAP participants, AA and nHW participants did not significantly differ in WMH volume (Aggarwal et al., 2010). Finally, using data from 2786 CHS participants, one study found that among those with dementia, the same proportion of AA and nHW participants had high WMH grades (Shadlen et al., 2006), but this was not tested for statistical significance.

In studies that measured WMH in Hispanic populations, three reported lesser burden among Hispanics compared to nHW. One analysis using data from 519 WHICAP participants found that in those ranging from cognitively normal to AD, Hispanics exhibited smaller total WMH volume compared to nHW and AA populations (Rizvi et al., 2018); this finding was independent of age and education. This study also found that after covariate adjustment, Hispanics displayed relatively smaller average WMH volumes for the frontal, temporal, and occipital regions. Burke et al. (2018) found in an ADRC-based cohort of 165 participants ranging from cognitively normal to dementia that race was significantly associated with visually-rated WMH volume such that Hispanics exhibited smaller WMH volumes relative to nHW. This was not explained by age, education, or geriatric depression score. This group additionally reported that after covariate adjustment, cognitively normal Hispanic participants displayed smaller WMH volumes relative to nHW; the magnitude of this difference increased with worsening cognition. Additionally, among 1229 older adults whose cognitive status were not reported in NOMAS, investigators reported that compared to nHW, mean WMH volume in Hispanic participants was larger after adjustment for age, sex, health behaviors, BMI, and vascular risk factors (Della-Morte et al., 2018).

In contrast to these findings, two studies that included cognitively normal and MCI participants from WHICAP reported greater WMH volumes among Hispanics relative to nHW. One of these subsets, which included 638 participants, reported that Hispanics displayed mean WMH volume that larger than that of nHW, with a detected effect size (Cohen's *d*) of 0.28 (Zahodne et al., 2015). The other subset reported greater WMH volume among Hispanics relative to nHW independent of age, sex, and vascular disease (Brickman et al., 2008).

Finally, one group reported no difference in visually-assessed WMH burden between Hispanics and nHW in a sample of 144 participants with AD reported (Minagar et al., 2000). Further, Hispanic race did not predict WMH burden after controlling for age, education, gender, MMSE score, and AD duration.

Racial Differences in Ischemic Lesions and Other Infarcts

Studies of comparisons by race for ischemic lesions and infarcts consistently reported a lack of statistically significant differences between nHW and either AA or Hispanic participants.

In studies that measured such biomarkers among AA and nHW populations, four found no racial differences. Among 575 CHAP participants who were classified as either having or not having dementia, no significant difference was detected between AA and nHW participants exhibiting more than one infarct (Aggarwal et al., 2010). Another community-based sample with 1255 participants ranging from cognitively normal to AD found that the proportion of AA participants who displayed lesions was not statistically different than that of nHW (Morris et al., 2019). DeCarli et al. (2008) found in an ADRC- and community-based cohort of 401 participants that subcortical infarcts did not vary significantly between AA and nHW after adjustment for age, gender, education, and cognitive status. Additionally, in a subset of 638 WHICAP participants that only included those with normal cognition and MCI, the proportion of AA participants with infarcts did not differ significantly than that of nHW (Zahodne et al., 2015). In another WHICAP subset of 243 participants ranging from cognitively normal to dementia, Wiegman et al. (2014) reported that the same proportion of AA and nHW participants exhibited at least one microbleed, though this was not formally tested for statistical significance.

Of exception, two studies reported significant differences between AA and nHW in ischemic lesion burden. In both of these studies, prevalence of infarcts in AA participants was greater than that of nHW participants (Wright et al., 2008; Qiao et al., 2016). Analyses from NOMAS found that among 656 participants, prevalence of at least one lesion was higher in AA participants (22%) compared to nHW participants (14%) (Wright et al., 2008). In a subsample of 1755 ARIC participants, investigators reported that 41% of AA participants and 32% of nHW participants exhibited at least one infarct (Qiao et al., 2016). Neither of these studies reported participant cognitive status.

In terms of infarct or lesion comparisons between Hispanics and nHW, no study reported statistically significant differences (DeCarli et al., 2008; Wiegman et al., 2014; Zahodne et al., 2015). One study from WHICAP that restricted to 638 participants with normal cognition or MCI, reported that Hispanics and nHW did not differ in the proportion of participants with infarcts (Zahodne et al., 2015). DeCarli et al. (2008) also found that subcortical infarcts did not vary significantly between nHW and Hispanics in a community- and ADRC-based study after adjusting for age, gender, education, and clinical diagnosis. In a another subset of WHICAP with 243 participants ranging from normal cognition to dementia, Wiegman et al. (2014) reported that relative to nHW, the proportion of Hispanics who had at least one microbleed did not significantly differ.

Studies included this review found that both AA and Hispanic race modified the relationship of cSVD on cognition such that minorities exhibited worse cognition than nHW at similar levels of burden. For example, one community- and ADRC-based study revealed an interaction between race and WMH volume such that WMH was associated with greater cognitive impairment in AA compared to nHW participants (Howell et al., 2017). Additionally, in a subset of CHS participants, Shadlen et al. (2006) found that infarcts were more strongly associated with dementia in AA participants compared to nHW. In contrast, analyses from CHAP revealed that the rate of cognitive decline with increasing WMH volume did not differ between nHW and AA (Aggarwal et al., 2010). However, nHW participants consistently displayed better global cognition compared to AA. Finally, one ADRC-based study

found that cognitive impairment in Hispanics was worse at similar levels of visually-assessed WMH burden compared to nHW (Burke et al., 2018).

Sex Differences in cSVD Biomarkers by Race

Two studies assessed measurements of cSVD by sex and race. In a subsample of 702 WHICAP participants, among those with normal cognition and MCI, sex was not associated with WMH volume (Brickman et al., 2008). Additionally, there were no significant sex-by-race interactions for WMH. The other study that reported sex-by-race differences, a post-mortem analysis that did not report participants cognitive status, found varying sex differences by race as related to brain infarcts and lacunes (Riudavets et al., 2006). Among AA participants, 26% of women and 16% of men exhibited infarcts or lacunes. However, in nHW participants, 16% of women and 26% of men exhibited such lesions; this was not formally tested for statistical significance.

Summary and Synthesis

Before discussing results surrounding race-related differences in cSVD burden, it should be noted that the majority of comparisons for these biomarkers were conducted without covariate adjustment (**Figure 2**). Further, of those that did adjust for confounders, many did not control for cardiovascular risk factors (DeCarli et al., 2008; Burke et al., 2018; Rizvi et al., 2018). In addition to being risk factors for cSVD (Babulal et al., 2019), incidence and prevalence rates of cardiovascular diseases differ between races (Alzheimer's Association, 2019) and sexes (Mozaffarian et al., 2015). Thus, some estimates of race and sex-by-race-related differences of cSVD may be biased due to unmeasured confounding.

Overall, findings for group differences in WMH between AA and nHW participants were mixed. Notably, the studies that reported no significant racial differences did not adjust for other covariates (Aggarwal et al., 2010; Gottesman et al., 2016; Howell et al., 2017; Amariglio et al., 2020). In contrast, those that reported relatively greater WMH in AA populations accounted for confounders in their analyses (Brickman et al., 2008; Della-Morte et al., 2018; Rizvi et al., 2018); these studies are further strengthened by having been conducted in large epidemiological cohorts. However, because they only represent two cohorts, both from Manhattan, New York, the findings may not generalize to other populations. Thus, the true magnitude of racial differences is likely not captured by these studies.

Results related to WMH burden in Hispanic populations consistently suggested that Hispanics exhibit lower WMH than nHW. While three studies found either no differences or greater burden in Hispanics, these findings may be influenced by lack of covariate adjustment (Zahodne et al., 2015) or, as the authors note in Minagar et al. (2000), the use of semiquantitative assessments with limited sensitivity.

In terms of comparisons between AA and nHW, only two studies that measured infarcts or lesions reported significant findings when comparing AA and nHW participants; in both studies, AA showed more frequent lesions and infarcts compared to nHW (Wright et al., 2008; Qiao et al., 2016). These findings are consistent with those that found greater WMH burden in AA populations relative to nHW and are further strengthened by having been uncovered in two large epidemiological studies, NOMAS and ARIC. In contrast, no studies that compared lesions or infarcts between Hispanics and nHW reported significant differences.

Previous work has found that female sex is associated with greater presence of lacunes and progression of WMH (van Dijk et al., 2008). This was not reflected in the WHICAP cohort, which found that women and men exhibited similar average volume of WMH (Brickman et al., 2008). In contrast, Riudavets et al. (2006) did report greater prevalence of infarcts and lacunes in women, but only in the AA populations. However, this comparison was not formally tested for significance, and thus, should be interpreted cautiously.

Several studies additionally reported a differential association of cSVD and cognition based on race; AA and Hispanic populations both exhibited worse cognitive function at similar burden of cSVD, indicating worse cognitive resilience to such pathology. These findings are particularly relevant for minority women as previous work has suggested that progression of WMH in women is faster than that in men (van Dijk et al., 2008). Thus, AA and Hispanic women may face higher risk of quickly-progressing cSVD coupled with relatively worse cognitive outcomes in its presence.

DISCUSSION

There is currently a limited number of studies that have examined ADRD pathology between races; there are even fewer that have done so by both race and sex. Despite this, we found evidence that greater prevalence of clinical ADRD among AA populations may be driven in part by more severe AD region-specific neurodegeneration and cSVD compared to other races. These findings were consistent in population-based studies among participants with varying cognitive status. However, we also found that both Hispanic populations and women show less severe pathology relative to nHW populations and men, respectively; both Hispanics and women exhibit relatively less neurodegeneration in regions affected by AD and cSVD. Thus, differences in neurodegeneration and cSVD on their own likely do not capture the full picture of sex-by-race differences in ADRD.

We did not find that minorities exhibited worse risk profiles for amyloid and tau pathology relative to nHW. Sex-by-race differences in Aß were additionally inconsistent. Although weak, there was some evidence for sex differences in tau severity such that a greater proportion of AA and nHW women exhibited more advanced tauopathy compared to men of their same race. This finding is consistent with results from largely nHW cohorts (Buckley et al., 2019; Ossenkoppele et al., 2020), implying that female sex may be associated

with greater tau burden independent of race. Given that tau is neurodegenerative (Spillantini and Goedert, 2013), it follows that women should experience more atrophy in regions affected by such protein aggregation. However, as noted above, women generally show less AD-related neurodegeneration than men. Thus, it may be that independent of race, women exhibit better brain resilience than men in the face of tau accumulation.

Minorities also consistently showed worse cognition relative to nHW in the presence of comparable ADRD pathology, including Aß and cSVD. While the studies that reported these race-related differences did not additionally stratify by sex, previous work has found a similar pattern among women. At comparable levels of Aß deposition, women exhibit worse cognition relative to men (Gamberger et al., 2017; Koran et al., 2017). Others have found that among those with cSVD, women decline more quickly than men (van Dijk et al., 2008). Considering both lines of research, these findings suggest that minority women may be at especially high risk of cognitive decline in the presence of more than one neuropathological insult related to ADRD. It should be noted that while cognitive decline is highly correlated with cSVD in older adult populations (Prins and Scheltens, 2015), it is generally not strongly associated with Aß aggregation directly. Rather, severity of clinical AD symptomatology correlates more strongly with neurodegeneration caused by tau accumulation (Hedden et al., 2013; LaPoint et al., 2017; Jansen et al., 2018; Maass et al., 2018). Thus, it is possible that despite potentially exhibiting more tau and less neurodegeneration than men of their same race (e.g., more brain resilience), minority women are more cognitively vulnerable to neuropathological insult compared to both men and nHW populations (e.g., worse cognitive resilience), which may partially explain their relatively higher prevalence of ADRD.

There is a growing body of evidence that suggests pathological and clinical presentation of ADRD differ between men and women (Jack et al., 2015; Sundermann et al., 2016; Ferretti et al., 2018; Buckley et al., 2019; Ossenkoppele et al., 2020). Results of this review indicate that race may additionally alter these sex differences, potentially supporting emerging evidence of greater ADRD prevalence among AA and Hispanic women compared to other older adults (Matthews et al., 2019). Whether race functions additively or multiplicatively with sex on ADRD pathology, and further, downstream symptomatology, is currently unclear. This gap in knowledge is due to the small number of studies that have examined ADRD biomarkers by both race and sex. As such, sites with existing biomarker and cognitive outcomes in racially diverse cohorts should consider conducting additional analyses on their data by (1) testing sex-by-race interactions on pathology and cognition or (2) stratifying by both sex and race and making biomarker and cognition comparisons accordingly.

In terms of future data collection, investigators should aim to recruit more racially diverse cohorts from which ADRD biomarkers, cognition, and risk factors can be measured, and sex-by-race analyses can be conducted; relating biomarker and cognitive outcomes to risk factors may help identify differential ADRD mechanisms in minority women. Additionally, recruiting minority participants who are representative of the target population will increase the generalizability of study results. There is also a need for serial biomarker measurements in these populations as all studies included in this review were cross-sectional. As such, they provide limited insight into the natural history of ADRD in AA and Hispanic populations. Still, the data examined herein suggests potentially altered trajectories of ADRD progression in minorities and further, minority women. Finally, it should be noted that investigators have reported difficulty with recruiting minorities for neuroimaging and CSF studies (Morris et al., 2019; Amariglio et al., 2020). Thus, designing studies around emerging blood-based biomarkers, which are less invasive, may circumvent this obstacle, allowing for more generalizable samples and better retention rates in longitudinal studies. Such steps to better characterize ADRD pathology and its progression in minority women, especially in the context of the AT(N) framework, should inform better diagnostic and therapeutic techniques, ultimately benefiting those potentially most at risk.

In summary, through this review, we identified that women may exhibit more tau, but less neurodegeneration than men, independent of race. We additionally found that women are more likely to show relatively worse cognition at similar levels of pathology load. Thus, women of all races may have lower cognitive resilience to ADRD neuropathological insult compared to men, despite also possessing higher brain resilience. While race likely alters these sex differences, the specific mechanism and magnitude of this effect is currently unknown. Future studies should aim to fill this gap in knowledge by recruiting more diverse and representative cohorts, comparing ADRD biomarker severity by both race and sex, and collecting relevant risk factor and cognitive data.

AUTHOR CONTRIBUTIONS

SR and CR conceived and designed the review and drafted the manuscript. AC and BS reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We would like to thank Helena VonVille (Graduate School of Public Health Librarian Liaison, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, United States) for her guidance regarding the Ovid MEDLINE search.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnsys.2021.685957/full#supplementary-material

REFERENCES

Aggarwal, N. T., Wilson, R. S., Bienias, J. L., De Jager, P. L., Bennett, D. A., Evans, D. A., et al. (2010). The association of magnetic resonance imaging measures with cognitive function in a biracial population sample. *Arch. Neurol.* 67, 475–482. doi: 10.1001/archneurol.2010.42

Altmann, A., Tian, L., Henderson, V. W., Greicius, M. D., and Alzheimer's Disease Neuroimaging Initiative Investigators (2014). Sex modifies the APOE-related risk of developing Alzheimer disease. *Ann. Neurol.* 75, 563–573. doi: 10.1002/ana.24135

Alzheimer's Association (2019). 2019 Alzheimer's disease facts and figures. *Alzheimers Dement.* 15, 321–387. doi: 10.1016/j.jalz.2019.01.010

Amariglio, R. E., Buckley, R. F., Rabin, J. S., Papp, K. V., Quiroz, Y. T., Mormino, E. C., et al. (2020). Examining cognitive decline across black and white participants in the harvard aging brain study. *J. Alzheimers Dis.* 75, 1437–1446. doi: 10.3233/JAD-191291

Andersen, K., Launer, L. J., Dewey, M. E., Letenneur, L., Ott, A., Copeland, J. R. M., et al. (1999). Gender differences in the incidence of AD and vascular dementia: the EURODEM studies. *Neurology* 53, 1992–1992. doi: 10.1212/WNL.53.9.1992

Arenaza-Urquijo, E. M., and Vemuri, P. (2018). Resistance vs resilience to Alzheimer disease: clarifying terminology for preclinical studies. *Neurology* 90, 695–703. doi: 10.1212/WNL.0000000000005303

Arias, E. (2016). *Changes in Life Expectancy by Race and Hispanic Origin in the United States, 2013–2014* (NCHS Data Brief No. 244). Hyattsville, MD: National Center for Health Statistics.

Arias, E. (2021). *Provisional Life Expectancy Estimates for January through June, 2020*. Hyattsville, MD: Centersfor Disease Control and Prevention. doi: 10.15620/cdc:100392

Arias, E., Kochanek, K., and Anderson, R. (2015). *How Does Cause of Death Contribute to the Hispanic Mortality Advantage in the United States?* (NCHS Data Brief No. 221). Hyattsville, MD: National Center for Health Statistics.

Arruda, F., Rosselli, M., Greig, M. T., Loewenstein, D. A., Lang, M., Torres, V. L., et al. (2020). The association between functional assessment and structural brain biomarkers in an ethnically diverse sample with normal cognition, mild cognitive impairment, or dementia. *Arch. Clin. Neuropsychol.* 36, 51–61. doi: 10.1093/arclin/acaa065

Babulal, G. M., Quiroz, Y. T., Albensi, B. C., Arenaza-Urquijo, E., Astell, A. J., Babiloni, C., et al. (2019). Perspectives on ethnic and racial disparities in Alzheimer's disease and related dementias: update and areas of immediate need. *Alzheimers Dement.* 15, 292–312. doi: 10.1016/j.jalz.2018.09.009

Barnes, L. L., and Bennett, D. A. (2014). Alzheimer's disease in African Americans: risk factors and challenges for the future. *Health Aff.* 33, 580–586. doi: 10.1377/hlthaff.2013.1353

Barnes, L. L., Leurgans, S., Aggarwal, N. T., Shah, R. C., Arvanitakis, Z., James, B. D., et al. (2015). Mixed pathology is more likely in black than white decedents with Alzheimer dementia. *Neurology* 85, 528–534. doi: 10.1212/WNL.0000000000001834

Brickman, A. M., Schupf, N., Manly, J. J., Luchsinger, J. A., Andrews, H., Tang, M. X., et al. (2008). Brain morphology in older African Americans, Caribbean Hispanics, and whites from northern Manhattan. *Arch. Neurol.* 65, 1053–1061. doi: 10.1001/archneur.65.8.1053

Buckley, R. F., Mormino, E. C., Rabin, J. S., Hohman, T. J., Landau, S., Hanseeuw, B. J., et al. (2019). Sex differences in the association of global amyloid and regional tau deposition measured by positron emission tomography in clinically normal older adults. *JAMA Neurol.* 76, 542–551. doi: 10.1001/jamaneurol.2018.4693

Burke, S. L., Rodriguez, M. J., Barker, W., Greig-Custo, M. T., Rosselli, M., Loewenstein, D. A., et al. (2018). Relationship between cognitive performance and measures of neurodegeneration among hispanic and white non-hispanic individuals with normal cognition, mild cognitive impairment, and dementia. *J. Int. Neuropsychol. Soc.* 24, 176–187. doi: 10.1017/S1355617717000820

DeCarli, C., Reed, B. R., Jagust, W., Martinez, O., Ortega, M., and Mungas, D. (2008). Brain behavior relationships among African Americans, whites, and Hispanics. *Alzheimer Dis. Assoc. Disord.* 22, 382–391. doi: 10.1097/wad.0b013e318185e7fe

Della-Morte, D., Dong, C., Markert, M. S., Elkind, M. S. V., Sacco, R. L., Wright, C. B., et al. (2018). Carotid intima-media thickness is associated with white matter hyperintensities: the Northern Manhattan study. *Stroke* 49, 304–311. doi: 10.1161/STROKEAHA.117.018943

Duara, R., Loewenstein, D. A., Lizarraga, G., Adjouadi, M., Barker, W. W., Greig-Custo, M. T., et al. (2019). Effect of age, ethnicity, sex, cognitive status and APOE genotype on amyloid load and the threshold for amyloid positivity. *NeuroImage* 22:101800. doi: 10.1016/j.nicl.2019.101800

Farrer, L. A., Cupples, L. A., Haines, J. L., Hyman, B., Kukull, W. A., Mayeux, R., et al. (1997). Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. *J. Am. Med. Assoc.* 278:1349. doi: 10.1001/jama.1997.03550160069041

Ferretti, M. T., Iulita, M. F., Cavedo, E., Chiesa, P. A., Schumacher Dimech, A., Santuccione Chadha, A., et al. (2018). Sex differences in Alzheimer disease – the gateway to precision medicine. *Nat. Rev. Neurol.* 14, 457–469. doi: 10.1038/s41582-018-0032-9

Fisher, D. W., Bennett, D. A., and Dong, H. (2018). Sexual dimorphism in predisposition to Alzheimer's disease. *Neurobiol. Aging* 70, 308–324. doi: 10.1016/j.neurobiolaging.2018.04.004

Gamberger, D., Lavrač, N., Srivatsa, S., Tanzi, R. E., and Doraiswamy, P. M. (2017). Identification of clusters of rapid and slow decliners among subjects at risk for Alzheimer's disease. *Sci. Rep.* 7:6763. doi: 10.1038/s41598-017-06624-y

Garrett, S. L., McDaniel, D., Obideen, M., Trammell, A. R., Shaw, L. M., Goldstein, F. C., et al. (2019). Racial disparity in cerebrospinal fluid amyloid and tau biomarkers and associated cutoffs for mild cognitive impairment. *JAMA Netw. Open* 2:e1917363. doi: 10.1001/jamanetworkopen.2019.17363

Gottesman, R. F., Schneider, A. L. C., Zhou, Y., Chen, X., Green, E., Gupta, N., et al. (2016). The ARIC-PET amyloid imaging study: brain amyloid differences by age, race, sex, and APOE. *Neurology* 87, 473–480. doi: 10.1212/WNL.0000000000002914

Grandner, M. A., Williams, N. J., Knutson, K. L., Roberts, D., and Jean-Louis, G. (2016). Sleep disparity, race/ethnicity, and socioeconomic position. *Sleep Med.* 18, 7–18. doi: 10.1016/j.sleep.2015.01.020

Gu, Y., Razlighi, Q. R., Zahodne, L. B., Janicki, S. C., Ichise, M., Manly, J. J., et al. (2015). Brain amyloid deposition and longitudinal cognitive decline in nondemented older subjects: results from a multi-ethnic population. *PLoS One* 10:e0123743. doi: 10.1371/journal.pone.0123743

Hales, C. M., Carroll, M. D., Fryar, C. D., and Ogden, C. L. (2020). *Prevalence of Obesity and Severe Obesity Among Adults: United States, 2017–2018* (Data Brief No. 360). Hyattsville, MD: National Center for Health Statistics.

Han, J. W., Maillard, P., Harvey, D., Fletcher, E., Martinez, O., Johnson, D. K., et al. (2020). Association of vascular brain injury, neurodegeneration, amyloid, and cognitive trajectory. *Neurology* 95, e2622–e2634. doi: 10.1212/WNL.0000000000010531

Hedden, T., Oh, H., Younger, A. P., and Patel, T. A. (2013). Meta-analysis of amyloid-cognition relations in cognitively normal older adults. *Neurology* 80, 1341–1348. doi: 10.1212/WNL.0b013e31828ab35d

Howell, J. C., Watts, K. D., Parker, M. W., Wu, J., Kollhoff, A., Wingo, T. S., et al. (2017). Race modifies the relationship between cognition and Alzheimer's disease cerebrospinal fluid biomarkers. *Alzheimers Res. Ther.* 9:88. doi: 10.1186/s13195-017-0315-1

Irwin, M. R., and Vitiello, M. V. (2019). Implications of sleep disturbance and inflammation for Alzheimer's disease dementia. *Lancet Neurol.* 18, 296–306. doi: 10.1016/S1474-4422(18)30450-2

Jack, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., et al. (2018). NIA-AA research framework: toward a biological definition of Alzheimer's disease. *Alzheimers Dement.* 14, 535–562. doi: 10.1016/j.jalz.2018.02.018

Jack, C. R., Knopman, D. S., Jagust, W. J., Petersen, R. C., Weiner, M. W., Aisen, P. S., et al. (2013). Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *Lancet Neurol.* 12, 207–216. doi: 10.1016/S1474-4422(12)70291-0

Jack, C. R., Wiste, H. J., Weigand, S. D., Knopman, D. S., Vemuri, P., Mielke, M. M., et al. (2015). Age, sex, and APOE ε4 effects on memory, brain structure, and β-amyloid across the adult life span. *JAMA Neurol.* 72, 511–519. doi: 10.1001/jamaneurol.2014.4821

Jagust, W. (2018). Imaging the evolution and pathophysiology of Alzheimer disease. *Nat. Rev. Neurosci.* 19, 687–700. doi: 10.1038/s41583-018-0067-3

Jansen, W. J., Ossenkoppele, R., Tijms, B. M., Fagan, A. M., Hansson, O., Klunk, W. E., et al. (2018). Association of cerebral amyloid-β aggregation with cognitive functioning in persons without dementia. *JAMA Psychiatry* 75, 84–95. doi: 10.1001/jamapsychiatry.2017.3391

Jorgensen, D. R., Shaaban, C. E., Wiley, C. A., Gianaros, P. J., Mettenburg, J., and Rosano, C. (2018). A population neuroscience approach to the study of cerebral small vessel disease in midlife and late life: an invited review. *Am. J. Physiol. Heart Circ. Physiol.* 314, H1117–H1136. doi: 10.1152/ajpheart.00535.2017

Koran, M. E. I., Wagener, M., Hohman, T. J., and Alzheimer's Neuroimaging Initiative. (2017). Sex differences in the association between AD biomarkers and cognitive decline. *Brain Imaging Behav.* 11, 205–213. doi: 10.1007/s11682-016-9523-8

Landau, S. M., Breault, C., Joshi, A. D., Pontecorvo, M., Mathis, C. A., Jagust, W. J., et al. (2013). Amyloid-β imaging with Pittsburgh compound B and florbetapir: comparing radiotracers and quantification methods. *J. Nuclear Med.* 54, 70–77. doi: 10.2967/jnumed.112.109009

LaPoint, M. R., Chhatwal, J. P., Sepulcre, J., Johnson, K. A., Sperling, R. A., and Schultz, A. P. (2017). The association between tau PET and retrospective cortical thinning in clinically normal elderly. *Neuroimage* 157, 612–622. doi: 10.1016/j.neuroimage.2017.05.049

Lee, C. M., Jacobs, H. I. L., Marquié, M., Becker, J. A., Andrea, N. V., Jin, D. S., et al. (2018). 18F-flortaucipir binding in choroid plexus: related to race and hippocampus signal. *J. Alzheimers Dis.* 62, 1691–1702. doi: 10.3233/JAD-170840

Lowe, V. J., Wiste, H. J., Senjem, M. L., Weigand, S. D., Therneau, T. M., Boeve, B. F., et al. (2018). Widespread brain tau and its association with ageing, Braak stage and Alzheimer's dementia. *Brain* 141, 271–287. doi: 10.1093/brain/awx320

Maass, A., Lockhart, S. N., Harrison, T. M., Bell, R. K., Mellinger, T., Swinnerton, K., et al. (2018). Entorhinal tau pathology, episodic memory decline, and neurodegeneration in aging. *J. Neurosci.* 38, 530–543. doi: 10.1523/JNEUROSCI.2028-17.2017

Manouchehrinia, A., Piehl, F., Hillert, J., Kuhle, J., Alfredsson, L., Olsson, T., et al. (2020). Confounding effect of blood volume and body mass index on blood neurofilament light chain levels. *Ann. Clin. Transl. Neurol.* 7, 139–143. doi: 10.1002/acn3.50972

Matthews, K. A., Xu, W., Gaglioti, A. H., Holt, J. B., Croft, J. B., Mack, D., et al. (2019). Racial and ethnic estimates of Alzheimer's disease and related dementias in the United States (2015-2060) in adults aged ≥65 years. *Alzheimers Dement.* 15, 17–24. doi: 10.1016/j.jalz.2018.06.3063

McDonough, I. M. (2017). Beta-amyloid and cortical thickness reveal racial disparities in preclinical Alzheimer's disease. *NeuroImage* 16, 659–667. doi: 10.1016/j.nicl.2017.09.014

Mielke, M. M., Wiste, H. J., Weigand, S. D., Knopman, D. S., Lowe, V. J., Roberts, R. O., et al. (2012). Indicators of amyloid burden in a population-based study of cognitively normal elderly. *Neurology* 79, 1570–1577. doi: 10.1212/WNL.0b013e31826e2696

Minagar, A., Sevush, S., and Bertran, A. (2000). Cerebral ventricles are smaller in Hispanic than non-Hispanic patients with Alzheimer's disease. *Neurology* 55, 446–448. doi: 10.1212/wnl.55.3.446

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta- analyses: the PRISMA statement. *PLoS Med.* 6:e1000097. doi: 10.1371/journal.pmed1000097

Morris, J. C., Schindler, S. E., McCue, L. M., Moulder, K. L., Benzinger, T. L. S., Cruchaga, C., et al. (2019). Assessment of racial disparities in biomarkers for alzheimer disease. *JAMA Neurol.* 76, 264–273. doi: 10.1001/jamaneurol.2018.4249

Mozaffarian, D., Benjamin, E. J., Go, A. S., Arnett, D. K., Blaha, M. J., Cushman, M., et al. (2015). Heart disease and stroke statistics–2015 update: a report from the American Heart Association. *Circulation* 131, e29–e322. doi: 10.1161/CIR.0000000000000152

Nebel, R. A., Aggarwal, N. T., Barnes, L. L., Gallagher, A., Goldstein, J. M., Kantarci, K., et al. (2018). Understanding the impact of sex and gender in Alzheimer's disease: a call to action. *Alzheimers Dement.* 14, 1171–1183. doi: 10.1016/j.jalz.2018.04.008

Olsson, B., Lautner, R., Andreasson, U., Öhrfelt, A., Portelius, E., Bjerke, M., et al. (2016). CSF and blood biomarkers for the diagnosis of Alzheimer's disease: a systematic review and meta-analysis. *Lancet Neurol.* 15, 673–684. doi: 10.1016/S1474-4422(16)00070-3

Ossenkoppele, R., Lyoo, C. H., Jester-Broms, J., Sudre, C. H., Cho, H., Ryu, Y. H., et al. (2020). Assessment of demographic, genetic, and imaging variables associated with brain resilience and cognitive resilience to pathological tau in patients with Alzheimer disease. *JAMA Neurol.* 77, 632–642. doi: 10.1001/jamaneurol.2019.5154

Petzold, A. (2005). Neurofilament phosphoforms: surrogate markers for axonal injury, degeneration and loss. *J. Neurol. Sci.* 233, 183–198. doi: 10.1016/j.jns.2005.03.015

Prins, N. D., and Scheltens, P. (2015). White matter hyperintensities, cognitive impairment and dementia: an update. *Nat. Rev. Neurol.* 11, 157–165. doi: 10.1038/nrneurol.2015.10

Qiao, Y., Guallar, E., Suri, F. K., Liu, L., Zhang, Y., Anwar, Z., et al. (2016). MR imaging measures of intracranial atherosclerosis in a population-based study. *Radiology* 280, 860–868. doi: 10.1148/radiol.2016151124

Rajan, K. B., Aggarwal, N. T., McAninch, E. A., Weuve, J., Barnes, L. L., Wilson, R. S., et al. (2020). Remote blood biomarkers of longitudinal cognitive outcomes in a population study. *Ann. Neurol.* 88, 1065–1076. doi: 10.1002/ana.25874

Riudavets, M. A., Rubio, A., Cox, C., Rudow, G., Fowler, D., and Troncoso, J. C. (2006). The prevalence of Alzheimer neuropathologic lesions is similar in blacks and whites. *J. Neuropathol. Exp. Neurol.* 65, 1143–1148. doi: 10.1097/01.jnen.0000248548.20799.a3

Rizvi, B., Narkhede, A., Last, B. S., Budge, M., Tosto, G., Manly, J. J., et al. (2018). The effect of white matter hyperintensities on cognition is mediated by cortical atrophy. *Neurobiol. Aging* 64, 25–32. doi: 10.1016/j.neurobiolaging.2017.12.006

Schneider, J. A., Arvanitakis, Z., Yu, L., Boyle, P. A., Leurgans, S. E., and Bennett, D. A. (2012). Cognitive impairment, decline and fluctuations in older community-dwelling subjects with Lewy bodies. *Brain* 135(Pt 10), 3005–3014. doi: 10.1093/brain/aws234

Shaaban, C. E., Aizenstein, H. J., Jorgensen, D. R., MacCloud, R. L., Meckes, N. A., Erickson, K. I., et al. (2017). In vivo imaging of venous side cerebral small-vessel disease in older adults: an MRI method at 7T. *Am. J. Neuroradiol.* 38, 1923–1928. doi: 10.3174/ajnr.A5327

Shadlen, M.-F., Siscovick, D., Fitzpatrick, A. L., Dulberg, C., Kuller, L. H., and Jackson, S. (2006). Education, cognitive test scores, and black-white differences in dementia risk. *J. Am. Geriatr. Soc.* 54, 898–905. doi: 10.1111/j.1532-5415.2006.00747.x

Snitz, B. E., Wang, T., Cloonan, Y. K., Jacobsen, E., Chang, C.-C. H., Hughes, T. F., et al. (2018). Risk of progression from subjective cognitive decline to mild cognitive impairment: the role of study setting. *Alzheimers Dement.* 14, 734–742. doi: 10.1016/j.jalz.2017.12.003

Spillantini, M. G., and Goedert, M. (2013). Tau pathology and neurodegeneration. *Lancet Neurol.* 12, 609–622. doi: 10.1016/S1474-4422(13)70090-5

Sundermann, E. E., Biegon, A., Rubin, L. H., Lipton, R. B., Mowrey, W., Landau, S., et al. (2016). Better verbal memory in women than men in MCI despite similar levels of hippocampal atrophy. *Neurology* 86, 1368–1376. doi: 10.1212/WNL.0000000000002570

van Dijk, E. J., Prins, N. D., Vrooman, H. A., Hofman, A., Koudstaal, P. J., and Breteler, M. M. B. (2008). Progression of cerebral small vessel disease in relation to risk factors and cognitive consequences: Rotterdam scan study. *Stroke* 39, 2712–2719. doi: 10.1161/STROKEAHA.107.513176

Wiegman, A. F., Meier, I. B., Schupf, N., Manly, J. J., Guzman, V. A., Narkhede, A., et al. (2014). Cerebral microbleeds in a multiethnic elderly community: demographic and clinical correlates. *J. Neurol. Sci.* 345, 125–130. doi: 10.1016/j.jns.2014.07.024

Wong, D. F., Rosenberg, P. B., Zhou, Y., Kumar, A., Raymont, V., Ravert, H. T., et al. (2010). In vivo imaging of amyloid deposition in Alzheimer disease using the radioligand 18F-AV-45 (florbetapir [corrected] F 18). *J. Nuclear Med.* 51, 913–920. doi: 10.2967/jnumed.109.069088

Wright, C. B., Festa, J. R., Paik, M. C., Schmiedigen, A., Brown, T. R., Yoshita, M., et al. (2008). White matter hyperintensities and subclinical infarction: associations with psychomotor speed and cognitive flexibility. *Stroke* 39, 800–805. doi: 10.1161/STROKEAHA.107.484147

Zahodne, L. B., Manly, J. J., Narkhede, A., Griffith, E. Y., DeCarli, C., Schupf, N. S., et al. (2015). Structural MRI predictors of late-life cognition differ across African Americans. hispanics, and whites. *Curr. Alzheimer Res.* 12, 632–639. doi: 10.2174/1567205012666150530203214

Zetterberg, H. (2019). Blood-based biomarkers for Alzheimer's disease-an update. *J. Neurosci. Methods* 319, 2–6. doi: 10.1016/j.jneumeth.2018.10.025

Zlokovic, B. V., Gottesman, R. F., Bernstein, K. E., Seshadri, S., McKee, A., Snyder, H., et al. (2020). Vascular contributions to cognitive impairment and dementia (VCID): a report from the 2018 national heart, lung, and blood institute and national institute of neurological disorders and stroke workshop. *Alzheimers Dement.* 16, 1714–1733. doi: 10.1002/alz.12157

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Systems Neuroscience

ORIGINAL RESEARCH
published: 03 August 2021
doi: 10.3389/fnsys.2021.692152

Age-Related Variations in Regional White Matter Volumetry and Microstructure During the Post-adolescence Period: A Cross-Sectional Study of a Cohort of 1,713 University Students

Ami Tsuchida[1,2,3], Alexandre Laurent[1,2,3], Fabrice Crivello[1,2,3], Laurent Petit[1,2,3], Antonietta Pepe[1,2,3], Naka Beguedou[1,2,3], Stephanie Debette[4,5], Christophe Tzourio[4,5] and Bernard Mazoyer[1,2,3,4,5]**

[1] Groupe d'Imagerie Neurofonctionnelle, Institut des Maladies Neurodégénératives, UMR 5293, Université de Bordeaux, Bordeaux, France, [2] Groupe d'Imagerie Neurofonctionnelle, Institut des Maladies Neurodégénératives, UMR 5293, CNRS, Bordeaux, France, [3] Groupe d'Imagerie Neurofonctionnelle, Institut des Maladies Neurodégénératives, UMR 5293, CEA, Bordeaux, France, [4] Université de Bordeaux, Inserm, Bordeaux Population Health Research Center, U1219, CHU Bordeaux, Bordeaux, France, [5] Centre Hospitalier Universitaire, Bordeaux, France

OPEN ACCESS

Edited by:
Patricio O'Donnell,
Takeda, United States

Reviewed by:
Ivan I. Maximov,
Western Norway University of Applied Sciences, Norway
Christian Beaulieu,
University of Alberta, Canada

***Correspondence:**
Ami Tsuchida
atsuch@gmail.com
Bernard Mazoyer
mazoyerb@gmail.com

Received: 07 April 2021
Accepted: 05 July 2021
Published: 03 August 2021

Citation:
Tsuchida A, Laurent A, Crivello F, Petit L, Pepe A, Beguedou N, Debette S, Tzourio C and Mazoyer B (2021) Age-Related Variations in Regional White Matter Volumetry and Microstructure During the Post-adolescence Period: A Cross-Sectional Study of a Cohort of 1,713 University Students. Front. Syst. Neurosci. 15:692152. doi: 10.3389/fnsys.2021.692152

Human brain white matter undergoes a protracted maturation that continues well into adulthood. Recent advances in diffusion-weighted imaging (DWI) methods allow detailed characterizations of the microstructural architecture of white matter, and they are increasingly utilized to study white matter changes during development and aging. However, relatively little is known about the late maturational changes in the microstructural architecture of white matter during post-adolescence. Here we report on regional changes in white matter volume and microstructure in young adults undergoing university-level education. As part of the MRi-Share multi-modal brain MRI database, multi-shell, high angular resolution DWI data were acquired in a unique sample of 1,713 university students aged 18–26. We assessed the age and sex dependence of diffusion metrics derived from diffusion tensor imaging (DTI) and neurite orientation dispersion and density imaging (NODDI) in the white matter regions as defined in the John Hopkins University (JHU) white matter labels atlas. We demonstrate that while regional white matter volume is relatively stable over the age range of our sample, the white matter microstructural properties show clear age-related variations. Globally, it is characterized by a robust increase in neurite density index (NDI), and to a lesser extent, orientation dispersion index (ODI). These changes are accompanied by a decrease in diffusivity. In contrast, there is minimal age-related variation in fractional anisotropy. There are regional variations in these microstructural changes: some tracts, most notably cingulum bundles, show a strong age-related increase in NDI coupled with decreases in radial and mean diffusivity, while others, mainly cortico-spinal projection tracts, primarily show an ODI increase and axial diffusivity decrease. These age-related variations are not

different between males and females, but males show higher NDI and ODI and lower diffusivity than females across many tracts. These findings emphasize the complexity of changes in white matter structure occurring in this critical period of late maturation in early adulthood.

Keywords: MRI, diffusion, white matter, DTI, NODDI, post-adolescence, cohort, cross-sectional

INTRODUCTION

Early adulthood is characterized by significant changes in lifestyle and behavior for many, when individuals explore their identity and various life possibilities to become fully independent. For some, it involves the attainment of higher education and training to acquire new skills and knowledge necessary for their planned vocation. Although the most dramatic development in the human brain takes place earlier in life, with the total brain volume reaching 90% of the adult volume by the age of 5 years (Dekaban, 1978; Lenroot and Giedd, 2006), both global and regional changes in brain structure and function persist throughout childhood and adolescence, and some of the maturational changes continue well into adulthood (Dumontheil, 2016). In particular, the white matter (WM) of the brain shows a protracted course of development, with its total volume continuing to increase up to the fourth or fifth decade of life (Walhovd et al., 2011; Lebel et al., 2012). The development of WM microstructure is also sensitive to common life experiences in young adults, including exposure to alcohol and tobacco, and other recreational drugs (Bava et al., 2013; Gogliettino et al., 2016; Silveri et al., 2016), changes in sleep patterns (Elvsåshagen et al., 2015; Telzer et al., 2015), and intensive motor and cognitive training (Scholz et al., 2009; Lövdén et al., 2010; Mackey et al., 2012; Schlegel et al., 2012; Lakhani et al., 2016). Detailed characterization of the late maturational processes of the WM in young adults is crucial for elucidating how the learning and other life experiences may shape the structural and functional organization of the brain through their impact on the brain wiring. Understanding the normative development during this period may also shed light on the vulnerability of this particular period in life to various neuropsychiatric disorders, such as substance abuse, mood and anxiety disorders (Kessler et al., 2007).

What we know about normative WM development primarily comes from non-invasive neuroimaging of typically developing individuals with magnetic resonance imaging (MRI). In addition to the macro-structural changes that can be measured with T1-weighted images, diffusion-weighted imaging (DWI) methods allow detailed characterizations of the WM microstructural properties. Over the past two decades, studies using DWI have provided much insight into the WM microstructural changes during development (reviewed in Lebel and Deoni, 2018; Tamnes et al., 2018; Lebel et al., 2019). The majority of these studies quantify DWI through a diffusion tensor imaging (DTI) model representing the direction and magnitude of diffusion of tissue water molecules as a single tensor in each voxel (Tournier et al., 2011). Most commonly, fractional anisotropy (FA), which measures the degree of diffusion directionality, is used to quantify maturational changes, with an increase in FA attributed to myelination and increased axonal size or packing. Other DTI measures include axial and radial diffusivity (AD/RD), representing diffusion along the longest and shortest axis, respectively, of the tensor modeled in each voxel, and mean diffusivity (MD), representing the average magnitude of diffusion. Across studies, FA increases and overall decreases in diffusivity with increasing age are observed in most WM regions through childhood and adolescence (e.g., Bonekamp et al., 2007; Lebel et al., 2008; Giorgio et al., 2010; Tamnes et al., 2010; Lebel and Beaulieu, 2011; Brouwer et al., 2012; Simmonds et al., 2014; Pohl et al., 2016). In a large-scale, multi-cohort study, we have recently demonstrated that such changes continue up to early to mid-adulthood (Beaudet et al., 2020).

However, being a "signal" based model, the DTI model only describes the diffusion process in each voxel and does not attempt to delineate signals attributable to different biological tissue components (Ferizi et al., 2017). Thus, changes in DTI metrics only indicate alterations in magnitude or directionality of diffusivity, and different biological processes that affect diffusion properties of the tissue cannot be distinguished (Jones et al., 2013). More concretely, FA can be increased due to myelination or increased axonal packing but would decrease with increasing fiber population complexity (e.g., crossing fibers). In contrast, "tissue" based models attempt to estimate the components of underlying tissue, typically using DWI acquisitions with multiple *b*-values, and likely provide more biologically specific insights (Alexander et al., 2019). One such model is neurite orientation dispersion and density imaging (NODDI), which models three tissue compartments (intra- and extra-cellular and cerebrospinal fluid). It estimates separate indices for neurite density (neurite density index, NDI) and fiber orientation complexity (orientation dispersion index, ODI), together with the isotropic volume fraction (i.e., cerebrospinal fluid compartment, IsoVF) (Zhang et al., 2012). Several recent studies have used NODDI to examine developmental changes in the WM microstructural properties through infancy (Jelescu et al., 2015; Dean et al., 2017), childhood to adolescence (Genc et al., 2017; Mah et al., 2017; Dimond et al., 2020; Lynch et al., 2020). These studies have indicated an age-related increase in NDI, with very little change observed in ODI in the first two decades of life (Mah et al., 2017; Dimond et al., 2020; Lynch et al., 2020), although studies covering a wider age range indicate that ODI in many WM tracts starts to increase in early adulthood (Chang et al., 2015; Slater et al., 2019). Nevertheless, a large-scale study focusing on the period of early adulthood to detail the late maturational changes in regional WM properties is still lacking.

In the present study, we characterize variations in WM-related metrics, including regional volumes and microstructural properties measured using both DTI and NODDI, in the

MRiShare database, a large cross-sectional cohort of young adults undergoing university-level education (Tsuchida et al., 2020). This study's primary goal is to document the age-related variations in the regional WM properties in this cohort. We also report on the interrelations among the age effects on different WM metrics in an effort to better understand biophysical processes underlying the late maturational changes in the WM. The secondary goal is to gain much-needed insights into the sexual dimorphism of developmental processes (Lebel et al., 2019) by investigating the effects of sex on these WM metrics and their age-related variations.

MATERIALS AND METHODS

Participants

The MRi-Share study protocol was approved by the local ethics committee (CPP2015-A00850-49). All participants were recruited through the larger i-Share cohort study (for internet-based Student Health Research enterprise).[1] Participants signed an informed written consent form and received compensation for their contribution. Out of 2,000 individuals who were enrolled between October 2015 and June 2017, 1,823 completed the MRI acquisition protocol for both structural (T1-weighted and FLAIR) and diffusion imaging. While the study protocol allowed enrollment of students up to 35 years of age, almost 95% of our sample was under 26 years old. In this study, we present the estimated age effect on WM metrics in the sub-sample of participants aged 18–26 (mean ± SD = 21.7 ± 1.8 years, N = 1,713). Age distribution was similar in males (mean ± SD = 21.9 ± 1.8 years, N = 467) and females (mean ± SD = 21.7 ± 1.7 years, N = 1,246), with only a marginal difference in their mean (2 months difference in age, p = 0.066, Welch's t-test). The higher proportion of females relative to males in MRi-Share is a feature observed among university students at the French national level that is amplified in the i-Share cohort due to an over-recruitment of students coming from faculties in which an even greater proportion of women are observed.

MRI Acquisition

The complete MRi-Share brain imaging acquisition and analysis protocols of the MRi-Share study have been detailed in Tsuchida et al. (2020). Briefly, all MRI data were acquired on the same Siemens 3T Prisma scanner with a 64-channels head coil (gradients: 80 mT/m–200 T/m/s) in the 2 years between November 2015 and November 2017. The MRi-Share acquisition protocol closely emulated that of the UKB MR brain imaging study (Alfaro-Almagro et al., 2018), in terms of both modalities and scanning parameters, with the exception of task-related functional MRI that was not acquired in MRi-Share participants. Here, we will focus on the MRi-Share structural (T1 and T2-FLAIR) and DWI brain imaging protocol. The key acquisition parameters for these scans were as follows;

- T1-weighted sagittal 3D-MPRAGE [repetition time (TR)/echo time (TE)/inversion time (TI) = 2,000/2.0/880 ms, in-plane acceleration factor (R) = 2, spatial resolution = 1 × 1 × 1 mm^3 isotropic, matrix size = 192 × 256 × 256, duration = 4 min 54 s].
- T2-weighted sagittal 3D-SPACE-FLAIR [TR/TE/TI = 5,000/394.0/1,800 ms, R = 2, partial Fourier (PF) = 7/8, spatial resolution = 1 × 1 × 1 mm^3 isotropic, matrix size = 192 × 256 × 256, duration = 5 min 50 s].
- 2D axial DWI (multi-band factor = 3, TR/TE = 3,540/75.0 ms, R = 1, PF = 6/8, fat-saturation, spatial resolution = 1.75 × 1.75 × 1.75 mm^3 isotropic, matrix size = 118 × 118 × 84, duration = 9 min 45 s).

For the DWI we acquired 8, 32, and 64 directions each for b-values 300, 1,000, and 2,000 s/mm^2, respectively, and acquired eight pairs of b = 0 images acquired in Anterior-Posterior (AP) and the reverse PA phase encoding, interleaved during the $b > 0$ acquisition. The spatial resolution of the DWI was 1.75 × 1.75 × 1.75 mm^3 isotropic, which was slightly better than that of UKB (2 × 2 × 2 mm^3 isotropic).

Image Processing

The acquired images were managed and processed with the Automated Brain Anatomy for Cohort Imaging platform (ABACI, IDDN.FR.001.410013.000.S.P.2016.000.31235; details in Tsuchida et al., 2020). Below we briefly describe the processing steps in each pipeline pertaining to the generation of the JHU atlas ROI image-derived phenotypes presented in the current paper.

T1 and T2-FLAIR Structural Pipeline

Our structural pipeline processed T1 and FLAIR images for multi-channel volume- and surface-based morphometry, primarily with SPM12[2] and Freesurfer v6.0.[3] For generating the regional WM volumes based on JHU atlas, we used the Jacobian-modulated WM probability map (1 mm isotropic) outputted by the "Unified Segmentation" framework (Ashburner and Friston, 2005) in the SPM-based volume processing branch of our pipeline (for details, see Tsuchida et al., 2020). The same Jacobian-modulated WM map was also used to obtain the total WM volume (TWMV). We also obtained the total intracranial volume (TIV) estimate based on the Freesurfer-branch of our pipeline.

Field Map Generation Pipeline

As in the UKB (Alfaro-Almagro et al., 2018), we estimated the fieldmap images from the b = 0 images with opposing AP-PA phase-encoding directions from DWI scans rather than from "traditional" fieldmaps based on dual echo-time gradient-echo images. We used all eight pairs of AP/PA b = 0 images that were interspersed in the DWI scan to estimate the susceptibility induced field and motion across the interspersed b = 0 scans using the topup tool (Andersson et al., 2003) from the FMRIB Software Library (FSL, v5.0.10).[4] The resulting subject motion parameters

[1] www.i-share.fr

[2] https://www.fil.ion.ucl.ac.uk/spm/

[3] http://surfer.nmr.mgh.harvard.edu/

[4] https://fsl.fmrib.ox.ac.uk

and the estimate of susceptibility induced off-resonance field were passed to the DWI pipeline. It also generated the brain mask based on the average distortion-corrected b0 maps, also used for the distortion corrections in the DWI pipeline.

Diffusion MRI Pipeline

A detailed description of the preprocessing steps of DWI is provided by Tsuchida et al. (2020). Briefly, the DWI data were first corrected for susceptibility and eddy-current distortion using the FSL Eddy tool, with replacement of outlier slices (*eddy_openmp* as implemented in FSL v5.0.10 patch; Andersson et al., 2016; Andersson and Sotiropoulos, 2016) and denoised by applying non-local means filter using "*nlmeans*" denoising tool (Coupe et al., 2008, 2011) as implemented in the *Dipy* package (0.12.0; Garyfallidis et al., 2014).[5] The resulting image was then used to fit (1) DTI (Diffusion-Tensor Imaging; Basser et al., 1994) modeling and (2) microstructural model fitting with NODDI (Neurite Orientation Dispersion and Density Imaging; Zhang et al., 2012). For fitting DTI, volumes with the highest b-value (b = 2,000 s/mm^2) were stripped from the data, as the accuracy of the fit starts to decrease above b = 1,000 s/mm^2 (Jensen and Helpern, 2010). Note that it still used multi-shell data, using volumes with both b = 300 and 1,000 s/mm^2 in addition to b = 0 images. The diffusivity maps were further cleaned by removing diffusivity value outliers using Random Sample Consensus (RANSAC) approach (Choi et al., 2009), as implemented in the *scikit-learn* package (0.19.1).[6] The denoising, DTI computation, and the RANSAC outlier removal were performed by wrapping *Scipy* scripts, developed by Sherbrooke Connectivity Imaging Lab.[7] For NODDI, the full set of multi-shell data was used for the fitting. We also used the empirical values of cohort-specific isotropic and parallel diffusivity as the dPar and dIso parameters for fitting NODDI (set to 1.5×10^{-3} and 2.4×10^{-3} mm^2/s, respectively), which were obtained by computing the mean MD within lateral ventricles and mean AD within the corpus callosum in individual T1 space for each subject. The preprocessing and DTI fitting were performed using tools from FSL and the *Dipy* package, while the AMICO (Accelerated Microstructure Imaging via Convex Optimization) tool (Daducci et al., 2015) was used for NODDI fitting. For each participant, the DWI processing pipeline produced seven images in native space: fractional anisotropy (FA), mean, axial, and radial diffusivity (MD, AD, and RD), based on DTI modeling, neurite density index (NDI), orientation dispersion index (ODI), and isotropic volume fraction (IsoVF), derived from NODDI.

Generation of JHU Atlas Region WM Phenotypes

We used the JHU ICBM-DTI-81 white matter labels atlas (Mori et al., 2008; Oishi et al., 2008) to generate regional phenotypes for each of the following metrics: regional WM volume and mean values for 4 DTI (FA, MD, AD, and RD) and 3 NODDI (NDI, ODI, and IsoVF) metrics. We used the atlas packaged with FSL v5.0.10, which does not have the orientation or labeling issues noted in other versions (Rohlfing, 2013) but is missing medial longitudinal fasciculus and inferior fronto-occipital fasciculus ROIs described by the authors of the atlas (Mori et al., 2008). We extracted the WM volume and mean DTI/NODDI values for 48 ROIs in this atlas, but in the absence of strong evidence for the hemispheric asymmetry in the age-related changes (Lebel and Beaulieu, 2011; Slater et al., 2019; Dimond et al., 2020), we combined values across the right and left hemispheres for the 21 pairs of ROIs present in each hemisphere by taking the average between the pair of ROIs, which were weighted by the respective volumes of each ROI in the case of DTI/NODDI metrics, to reduce the number of comparisons. **Table 1** provides the abbreviations of ROIs used in the figures and tables throughout the manuscript, and **Figure 1** presents the locations of these ROIs. They are organized according to the broad classification used by the author of the atlas: (1) tracts in the brainstem, (2) projection fibers, (3) association fibers, and (4) commissural fibers (Mori et al., 2008).

For extracting the regional DTI and NODDI values, we first computed the rigid transform for aligning DTI and NODDI maps to the native T1 reference space ($1 \times 1 \times 1$ mm^3 isotropic) with the SPM12 "Coregister" function. This transform was then aggregated with the deformation field generated in the structural pipeline to transform DTI/NODDI maps in the native DWI space to the standard template space in one step, using the SPM12 "Normalize" function. When computing the mean values within each of the 48 ROIs, we used the subject-specific, spatially normalized WM probability map, thresholded at 0.5, as an inclusive mask. It ensured that the mean values were computed within regions that are primarily WM, and minimized the partial volume effects from the surrounding non-WM tissues. **Figure 2** provides the example images of WM tissue map and DTI/NODDI maps from a representative subject, with the outlines of JHU ROIs to show the quality of alignment.

Quality Control

A detailed description of the quality control (QC) procedure for image analysis is provided in Tsuchida et al. (2020). Briefly, all structural scans were reviewed by one of the three experienced MD investigators of the MRiShare study to check for major artifacts or structural abnormalities before processing. During image processing of the structural or DWI pipelines, pipeline-specific QC images were generated for each subject. For the structural pipeline that generated reference T1 images for other modalities, a trained rater (N.B.) reviewed individual subject-specific QC images for each step of the processing for all subjects and verified that the quality of the SPM-based tissue segmentation and spatial normalization were satisfactory. For the DWI pipeline, a number of subject-specific QC images and quantifiable QC metrics mainly related to the quality of DWI data were generated (see **Supplementary Material**). Additional QC metrics for the spatial normalization were extracted by computing the image similarity of individual WM tissue probability map and DTI and NODDI scalar maps to the cohort-average maps, using Fisher z-transformed Pearson's correlation r between the two images. Two investigators (A.T. and L.P.) identified and reviewed the subject-specific QC images

[5] https://dipy.org

[6] https://scikit-learn.org/stable/index.html

[7] https://scilpy.readthedocs.io/en/latest/

TABLE 1 | Abbreviations of JHU atlas ROI names.

ROI name	Abbreviation	Hemisphere side	ROI name	Abbreviation	Hemisphere side
Brainstem			**Association**		
Middle cerebellar peduncle	MCP	Both	Fornix	FX	Both
Pontine crossing tract	PCT	Both	Fornix cres or stria terminalis	FX/ST	Right/Left
Corticospinal tract	CST	Right/Left	Cingulum cingulate gyrus	CgC	Right/Left
Medial lemniscus	ML	Right/Left	Cingulum hippocampus	CgH	Right/Left
Superior cerebellar peduncle	SCP	Right/Left	Superior fronto-occipital fasciculus	SFO	Right/Left
Inferior cerebellar peduncle	ICP	Right/Left	Superior longitudinal fasciculus	SLF	Right/Left
Projection			External capsule	EC	Right/Left
Anterior corona radiata	ACR	Right/Left	Uncinate fasciculus	UNC	Right/Left
Superior corona radiata	SCR	Right/Left	Sagittal stratum	SS	Right/Left
Posterior corona radiata	PCR	Right/Left	**Commissural**		
Anterior limb of the internal capsule	ALIC	Right/Left	Genu corpus callosum	GCC	Both
Posterior limb of the internal capsule	PLIC	Right/Left	Body corpus callosum	BCC	Both
Retrolenticular part of the internal capsule	RLIC	Right/Left	Splenium corpus callosum	SCC	Both
Posterior thalamic radiation	PTR	Right/Left	Tapetum	TAP	Right/Left
Cerebral peduncle	CP	Right/Left			

for those with extreme values in any of the QC metrics, but none of them showed any obvious signs of noticeable problems in the raw DWI or the scalar DTI and NODDI maps and their spatial normalization, except in a few cases where midsagittal plots of the raw DWI revealed a zig-zag pattern indicative of the within-volume motion in a few volumes.

Similarly, we checked the group-level distributions at the level of individual phenotypes for any missing values and the extreme outliers. Four subjects did not have any volumetric or DTI/NODDI values for fornix (FX), as the WM probability map did not overlap with this small ROI in the standard space. For the same reason, one subject was missing data for the tapetum (TAP). In addition, for corticospinal tract (CST; $n = 4$) and inferior cerebellar peduncle (ICP; $n = 6$) ROIs, mean DTI/NODDI values were not computed in the pipeline since these ROIs extended beyond the bounding box of the DWI-derived images in the standard space. Beyond these missing data, the extreme outliers were rare, and each phenotype was roughly normally distributed. Exceptions were some ROIs, in particular those surrounded by cerebrospinal fluid and/or relatively small ROIs (e.g., FX, TAP, brainstem ROIs), which had slightly skewed distributions, most likely caused by slight misalignments in DWI-derived images and structural images in standard space.

We checked for the impact of both phenotypic and QC metric outliers by removing the "far out" outliers (Tukey, 1977), defined as those with values below or above three times interquartile range (IQR) from the first or third quartile, respectively, for either the individual phenotype or any of the quantitative QC metrics. In addition to the phenotypic and QC metric outlier removal, we investigated the effect of including a global image quality metric as a covariate in the model. For the WM volume, we used the Euler number computed by Freesurfer that has been shown to be consistently correlated with the manual rating of the quality of the structural image (Rosen et al., 2018). For the DWI-based metrics, we used the mean relative RMS of the volume to volume displacement that quantifies the in-scanner motion since a recent study has demonstrated that both DTI and NODDI mean values were impacted by this QC metric (Pines et al., 2020). However, the effects of outliers or inclusion of these global quality metrics on the analyses were relatively minor (see **Supplementary Material**). For simplicity, here we report the results without any outlier removal, with total sample size of 1,713 for all ROI-metric combinations, except for FX ($N = 1{,}709$), TAP ($N = 1{,}712$), CST ($N = 1{,}709$), and ICP ($N = 1{,}707$) ROIs.

Statistical Analysis

The primary goal of the present manuscript is to describe the age-related variations in the regional WM volumes and microstructural properties in young adults. Although not our primary focus, we included sex as a covariate, and report the global pattern, mainly to characterize any overall differences between the two sexes at this age range and to examine any sex dependency in the observed age effects by including age by sex interaction term. Given our sample's narrow target age range, we expected most of the age-related variations in the volumetric and diffusion metrics to be captured by a linear age model. Indeed, the inspection of raw scatter plots (see **Supplementary Material**) did not suggest any ROIs showing any clear non-linear patterns of age-dependency. Also, a preliminary comparison of models with and without quadratic age effect to capture any non-linear trend showed that linear age effect models were sufficient for each metric and ROI combinations, as judged by the Bayesian information criterion (BIC; data not shown). Thus, for all metrics, we tested the following model;

$$Y \sim \alpha + \beta_{Age}Age + \beta_{Sex}Sex + \beta_{Age \times Sex}Age \times Sex$$

We also checked the consistency of the reported age effect estimates on the regional WM volumes when correcting for the global volume (TIV), and in the case of the DTI/NODDI metrics, examined the effects of correcting for both the

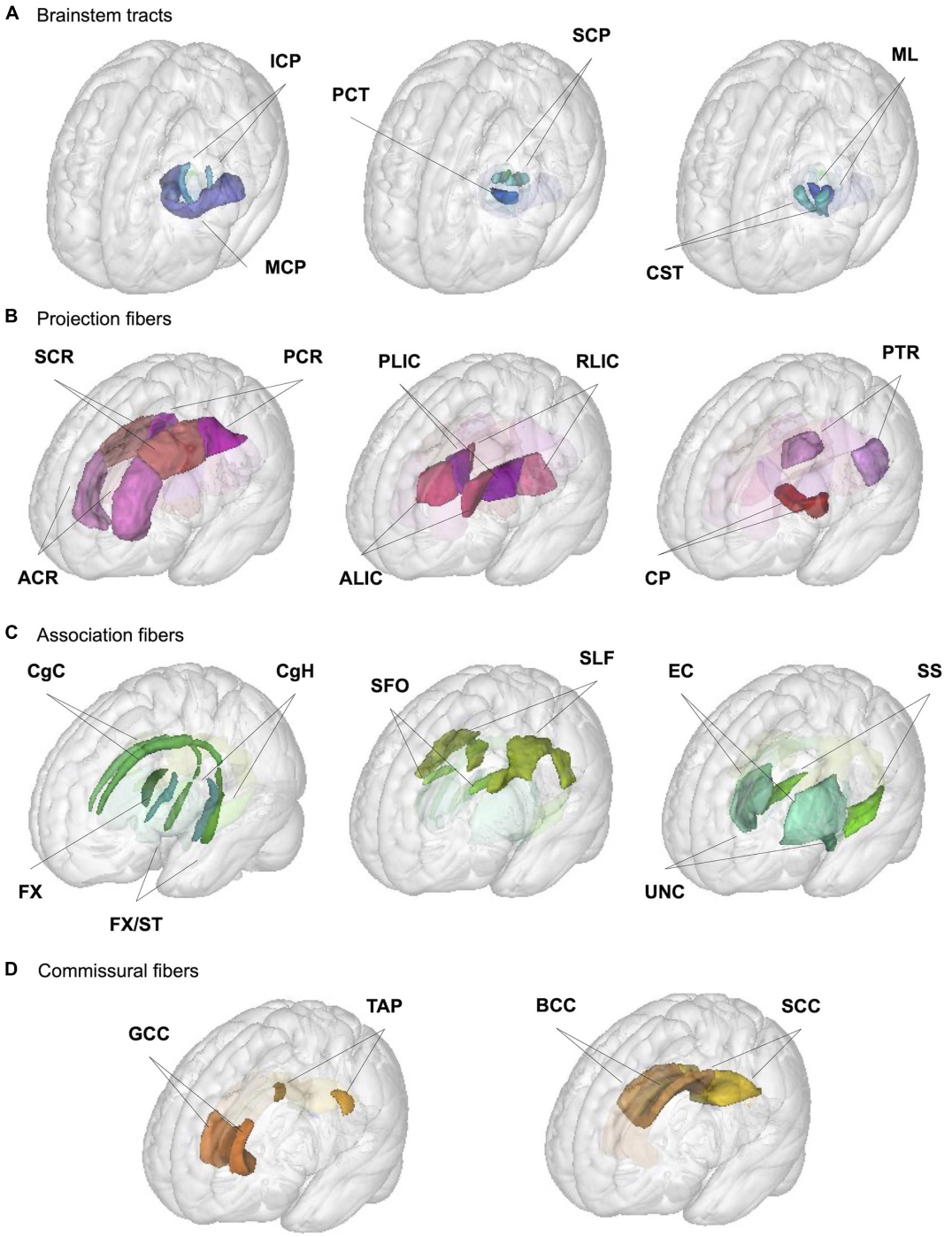

FIGURE 1 | Illustration of JHU ROIs used in the analysis. Locations of the 27 ROIs (6 medially located plus 21 pairs of ROIs in each hemisphere) from the JHU ICBM-DTI-81 white matter labels atlas are shown in the glass brain for each broad group; **(A)** brainstem, **(B)** projection, **(C)** association, and **(D)** commissural fibers. See **Table 1** for the full ROI name corresponding to the abbreviations in the figure.

global (TIV) and regional (ROI) volumes, and report them in **Supplementary Material**.

In an effort to better understand biophysical processes underlying the late maturational changes in the WM, we performed an exploratory analysis of the interrelations among the age effect estimates of the WM metrics. For this, we first computed the standardized parameter estimates (β*) of the age effect for each of the eight WM metrics across the 27 ROIs, and calculated pairwise Pearson's correlations between the β* values in the 27 ROIs for given metrics (e.g., FA vs. NDI, NDI

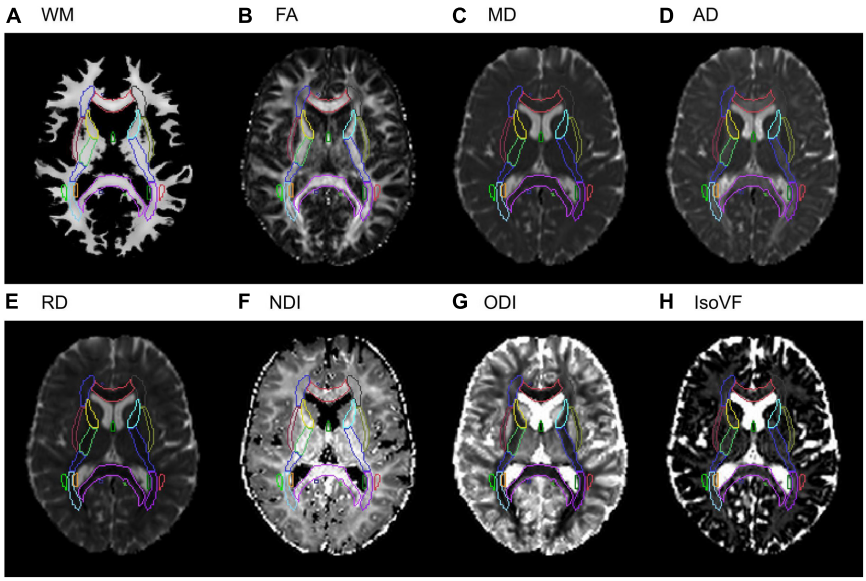

FIGURE 2 | Examples of the WM tissue map and DTI/NODDI maps from a representative subject in the stereotaxic space. A selected axial slice from **(A)** the Jacobian-modulated WM tissue map, **(B)** FA, **(C)** MD, **(D)** AD, **(E)** RD, **(F)** NDI, **(G)** ODI, **(H)** IsoVF maps in the stereotaxic space is shown for a representative subject. Outlines of the JHU ICBM-DTI-81 white matter labels atlas are shown for each image to show the quality of the alignment.

vs. ODI, etc.). Note that it quantifies the correlation between the estimates of age effects across the 27 ROIs, and not the raw correlations between metric values in the ROIs, although perfect correlations in the underlying raw data would result in the perfect correlations in the estimated effects of age and sex as well. That is, if two metrics measure a single property of WM and are perfectly correlated, the age or sex effect estimates for such hypothetical metrics would also be perfectly correlated. In reality, if two metrics represent related but distinct properties that are differentially sensitive to age or sex, correlation structures for the respective effects would be different. To illustrate this point, we also present a similar correlation structure for the estimated mean values in the ROIs across the sexes (to account for the fact that the age effect estimates also represent the value across both sexes) using the metric values standardized across the ROIs.

All model fits were performed in R, version 3.4.4 (R Core Team, 2018). We used the *lm* function as implemented in the stats library for fitting the model. The goodness of fit was assessed with adjusted R^2. The Sex contrast was deviation-coded using "contr.sum" setting so that parameter estimate (β) and *t* statistics for non-categorical variables (i.e., age in our case) represent those across sexes, and not for the specific reference sex (as would be in the case of treatment-coding, in the presence of interaction terms). Age was mean-centerd so that the intercept represented the value at group mean age. For all analyses, we report *p*-values as significant when below 0.05 after Bonferroni correction for multiple tests (27 ROIs × 8 measures, nominal *p* threshold = 0.05/216 = 0.00023). We also report generalized eta squared (η^2_G) as a measure of effect size (Olejnik and Algina, 2003), obtained using *aov_car* function in *afex* package (Singmann et al., 2021), including all terms in the model as the "observed" variables. The specification of the observed variables (as opposed to manipulated variables in other research designs) allows the correction of the effect size estimate, which makes this measure less dependent on specific research design features (Olejnik and Algina, 2003).

Visualizations of statistical summaries were created with *ggplot2* (Wickham, 2016), and tables were created with the *gt* package (Iannone et al., 2020) in R. Linear fitting of age effects for each sex was performed by predicting the given WM property in each sex using the *emmeans* package (Lenth, 2021). For evaluating the interrelations between the age-related variations in the regional WM volumes and diffusion metrics, we first computed the β* values for the respective terms in each metric using the robust standardization through refitting, implemented with the *effectsize* package (Ben-Shachar et al., 2020). Then, the β* values across 27 ROIs were used to compute Pearson's correlation between the pairwise metrics. The computation of correlation values and visualization of the results was performed using the *Ggally* package (Schloerke et al., 2021).

RESULTS

The Main Effect of Age

Table 2 presents the parameter estimates (β) for age effects for each metric (WM volume, 4 DTI and 3 NODDI metrics) across the ROIs, and **Figure 3** visually presents the summary by showing the *t* statistics and effect sizes (η^2_G) as heatmaps, filtering out those that did not survive Bonferroni corrections. **Supplementary Tables 1–8** provide the complete model results, including the confidence intervals of age β, uncorrected *p*-values and η^2_G, and total variance explained by the model for each metric and ROI. **Figure 4** provides selected scatter plots of age

TABLE 2 | Summary of age effects for each diffusion phenotype across JHU ROIs.

	Volume	FA ($\times 10^{-3}$)	MD ($\times 10^{-6}$)	AD ($\times 10^{-6}$)	RD ($\times 10^{-6}$)	NDI ($\times 10^{-3}$)	ODI ($\times 10^{-3}$)	IsoVF ($\times 10^{-3}$)
Brainstem								
MCP	−25.2	−0.0	−0.9*	−1.4*	−0.6	1.1*	0.7*	−0.6*
PCT	−2.7	−0.7	−1.9	−3.3*	−1.2	0.3	1.3*	−1.9*
CST	−1.8	−1.3*	−2.6*	**−5.6*** **	−1.1	0.2	1.9*	−2.0*
ML	−1.2	−1.1	−2.4*	**−5.1*** **	−1.1	2.2*	1.3	−1.5*
SCP	−2.3*	−0.3	**−2.0*** **	**−4.1*** **	−0.9*	**2.0*** **	0.9*	**−1.2*** **
ICP	−0.4	−0.3	−2.1**	**−3.6*** **	−1.4*	**2.4*** **	1.1*	−1.2*
Projection								
ACR	3.2	1.4**	**−2.0*** **	−1.4*	**−2.3*** **	**2.9*** **	−0.2	−0.1
SCR	5.6	−0.6	**−1.3*** **	**−2.6*** **	−0.6*	**2.2*** **	**1.2*** **	−0.1
PCR	3.3	0.6	−1.3**	−1.3*	−1.2*	**2.8*** **	0.4	0.5*
ALIC	1.7	0.6	**−2.0*** **	**−2.8*** **	**−1.6*** **	**3.1*** **	0.7*	−0.4
PLIC	−0.5	−0.7*	**−1.4*** **	**−3.4*** **	−0.5	**1.8****	**1.3*** **	−0.6**
RLIC	−0.9	0.8*	**−1.6*** **	**−1.8****	**−1.5*** **	**3.4*** **	0.5*	−0.1
PTR	−1.4	0.3	**−1.4*** **	**−2.2****	−1.0*	**2.1*** **	0.5*	0.1
CP	−3.9	−1.2*	−2.2*	**−5.8*** **	−0.4	1.4	**2.4*** **	−1.6*
Association								
FX	**2.6*** **	0.7	0.3	1.1	−0.1	**3.5*** **	1.4*	2.0*
FX/ST	−0.2	0.9*	**−2.2*** **	**−3.0*** **	**−1.7*** **	**3.7*** **	0.8*	−0.6*
CgC	8.3**	**2.3*** **	**−2.0*** **	−0.5	**−2.7*** **	**3.8*** **	−0.5	−0.3
CgH	**5.1*** **	1.2*	**−3.0*** **	**−3.6*** **	**−2.8*** **	**6.6*** **	1.6**	−0.2
SFO	0.5	0.1	**−1.8*** **	**−2.8*** **	−1.3*	**3.5*** **	0.6	−0.1
SLF	6.3	1.0*	**−1.3*** **	−0.8	**−1.5*** **	**2.6*** **	−0.1	0.1
EC	3.8	**1.2****	**−1.7*** **	−1.4**	**−1.9*** **	**3.3*** **	0.3	0.2
UNC	−0.3	1.0	**−1.8*** **	−2.1*	−1.7**	**4.1*** **	0.8*	0.7*
SS	−0.8	**1.6*** **	**−2.0*** **	−1.2*	**−2.4*** **	**3.9*** **	−0.0	0.1
Commissural								
GCC	5.1	**1.7*** **	**−1.7*** **	−0.8	**−2.1*** **	**2.4*** **	−0.4	−0.3
BCC	21.7	0.7	**−1.3*** **	−1.5*	−1.2**	**2.4*** **	0.3	−0.1
SCC	31.5*	**1.4*** **	**−1.5*** **	−1.0	**−1.7*** **	**2.6*** **	0.1	−0.5*
TAP	−2.7*	2.0*	−1.4*	0.5	−2.4*	2.3*	−0.3	0.5

*Non-standardized parameter estimates (β) for age effects for each phenotype and ROI are shown (see **Table 1** for the full names of abbreviated ROIs). The unit of the age effect is mm^3/year for the volume, mm^2/s/year for the diffusivity measures (MD, AD, and RD), and/year for FA and the NODDI phenotypes (NDI, ODI, IsoVF). Statistical significance symbols (uncorrected for multiple comparisons) *0.05 < p < 0.001, **0.001 < p < 0.0001, ***p < 0.0001. Bold symbols indicate Bonferroni-corrected significant p-values.*

effects for each sex to present examples of such effects. Similar plots of age effects for the entire metrics and ROIs are also provided in **Supplementary Figures 5–12**. As evident in **Table 2** and **Figure 3**, a number of WM ROIs showed robust age-related variations in one or more metrics we examined.

Significant age-related increases in WM volumes were observed only in cingulum hippocampus (CgH) and fornix (FX). The cingulum in the cingulate gyrus (CgC) showed a significant age-related increase when TIV or TWMV was accounted for by including them in the model (see **Supplementary Material**).

In contrast, robust age effects in DTI and NODDI metrics were observed across many ROIs, most pronounced for MD and NDI (**Figure 3**). Those with significant age effects all showed an age-related increase in NDI, and decreases in diffusivity metrics. Many of these ROIs showed a tendency for the volumetric increase as well, but some showed a significant NDI increase and diffusivity decrease without any trend for volumetric increase (see **Figure 4** for examples in CgH, with the volumetric increase, and uncinate fasciculus (UNC), without). CgH additionally showed a significant age-related decrease in AD and a trend for an ODI increase. The AD decrease was also observed across many ROIs in projection fibers and brainstem ROIs with varying degrees but was particularly pronounced in the ROIs that represent a connected pathway of projection fibers: superior corona radiata (SCR), posterior limb of the internal capsule (PLIC), and cerebral peduncle (CP) (see **Figure 4** for example in PLIC), all of which also showed a significant ODI increase with age.

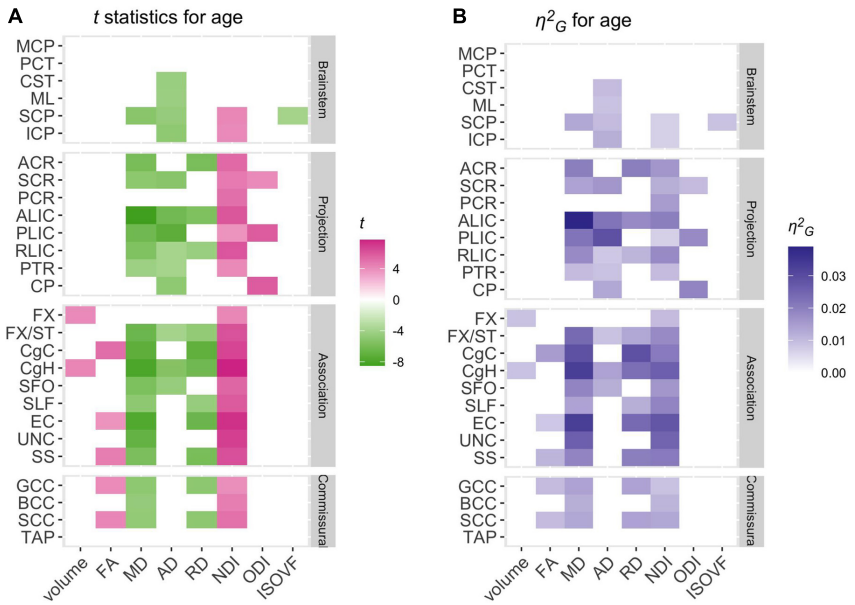

FIGURE 3 | Patterns of significant age effects across WM volume and diffusion phenotypes and ROIs. Relative statistical strengths and effect sizes of age effects across diffusion phenotypes and ROIs are shown as heatmaps of **(A)** *t* statistics and **(B)** η^2_G values (see **Table 1** for the full names for the abbreviated ROIs). Those that did not survive Bonferroni corrections for multiple comparisons were filtered out (set to 0) to facilitate comparisons within significant results. Positive *t*-scores in pink indicate an age-related increase and negative values in green indicate an age-related decrease.

Interrelations Among the Age Effects on the WM Properties

Figure 5 shows the correlation plot of the standardized parameter estimates (β*) for the age effect between pairs of metrics across the 27 ROIs. For a comparison, **Supplementary Figure 13** shows a similar plot computed for the simple regional mean values of these metrics, calculated after standardizing values across the ROIs.

The correlation structure of the age effect β* values indicated that overall, the degree of age-related variations in the regional mean FA values was negatively associated with RD and positively with AD. Thus, although both AD and RD decreased with age across the most ROIs, regions with faster age-related decreases in RD relative to AD showed overall age-related increases in FA. The degree of age-related variations in FA was also negatively associated with ODI. These patterns are expected since FA is, by definition, higher when diffusivity along the axial axis is higher than along the radial axis and when fiber orientation dispersion is lower. Indeed, such patterns were more evident in the correlations of simple mean values of the regional WM metrics, which showed a strong positive correlation between the regional FA and AD values and also strong negative correlations between the regional FA and RD or ODI values.

In contrast, the correlation patterns for NDI were distinct between the regional age effects and the simple mean values: the degree of age-related increases in NDI was positively associated with the degree of age-related variations in the regional WM volume and FA, and negatively associated with the age-related decrease in RD (i.e., regions with more NDI increases showing more volumetric and FA increases and RD decreases). In the regional mean values, the higher NDI values were not strongly associated with the regional WM volumes or FA and RD values. Another difference was the non-significant but negative correlation between the age effects on NDI and ODI, indicating the ROIs showing more age-related increases in NDI tended to show less ODI increases, while in the regional mean values, NDI and ODI were weakly but positively associated, indicating higher NDI values in ROIs with higher ODI. Note that despite the weak correlations between the age effects on the regional volume and FA, RD, and NDI values, the age effects in these microstructural properties were not affected by the inclusion of the regional volumes as a covariate in the model (see **Supplementary Material**).

Dependency of Age-Related Variations on Sex

Summary of *t* statistics and η^2_G values for the sex effect on each of the eight WM properties across the JHU ROIs are presented in **Figure 6** (see also **Supplementary Tables 1–8**).

Not surprisingly, males had larger WM volumes than females across most of the ROIs examined. However, the difference diminished considerably when global volume differences were taken into account by including either TIV or TWMV in the model (see **Supplementary Material**). For the diffusion metrics, females showed higher diffusivity than males across many ROIs, while males showed higher NDI and ODI overall. There were relatively few regionally specific patterns in the sex effects, although the differences were most robust in the brainstem ROIs.

Despite the widespread main effects of sex, we did not observe any significant sex differences in the age-related variations in

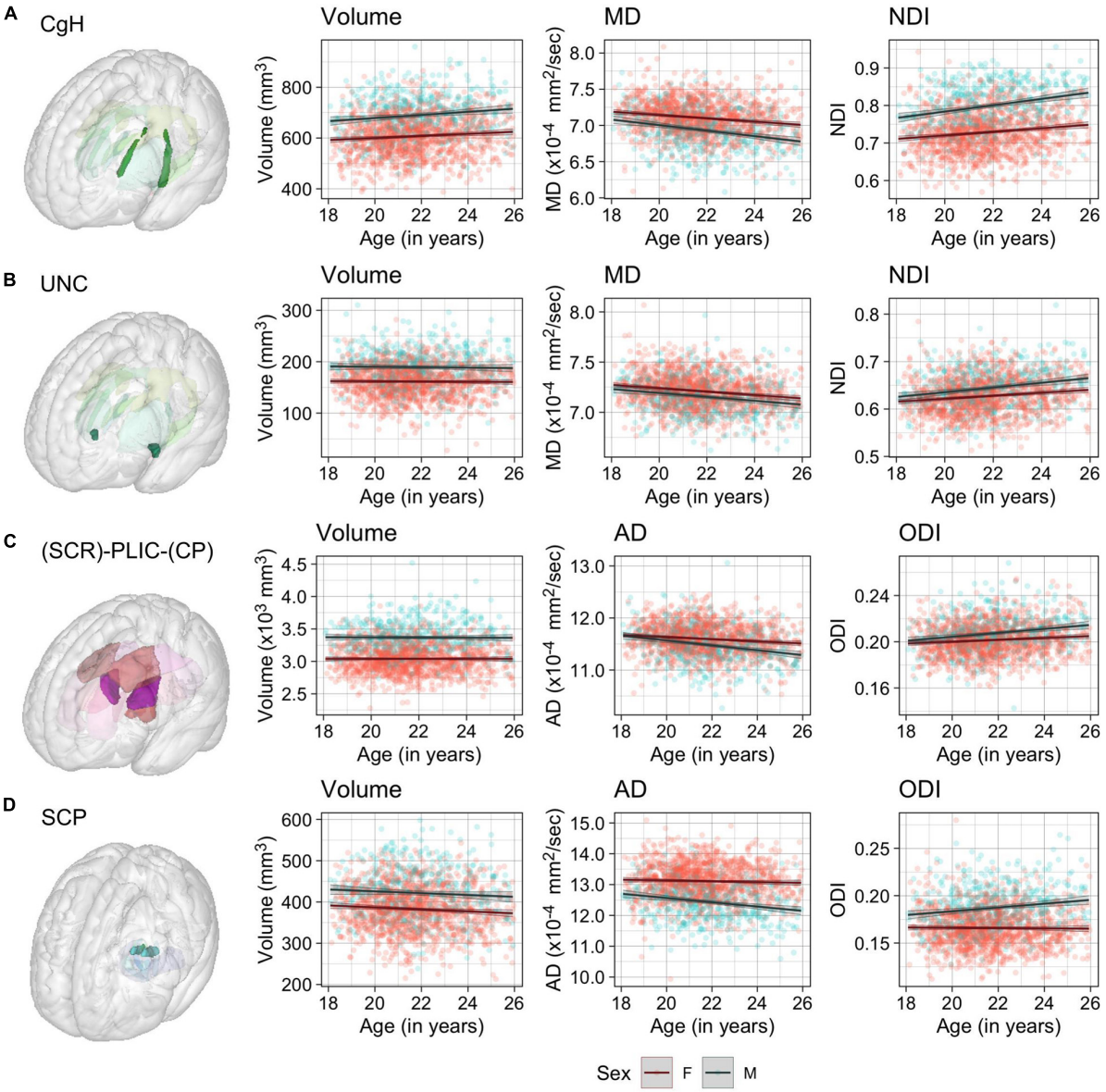

FIGURE 4 | Scatter plots of individual age effects in **(A)** the cingulum-hippocampus (CgH), **(B)** uncinate fasciculus (UNC), **(C)** posterior limb of the internal capsule (PLIC), and **(D)** superior cerebellar peduncle (SCP) ROIs. Predicted linear regression lines are superimposed for each sex (dark red: females, dark cyan: males), with shades indicating the 95% confidence intervals.

the WM properties (the lowest uncorrected $p = 0.0008$). Overall, any non-significant sex differences in the age-related trajectory tended to show a steeper slope in males than in females, in particular for AD and ODI (see for example in the SCP, **Figure 4** and **Supplementary Figures 5–12**).

DISCUSSION

The primary objective of the present study was to characterize the late maturational changes in the regional WM properties during post-adolescence in the large and unique sample from the MRi-Share database. We also examined sex differences in the WM of this sample and assessed whether the age-related changes differed between the two sexes. Below we discuss our main findings in relation to the existing literature, comment on the specific features of our dataset, and methodological strengths and limitations of the present study.

Age-Related Variations in Regional WM Properties

We observed widespread age-related increases in the NDI as well as decreases in diffusivity (MD, AD, and RD) across many of the JHU ROIs in our sample of young adults aged between

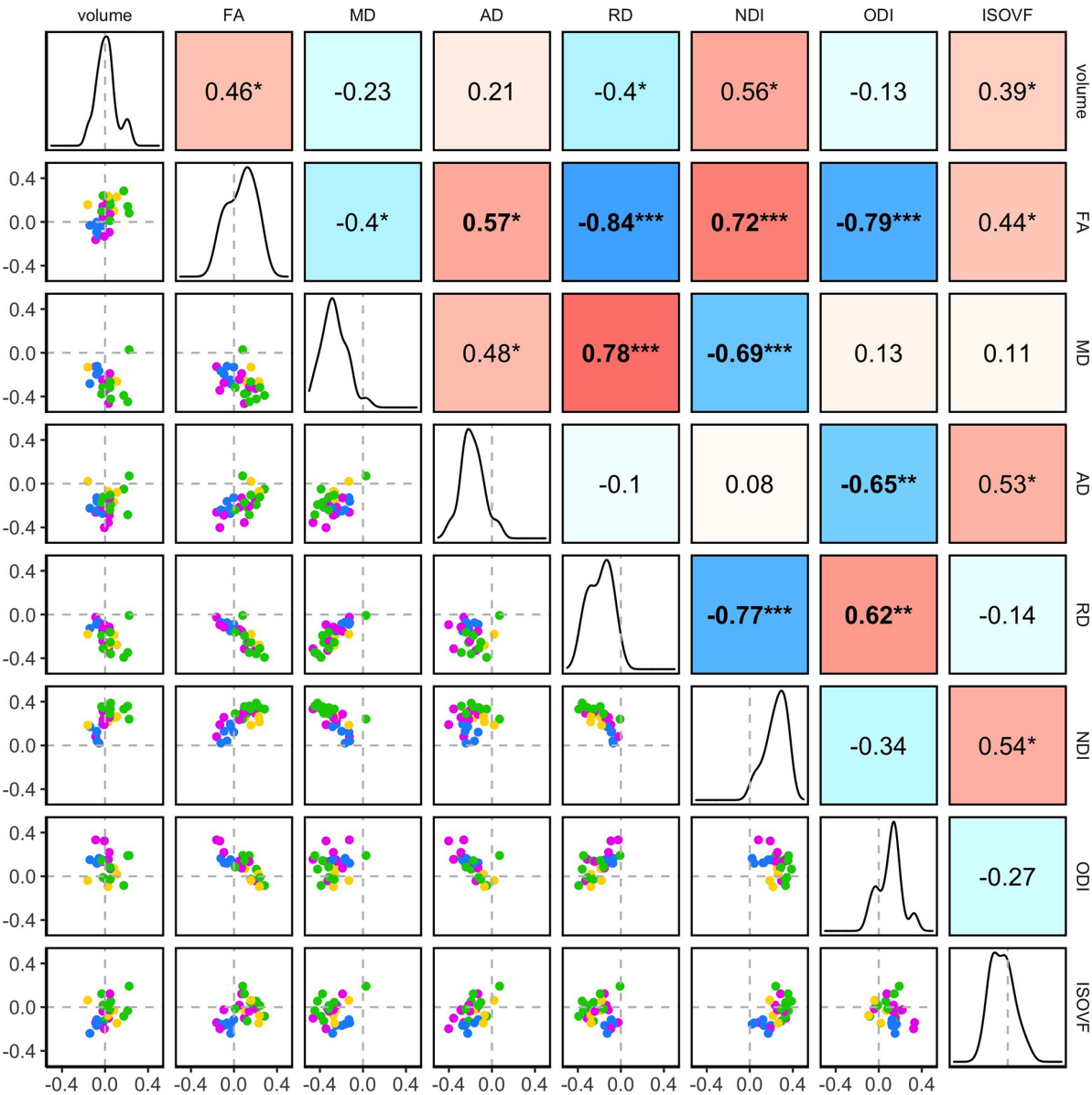

FIGURE 5 | The inter-relations between the age-related variations in the regional WM properties. Pairwise correlations of the standardized parameter estimates (β*) for age effects in the 27 ROIs are shown. The diagonal of the plot matrix shows the distributions of β^*_{Age} for the regional WM volume and DTI/NODDI values. The upper triangle shows Pearson's correlation (*r*) values. The lower triangle shows the pairwise scatter plots of β^*_{Age}, with the colors indicating the ROI groups (*blue*: brainstem, *pink*: projection, *green*: association, *yellow*: commissural). Statistical significance symbols (uncorrected for multiple comparisons) *0.05 < p < 0.001, **0.001 < p < 0.0001, ***p < 0.0001. Bold-face indicates a significant correlation after Bonferroni correction for multiple comparisons (28 correlations).

18 and 26 years. Changes in FA were statistically weaker, but ROIs with significant age effects all showed an increase with age. Regional volumes did not vary significantly with age for the most part but showed trends for an age-related increase in some ROIs. The degree of age-related increases in FA and volume in each ROI were nonetheless correlated with the degree of age-related variations in the NDI and diffusivity. Regionally, we observed that many ROIs in projection and brainstem fiber groups showed primarily significant age-related decreases in AD. In contrast, those in association and commissural fiber groups were more characterized by decreases in RD. Several ROIs in the corticospinal pathway additionally showed age-related increases in ODI.

The global patterns we observed in our sample are consistent with a wealth of literature showing a relatively protracted maturation of human brain WM: both developmental and life-span studies of WM volume and DTI metrics have indicated continued increases in global WM volume and FA into young adulthood, together with decreases in diffusivity that peaks sometime in young to mid-adulthood (Hasan et al., 2007,

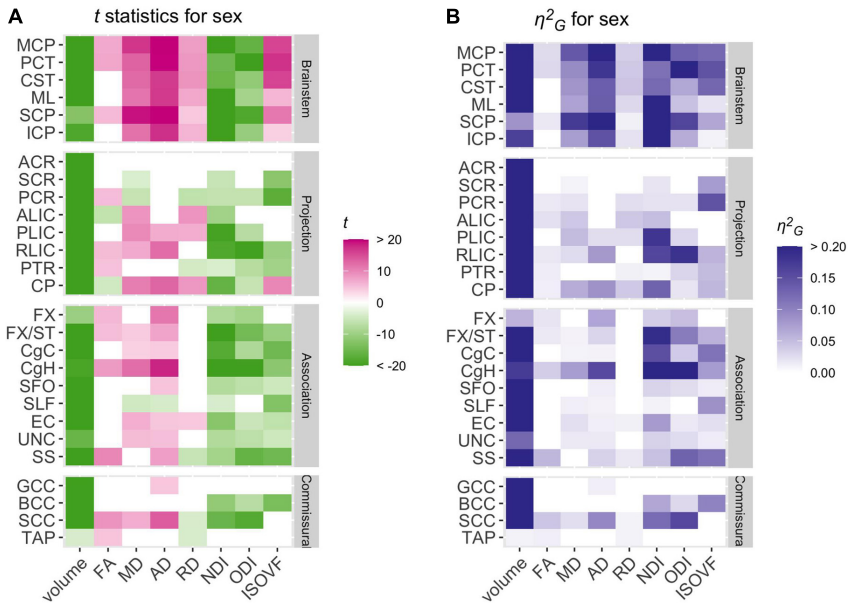

FIGURE 6 | Patterns of significant sex effects across WM volume and diffusion phenotypes and ROIs. Relative statistical strengths and effect sizes of sex effects across WM phenotypes and ROIs are shown as heatmaps of **(A)** *t* statistics and **(B)** η^2_G values (see **Table 1** for the full names for the abbreviated ROIs). Those that did not survive Bonferroni corrections for multiple comparisons were filtered out (set to 0) to facilitate comparisons within significant results. Positive *t*-scores in pink indicate higher values in females than in males and negative values in green indicate the opposite.

2010; Westlye et al., 2010; Lebel et al., 2012; Slater et al., 2019; Beaudet et al., 2020; Tsuchida et al., 2020). More recent studies using NODDI have also shown the continuous increase of NDI through development (Genc et al., 2017; Mah et al., 2017; Dimond et al., 2020; Lynch et al., 2020; Pines et al., 2020) and adulthood (Billiet et al., 2015; Chang et al., 2015; but see Kodiweera et al., 2016), peaking around the fourth and fifth decade of life (Slater et al., 2019; Qian et al., 2020). ODI, on the other hand, has not been reported to change noticeably during development (Dimond et al., 2020) or show a slight decrease in some tracts (Lynch et al., 2020) but starts to increase during young adulthood (Chang et al., 2015; Slater et al., 2019) that continues through aging (Billiet et al., 2015; Beck et al., 2021).

More robust and wide-spread increase in NDI than FA observed in our data likely results from the fact that we sampled the FA values from the entire WM regions within each ROI, rather than a limited "core" region with high FA values, a common approach in studies using the same JHU atlas, as discussed in the section on Potential limitations below. When sampling over regions with more complex fiber organizations, NODDI can provide more specific insights than FA, since FA can be influenced by both the fiber density and myelination as well as by the composition of fiber orientations (among other things) in the sampled voxel (Zhang et al., 2012; Jones et al., 2013). This point is corroborated by the relationships we observed between the age effects on the regional FA and NDI or ODI; while the age-related increase in FA was positively correlated with that of NDI, it was negatively correlated with the degree of age-related increase in ODI. It suggests that concomitant increases in NDI and ODI can have an opposing impact on the regional FA, rendering it less sensitive to the effects of age.

Regionally, we observed that cingulum WM showed a prominent age-related increase in NDI as well as MD and RD decreases. With concurrent RD reduction, the NDI increase is suggestive of increased myelination (Song et al., 2005). Cingulum WM in hippocampal region (CgH) also showed a robust volumetric increase as well, both in terms of raw volume and relative to TIV or TWMV. However, in these and other ROIs, the regional volume had little impact on the observed age effects on other WM properties, suggesting the distinct biological processes governing the age-related changes in WM volumes and other metrics related to microstructural properties (Lebel et al., 2019). Previous studies have indicated cingulum to be one of the last major tracts to mature during development, reaching peak values in FA or minimum values in MD later than other tracts (Tamnes et al., 2010; Westlye et al., 2010; Lebel et al., 2012). Similarly, a higher rate of NDI growth in limbic tracts that include CgC and CgH has been reported in a sample of 66 healthy subjects with a mean age of 25 years (Chang et al., 2015). A more recent and larger-scale lifespan study on regional DTI and NODDI metrics in 801 individuals aged 7–84 years has also indicated a relatively late peak age for NDI in CgC and CgH (Slater et al., 2019). The cingulum bundle primarily contains fibers that link cingulate gyrus and hippocampus (Mori et al., 2008), but also consists of many short association fibers that interconnect medial parts of the frontal, parietal, and temporal regions (Heilbronner and Haber, 2014). With the diverse fiber populations that make up this bundle, neuroimaging studies in healthy subjects as well as in clinical populations have implicated this region for a wide

range of cognitive functions: these include executive control, motivation, and pain in anterior/dorsal cingulate and memory in hippocampal region (reviewed in Bubb et al., 2018). Several studies have also shown the link between the microstructural integrity of the cingulum bundle and cognitive performance in children (Bathelt et al., 2019) and older adults (Kantarci et al., 2011; Bettcher et al., 2016). In this context, robust age-related changes observed in the cingulum ROIs in our sample of young adults undergoing higher-level education are particularly interesting. Future studies should investigate the relevance of volumetric and microstructural differences across subjects in cingulum to cognitive and academic performance and emotional and behavioral development.

Beyond the cingulum bundle, all association ROIs tended to show a higher increase in NDI (average annual percentage increase, computed from the base value at age 18) of 0.55%/year, ranging from 0.37%/year in SLF to 0.89%/year in CgH) than commissural (average of 0.35%/year, ranging from 0.32%/year in GCC and 0.41%/year in TAP) and projection ROIs (average of 0.34%/year, ranging from 0.16%/year in CP and 0.45%/year in ACR and RLIC). The NDI increase was smallest in the brainstem ROIs (average of 0.16%/year) and least statistically significant. Of note, the brainstem ROIs also had the highest estimated NDI at age 18 [mean (range) = 0.89 (0.84–0.96)], while association fiber ROIs had the lowest estimated NDI at the same age [mean (range) = 0.71 (0.62–0.80)]. It suggests that most of the NDI growth in brainstem ROIs likely takes place earlier than the age range of our sample. The observed pattern is broadly consistent with previous DTI studies suggesting earlier maturation in the commissural and projection fibers, followed by association fibers, especially in fronto-temporal regions (Tamnes et al., 2010; Westlye et al., 2010; Lebel et al., 2012). More recent studies with NODDI also support similar regional patterns of the developmental trajectory (Dean et al., 2017; Slater et al., 2019; Lynch et al., 2020). For instance, in a recent study examining the maturational timing of regional NODDI parameters in a cross-sectional sample of 104 subjects aged between 0 and 18 years, the NDI growth in callosal fibers reached a plateau the earliest, followed by projection and association fibers (Lynch et al., 2020).

While relatively modest in terms of NDI growth, we found that the connected ROIs of projection fibers, from superior corona radiata (SCR), through the posterior limb of the internal capsule (PLIC), then to cerebral peduncle (CP), showed the age-related increase in ODI and decrease in AD. It suggests the increasing fiber complexity in this large WM bundle that contains the pyramidal and cortico-pontine tracts. This observation is novel, and has not been reported in previous studies examining age-related variations in regional NODDI values in subjects with age-range that overlaps with our study (Billiet et al., 2015; Chang et al., 2015; Slater et al., 2019; Pines et al., 2020). None of these studies reported notable age-related ODI increase in this projection fiber pathway that stood out from other regions (e.g., non-brainstem projection fiber ROI in Chang et al., 2015 and tractography-based corticospinal tract in Slater et al., 2019). However, different methodology in defining the tract ROI as well as modeling strategies makes the direct comparison difficult. Future studies are needed to confirm the validity of our observation and investigate the functional relevance of such age-related changes.

Biophysical Interpretation of Age-Related Variations in DTI and NODDI

The present study demonstrates the usefulness of NODDI metrics in at least partially disambiguating the factors that can result in the observed patterns of age-related variations in DTI metrics: the overall age-related decreases in diffusivity were associated with two uncorrelated increases in NDI and ODI, with NDI increases associated with decreases in RD and ODI with decreases primarily in AD. It indicates that the age-related variations in DTI metrics at this age range likely result from changes in both the intra-neurite fraction and fiber complexity. However, it should be cautioned that as in any other models, NODDI makes certain assumptions that over-simplify the underlying microstructure, and it has been criticized in recent years that some of these assumptions are invalid and can introduce biases in the estimates (Jelescu et al., 2015; Lampinen et al., 2017). In particular, the assumption of a single and fixed intrinsic diffusivity for both intra- and extracellular space that causes non-negligible biases in ODI and IsoVF, as well as large uncertainty in the IsoVF estimation (Jelescu et al., 2015). NDI has also been shown to be overestimated in the tissue with lower diffusivity than assumed in the model, such as in the gray matter and in pathology (Lampinen et al., 2017). Even when the estimates are free from biases, the underlying biological phenomena are not as specific as the naming of NDI ("neurite density" index) suggests, since any microstructural changes that can affect intra-neurite fraction directly (increase in the number and density of axons) or indirectly by affecting the volume of the extra-axonal space (for example myelination). Such ambiguity is evident in a number of speculative interpretations in the clinical applications of NODDI in the literature (Kamiya et al., 2020). Ultimately, precise biological interpretations of observed changes or variations in NODDI should be validated through comparisons with histological studies and with complementary or higher-order diffusion models (Jelescu et al., 2020). Nonetheless, in the case of the white matter in normal development, it is likely that the observed patterns of NODDI and DTI metrics reflect the myelination and remodeling of myelin, rather than an increase in the number of axons (de Graaf-Peters and Hadders-Algra, 2006; Sampaio-Baptista and Johansen-Berg, 2017). The correlation of the age-related increases in NDI with decreases in RD, but not with AD, is consistent with this interpretation. Future studies should investigate the validity of this observation, and also examine how such changes in young adults are affected by cognitive and physical activities, and other lifestyle and environmental factors.

Sex Differences in the WM Properties and the Patterns of Age-Related Variations

In many ROIs, we detected significant sex differences in the regional WM properties but found very little evidence for sex

differences in the age-related variations in the WM properties at this age range. The sex differences in the regional WM volumes were most likely due to differences in overall head size, as the inclusion of TIV or TWMV diminished most of the differences. However, we also observed globally higher diffusivity (MD, AD, and RD) in females than in males and higher NDI and ODI in males than in females, which cannot be accounted for by the head size differences. While there are some studies reporting lower MD values in males than in females in young adults and adolescents (e.g., Lebel and Beaulieu, 2011; Herting et al., 2012), such sex differences are often more regionally specific and not universally detected across studies (e.g., Lebel et al., 2008; Tamnes et al., 2010). Nonetheless, we did observe greater MD and AD over the entire WM skeleton in females than in males in our recent large-scale multi-cohort study (total $N > 20{,}000$) that covered most of the adult life span (Beaudet et al., 2020), suggesting that the greater diffusivity in females is not unique to this sample. With regard to NODDI, one study with a young to middle-age sample (age range of 18–55 years) reported a robustly higher NDI and ODI in males than in females (Kodiweera et al., 2016), similar to our findings; however, most studies do not report any sex differences in childhood and adolescence (Genc et al., 2017; Mah et al., 2017; Dimond et al., 2020; Lynch et al., 2020).

Despite the main effects of sex, we did not detect strong evidence for the sex differences in the age-related variations in our data. It is consistent with prior studies that report no or minimal interaction between sex and age after post-childhood in DTI (Hsu et al., 2010; Hasan et al., 2010; Tamnes et al., 2010; Inano et al., 2011; Lebel et al., 2012; Pohl et al., 2016) or NODDI (Cox et al., 2016; Kodiweera et al., 2016; Slater et al., 2019; Lynch et al., 2020). It suggests that any sex-related differences in the WM properties develop relatively early in development. Indeed, some studies reported steeper age-related changes in both FA and MD in boys than girls during childhood (Simmonds et al., 2014; Reynolds et al., 2019). However, further studies are needed to determine factors that may influence apparent sex differences in the WM properties and their rate of change with age in specific cohorts, such as body mass index, physical and intellectual activities, and other behavioral differences between the sexes that may modulate the WM properties.

Potential Limitations

As we describe more in detail in Tsuchida et al. (2020), our sample from the MRi-Share database is drawn from students undergoing university-level education in Bordeaux, and as such, not necessarily a representative sample of healthy young adults. As a consequence, our sample is dominated by female participants, for example, and likely have different socio-demographic backgrounds and levels of education than the rest of the population of the same age range. They are also not guaranteed to be perfectly "healthy," as the i-Share study, from which the MRi-Share participants were drawn, was designed to investigate the physical and mental health of students, and did not exclude those with a past or current history of mental illness, alcohol intake, smoking habits, and/or use of any recreational drugs and psychotropic medications. While this undoubtedly increases the variance unaccounted for in our data, it also makes our data more representative of the sampled population.

The MRi-Share database is also currently cross-sectional, limiting our inference of maturational trajectory from the age-related variations in the data. The analysis of age effects based on cross-sectional data has been shown to lead to spurious findings unsupported from longitudinal analysis, especially when using quadratic models to describe non-linear patterns of age-related changes (Fjell et al., 2010; Pfefferbaum and Sullivan, 2015). In our sample with a relatively limited age range of 18–26 years, we found that linear age trends were sufficient for characterizing age-related variations in the data, thus avoiding some of the pitfalls of fitting quadratic age models. While we still need to exert caution when interpreting the apparent age-related variations in our data, our findings were found to be broadly consistent with the known age-related trajectories in WM properties.

Though our DWI preprocessing pipeline included standard steps with susceptibility and eddy-current distortion correction and was similar to the official UKB DWI pipeline [with additional denoising using non-local means filter (Coupe et al., 2008, 2011)], our study did not make use of additional preprocessing steps such as bias field correction and Gibbs ringing correction. Recent work has highlighted the potential impact of such preprocessing choices on diffusion metrics and the observed age associations (Maximov et al., 2019). We also used the version of Eddy (patch 5.0.10) before the option to correct for within-volume movement (Andersson et al., 2017) and interactions between susceptibility and motion (Andersson et al., 2018) implemented in the latest version of the tool. Future investigations with this dataset may benefit from the updated preprocessing pipeline that incorporates these steps and examine the reliability of the findings from the current study.

Regarding the specific methodology for characterizing the regional WM properties, we used the ROIs based on the JHU ICBM-DTI-81 white matter labels atlas, computing the mean DTI/NODDI values within regions with high WM probability based on the multi-channel tissue segmentation with T1 and FLAIR scans. The ROIs in this atlas represent the WM regions with relatively well-organized structures that are clearly visible in the color-coded map of the tensor fields and should not be conflated with tracts obtained through tractography-based methods: The naming of these ROIs is based on the primary WM fiber population passing through the region, but these ROIs often represent a limited portion of a given tract, with arbitrary boundaries, and also may contain different fiber populations. For example, the corticospinal tract (CST) ROI in this atlas represents a portion of the CST at the level of medulla and pons, whereas the CST in the tractography-based methods usually refers to the fiber population that spans from corona radiata, passing through the internal capsule, then to the midbrain (Thiebaut de Schotten et al., 2011; Chenot et al., 2019). Another example is the sagittal stratum (SS) ROI, which, according to the authors of the atlas, includes both the inferior longitudinal fasciculus and the projection fibers from the internal capsule, therefore including both projection and association fibers (Mori et al., 2008). We also note that recent anatomical studies have seriously questioned the presence of superior fronto-occipital fasciculus (SFO) in humans

(Türe et al., 1997; Forkel et al., 2014; Meola et al., 2015; Liu et al., 2020). Thus, this ROI most likely represents anterior thalamic radiation, as has been noted by the authors (Mori et al., 2008).

Although tractography-based methods allow a more direct characterization of any given tract in the WM, averaging of diffusion metrics along the entire length of tracts of interest can be problematic, in particular for DTI metrics, which can vary considerably along the tract due to the variability in the fiber tract geometry (Lebel et al., 2008; Vos et al., 2012). For this reason, more detailed comparisons of metrics at arbitrary points along the tract ("tract profiling") have been proposed (Jones et al., 2005; Colby et al., 2012; Yeatman et al., 2012; Cousineau et al., 2017). Regardless of how to sample values from the tracts of interests, the choice of specific tracts to be extracted, the tracking or extraction criteria (seeding and exclusion regions for tracking specific tracts, or inclusion or exclusion criteria when extracting specific tracts from a whole-brain tractogram), tracking algorithms and their hyperparameters can complicate comparisons across studies (Côté et al., 2013). To avoid the bias introduced by study-specific protocols, a number of automated or semi-automated methods to extract major WM tracts have been proposed in recent years (Zhang et al., 2008; Yendiki et al., 2011; Yeatman et al., 2012; Wassermann et al., 2016; Wasserthal et al., 2018; Warrington et al., 2020), but no one method has been applied widely to characterize age-related changes in WM properties (Lebel et al., 2019). Also, more work is needed to assess the reproducibility and anatomical validity of different protocols for tract reconstructions (Rheault et al., 2020).

Within the studies using the ROI-based approach, and in particular the ROIs based on the same JHU ICBM-DTI-81 atlas, many use the framework of Tract-Based Spatial Statistics (TBSS, Smith et al., 2006), included with the FSL package. TBSS was developed to overcome the limitations of voxel-based analyses as applied to DTI metrics, namely the difficulty of aligning complex fiber architecture across subjects and the problem of smoothing images with highly heterogeneous noise such as FA. Its approach is to project the highest local FA values onto the non-linearly aligned group average or a template FA map that has been "skeletonized" by only taking the regions with maximal FA values with low inter-subject variability (Smith et al., 2006). The DTI or any other maps of diffusion metrics can then be projected to the FA-based skeleton to perform a voxel-based comparison within the skeleton or an ROI-based comparison using the atlas, such as the JHU atlas used in the present study. The focus on the WM skeleton with high FA values across subjects resolves the issue of alignment and correspondence across multiple subjects, but by design, it biases the characterization of the WM microstructural properties to the very small portion of WM inside the skeleton that is only one voxel in width, with relatively simple fiber orientations (Lebel et al., 2019). When used together with the ROI-based approach, the number of voxels contributing to the analyses are further reduced. In the present study, we used less restrictive sampling based on the WM probability map rather than the TBSS-style FA skeleton to allow for a more complete characterization of the regional microstructural properties. This approach also allowed for the direct comparison of the variations in the regional volume based on the Jacobian-modulated WM probability map and the variations in the microstructural properties in the same region. The inclusion of voxels outside the FA skeleton likely explain the relative lack of age or sex effects for mean FA values in our study, since it averages over regions with more complex fiber geometry and makes it difficult to dissociate changes related to the axonal diffusion properties from those related to the complexity of fiber orientations. However, multi-component tissue models such as NODDI can offer more specific inferences about the variations or differences in the microstructural properties without restricting the analysis to the WM skeleton, as we demonstrated in our study.

Another critical difference between the TBSS-based approach and the current study is the method of spatial normalization: after non-linear alignment of FA map to the template space, the TBSS projects the highest FA values onto a template FA skeleton in the standard space. Although it is meant to improve the alignment of the core of WM tracts, concerns have been raised with regard to the anatomical inaccuracies introduced by such a method (Bach et al., 2014). In the present study, we used the "Unified Segmentation" framework (Ashburner and Friston, 2005) to perform spatial normalization based on tissue segmentation of the structural scans, a common approach in voxel-based morphometry studies (e.g., Takao et al., 2011; Powell et al., 2012; Shiino et al., 2017). The non-linear deformation field obtained from the spatial normalization of the structural scans was then applied to DTI and NODDI maps, together with affine transformations that co-register these maps to the reference T1 scan of each subject. Although this is not necessarily the best available method to non-linearly align images (Klein et al., 2009), we believe that the sampling and averaging of values within the regions comprising hundreds of voxels (or thousands, in many ROIs), defined based on both the template atlas label and subject-specific WM probability map, would limit the effects of small misalignments, especially with the large sample size in our study. Having said that, the robustness of the findings should be confirmed using state-of-the-art methods to align images, such as registrations based on diffusion tensor images (Zhang et al., 2006) or fiber orientation distributions (Raffelt et al., 2011).

CONCLUSION

In a large cohort of university students, we found a widespread increase in NDI, with a more regionally specific increase in ODI, indicating a continuing modulation of WM properties at this age range. We also demonstrated the distinct patterns of interrelations among the estimated age effects on different WM properties that were consistent with remodeling of myelin in post-adolescence. We did not find any evidence for a strong sex dependency in the patterns of age-related variations. These findings highlight the complexity of the patterns of regional WM properties and individual variations in such patterns. Although we focused on the basic characterization of age and sex effects in the present study, they represent a small portion of the variance in data, and there are large individual differences in the regional WM volumes and microstructure. Future studies should investigate how the maturational processes in the WM influence,

or are influenced by, genetic, cognitive, behavioral, lifestyle and social factors, and how they are altered in neuropsychiatric conditions that manifest in early adulthood.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because to access i-Share and MRi-Share de-identified data, a request can be submitted to the i-Share Scientific Collaborations Coordinator (ilaria.montagni@u-bordeaux.fr) with a letter of intent (explaining the rationale and objectives of the research proposal), and a brief summary of the planned means and options for funding. The i-Share Steering Committee will assess this request, and provide a response (principle agreement, request to reformulate the application or for further information, refusal with reasons). If positive, applicants will have to complete and return an application package which will be reviewed by the principal investigator, the Steering Committee, and the operational staff. Reviews will be based on criteria such as the regulatory framework and adherence to regulations (access to data, confidentiality), the scientific and methodological quality of the project, the relevance of the project in relation to the overall consistency of the cohort in the long term, the complementarity/competition with projects planned or currently underway, ethical aspects. De-identified data (and data dictionaries) will be shared after (i) final approval of the application, and (ii) formalization of the specifics of the collaboration. The JHU WM phenotypes and associated QC metrics presented in this study are available in the Dryad repository (https://doi.org/10.5061/dryad.cvdncjt4m), and source codes for the statistical analysis presented are available on GitHub (https://github.com/atsuch/MRiShare_regionalWM_Age_analysis) or Zenodo (https://doi.org/10.5281/zenodo.5072215) repository associated with the Dryad repository. Requests to access the datasets should be directed to Ilaria Montagni, ilaria.montagni@u-bordeaux.fr.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Comité de Protection des Personnes SUD-OUEST et Outre-Mer III. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

BM, CT, and SD contributed to conception and design of the study. AT and AL organized and processed imaging data to obtain JHU regional phenotypes described in the manuscript. AP and NB contributed to the QC of the image processing. AT and BM performed the statistical analysis and wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

The i-Share cohort has been funded by a grant ANR-10-COHO-05-01 as part of the Programme Investissements d'Avenir. Supplementary funding was received from the Conseil Régional of Nouvelle-Aquitaine, reference 4370420. The MRi-Share cohort and the ABACI software development have been supported by ANR-10-LABX-57 (TRAIL) and ANR-16-LCV2-0006 (GINESISLAB for the software) grants. Some regulatory and ethical aspects of MRi-Share have been supported by the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme under grant agreement no. 640643. AT, NB, and AL have been supported by a grant from the Fondation pour la Recherche Médicale (DIC202161236446) and AP by a grant ANR-15-HBPR-0001-03 (as part of the EU FLAG-ERA MULTI-LATERAL consortium). Additional support for AT and AL was provided by grant ANR-18-RHUS-002 (RHU SHIVA) as part of the Programme Investissements d'Avenir.

ACKNOWLEDGMENTS

We are indebted to Maxime Descoteaux (Sherbrooke University, Canada) for his help in implementing the DWI processing and QC pipelines.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnsys.2021.692152/full#supplementary-material

REFERENCES

Alexander, D. C., Dyrby, T. B., Nilsson, M., and Zhang, H. (2019). Imaging brain microstructure with diffusion MRI: practicality and applications. *NMR Biomed.* 32:e3841. doi: 10.1002/nbm.3841

Alfaro-Almagro, F., Jenkinson, M., Bangerter, N. K., Andersson, J. L. R., Griffanti, L., Douaud, G., et al. (2018). Image processing and quality control for the first 10,000 brain imaging datasets from UK Biobank. *Neuroimage* 166, 400–424. doi: 10.1016/j.neuroimage.2017.10.034

Andersson, J. L. R., Graham, M. S., Drobnjak, I., Zhang, H., and Campbell, J. (2018). Susceptibility-induced distortion that varies due to motion: correction in diffusion MR without acquiring additional data. *Neuroimage* 171, 277–295. doi: 10.1016/j.neuroimage.2017.12.040

Andersson, J. L. R., Graham, M. S., Drobnjak, I., Zhang, H., Filippini, N., and Bastiani, M. (2017). Towards a comprehensive framework for movement and distortion correction of diffusion MR images: within volume movement. *Neuroimage* 152, 450–466. doi: 10.1016/j.neuroimage.2017.02.085

Andersson, J. L. R., Graham, M. S., Zsoldos, E., and Sotiropoulos, S. N. (2016). Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. *Neuroimage* 141, 556–572. doi: 10.1016/j.neuroimage.2016.06.058

Andersson, J. L. R., Skare, S., and Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *Neuroimage* 20, 870–888. doi: 10.1016/S1053-8119(03)00336-7

Andersson, J. L. R., and Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* 125, 1063–1078. doi: 10.1016/j.neuroimage.2015.10.019

Ashburner, J., and Friston, K. J. (2005). Unified segmentation. *Neuroimage* 26, 839–851. doi: 10.1016/j.neuroimage.2005.02.018

Bach, M., Laun, F. B., Leemans, A., Tax, C. M. W., Biessels, G. J., Stieltjes, B., et al. (2014). Methodological considerations on tract-based spatial statistics (TBSS). *Neuroimage* 100, 358–369. doi: 10.1016/j.neuroimage.2014.06.021

Basser, P. J., Mattiello, J., and LeBihan, D. (1994). MR diffusion tensor spectroscopy and imaging. *Biophys. J.* 66, 259–267. doi: 10.1016/S0006-3495(94)80775-1

Bathelt, J., Johnson, A., Zhang, M., and Astle, D. E. (2019). The cingulum as a marker of individual differences in neurocognitive development. *Sci. Rep.* 9:2281. doi: 10.1038/s41598-019-38894-z

Bava, S., Jacobus, J., Thayer, R. E., and Tapert, S. F. (2013). Longitudinal changes in white matter integrity among adolescent substance users. *Alcohol. Clin. Exp. Res.* 37(Suppl. 1), E181–E189. doi: 10.1111/j.1530-0277.2012.01920.x

Beaudet, G., Tsuchida, A., Petit, L., Tzourio, C., Caspers, S., Schreiber, J., et al. (2020). Age-related changes of peak width skeletonized mean diffusivity (PSMD) across the adult lifespan: a multi-cohort study. *Front. Psychiatry* 11:342. doi: 10.3389/fpsyt.2020.00342

Beck, D., de Lange, A.-M. G., Maximov, I. I., Richard, G., Andreassen, O. A., Nordvik, J. E., et al. (2021). White matter microstructure across the adult lifespan: a mixed longitudinal and cross-sectional study using advanced diffusion models and brain-age prediction. *Neuroimage* 224, 117441. doi: 10.1016/j.neuroimage.2020.117441

Ben-Shachar, M., Lüdecke, D., and Makowski, D. (2020). effectsize: estimation of effect size indices and standardized parameters. *J. Open. Source Sci.* 5:2815. doi: 10.21105/joss.02815

Bettcher, B. M., Mungas, D., Patel, N., Elofson, J., Dutt, S., Wynn, M., et al. (2016). Neuroanatomical substrates of executive functions: beyond prefrontal structures. *Neuropsychologia* 85, 100–109. doi: 10.1016/j.neuropsychologia.2016.03.001

Billiet, T., Vandenbulcke, M., Mädler, B., Peeters, R., Dhollander, T., Zhang, H., et al. (2015). Age-related microstructural differences quantified using myelin water imaging and advanced diffusion MRI. *Neurobiol. Aging* 36, 2107–2121. doi: 10.1016/j.neurobiolaging.2015.02.029

Bonekamp, D., Nagae, L. M., Degaonkar, M., Matson, M., Abdalla, W. M. A., Barker, P. B., et al. (2007). Diffusion tensor imaging in children and adolescents: reproducibility, hemispheric, and age-related differences. *Neuroimage* 34, 733–742. doi: 10.1016/j.neuroimage.2006.09.020

Brouwer, R. M., Mandl, R. C. W., Schnack, H. G., van Soelen, I. L. C., van Baal, G. C., Peper, J. S., et al. (2012). White matter development in early puberty: a longitudinal volumetric and diffusion tensor imaging twin study. *PLoS One* 7:e32316. doi: 10.1371/journal.pone.0032316

Bubb, E. J., Metzler-Baddeley, C., and Aggleton, J. P. (2018). The cingulum bundle: anatomy, function, and dysfunction. *Neurosci. Biobehav. Rev.* 92, 104–127. doi: 10.1016/j.neubiorev.2018.05.008

Chang, Y. S., Owen, J. P., Pojman, N. J., Thieu, T., Bukshpun, P., Wakahiro, M. L. J., et al. (2015). White Matter changes of neurite density and fiber orientation dispersion during human brain maturation. *PLoS One* 10:e0123656. doi: 10.1371/journal.pone.0123656

Chenot, Q., Tzourio-Mazoyer, N., Rheault, F., Descoteaux, M., Crivello, F., Zago, L., et al. (2019). A population-based atlas of the human pyramidal tract in 410 healthy participants. *Brain Struct. Funct.* 224, 599–612. doi: 10.1007/s00429-018-1798-7

Choi, S., Kim, T., and Yu, W. (2009). "Performance evaluation of RANSAC family," in *Procedings of the British Machine Vision Conference 2009* (British Machine Vision Association), 81.1-81.12, London. doi: 10.5244/C.23.81

Colby, J. B., Soderberg, L., Lebel, C., Dinov, I. D., Thompson, P. M., and Sowell, E. R. (2012). Along-tract statistics allow for enhanced tractography analysis. *Neuroimage* 59, 3227–3242. doi: 10.1016/j.neuroimage.2011.11.004

Côté, M.-A., Girard, G., Boré, A., Garyfallidis, E., Houde, J.-C., and Descoteaux, M. (2013). Tractometer: towards validation of tractography pipelines. *Med. Image Anal.* 17, 844–857. doi: 10.1016/j.media.2013.03.009

Coupe, P., Manjon, J., Robles, M., and Collins, L. D. (2011). Adaptive multiresolution non-local means filter for 3d mr image denoising. *IET Image Process.*

Coupe, P., Yger, P., Prima, S., Hellier, P., Kervrann, C., and Barillot, C. (2008). An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. *IEEE Trans. Med. Imaging* 27, 425–441. doi: 10.1109/TMI.2007.906087

Cousineau, M., Jodoin, P.-M., Morency, F. C., Rozanski, V., Grand'Maison, M., Bedell, B. J., et al. (2017). A test-retest study on Parkinson's PPMI dataset yields statistically significant white matter fascicles. *Neuroimage Clin.* 16, 222–233. doi: 10.1016/j.nicl.2017.07.020

Cox, S. R., Ritchie, S. J., Tucker-Drob, E. M., Liewald, D. C., Hagenaars, S. P., Davies, G., et al. (2016). Ageing and brain white matter structure in 3,513 UK Biobank participants. *Nat. Commun.* 7:13629. doi: 10.1038/ncomms13629

Daducci, A., Canales-Rodríguez, E. J., Zhang, H., Dyrby, T. B., Alexander, D. C., and Thiran, J.-P. (2015). Accelerated microstructure imaging via convex optimization (AMICO) from diffusion MRI data. *Neuroimage* 105, 32–44. doi: 10.1016/j.neuroimage.2014.10.026

de Graaf-Peters, V. B., and Hadders-Algra, M. (2006). Ontogeny of the human central nervous system: what is happening when? *Early Hum. Dev.* 82, 257–266. doi: 10.1016/j.earlhumdev.2005.10.013

Dean, D. C., Planalp, E. M., Wooten, W., Adluru, N., Kecskemeti, S. R., Frye, C., et al. (2017). Mapping white matter microstructure in the one month human brain. *Sci. Rep.* 7:9759. doi: 10.1038/s41598-017-09915-6

Dekaban, A. S. (1978). Changes in brain weights during the span of human life: relation of brain weights to body heights and body weights. *Ann. Neurol.* 4, 345–356. doi: 10.1002/ana.410040410

Dimond, D., Heo, S., Ip, A., Rohr, C. S., Tansey, R., Graff, K., et al. (2020). Maturation and interhemispheric asymmetry in neurite density and orientation dispersion in early childhood. *Neuroimage* 221, 117168. doi: 10.1016/j.neuroimage.2020.117168

Dumontheil, I. (2016). Adolescent brain development. *Curr. Opin. Behav. Sci.* 10, 39–44. doi: 10.1016/j.cobeha.2016.04.012

Elvsåshagen, T., Norbom, L. B., Pedersen, P. Ø, Quraishi, S. H., Bjørnerud, A., Malt, U. F., et al. (2015). Widespread changes in white matter microstructure after a day of waking and sleep deprivation. *PLoS One* 10:e0127351. doi: 10.1371/journal.pone.0127351

Ferizi, U., Scherrer, B., Schneider, T., Alipoor, M., Eufracio, O., Fick, R. H. J., et al. (2017). Diffusion MRI microstructure models with *in vivo* human brain Connectome data: results from a multi-group comparison. *NMR Biomed.* 30:e3734. doi: 10.1002/nbm.3734

Fjell, A. M., Walhovd, K. B., Westlye, L. T., Østby, Y., Tamnes, C. K., Jernigan, T. L., et al. (2010). When does brain aging accelerate? Dangers of quadratic fits in cross-sectional studies. *Neuroimage* 50, 1376–1383. doi: 10.1016/j.neuroimage.2010.01.061

Forkel, S. J., Thiebaut de Schotten, M., Kawadler, J. M., Dell'Acqua, F., Danek, A., and Catani, M. (2014). The anatomy of fronto-occipital connections from early blunt dissections to contemporary tractography. *Cortex* 56, 73–84. doi: 10.1016/j.cortex.2012.09.005

Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., van der Walt, S., Descoteaux, M., et al. (2014). Dipy, a library for the analysis of diffusion MRI data. *Front. Neuroinformatics* 8:8. doi: 10.3389/fninf.2014.00008

Genc, S., Malpas, C. B., Holland, S. K., Beare, R., and Silk, T. J. (2017). Neurite density index is sensitive to age related differences in the developing brain. *Neuroimage* 148, 373–380. doi: 10.1016/j.neuroimage.2017.01.023

Giorgio, A., Watkins, K. E., Chadwick, M., James, S., Winmill, L., Douaud, G., et al. (2010). Longitudinal changes in grey and white matter during adolescence. *Neuroimage* 49, 94–103. doi: 10.1016/j.neuroimage.2009.08.003

Gogliettino, A. R., Potenza, M. N., and Yip, S. W. (2016). White matter development and tobacco smoking in young adults: a systematic review with recommendations for future research. *Drug Alcohol Depend.* 162, 26–33. doi: 10.1016/j.drugalcdep.2016.02.015

Hasan, K. M., Kamali, A., Abid, H., Kramer, L. A., Fletcher, J. M., and Ewing-Cobbs, L. (2010). Quantification of the spatiotemporal microstructural organization of the human brain association, projection and commissural pathways across the lifespan using diffusion tensor tractography. *Brain Struct. Funct.* 214, 361–373. doi: 10.1007/s00429-009-0238-0

Hasan, K. M., Sankar, A., Halphen, C., Kramer, L. A., Brandt, M. E., Juranek, J., et al. (2007). Development and organization of the human brain tissue compartments across the lifespan using diffusion tensor imaging. *Neuroreport* 18, 1735–1739. doi: 10.1097/WNR.0b013e3282f0d40c

Heilbronner, S. R., and Haber, S. N. (2014). Frontal cortical and subcortical projections provide a basis for segmenting the cingulum bundle: implications for neuroimaging and psychiatric disorders. *J. Neurosci.* 34, 10041–10054. doi: 10.1523/JNEUROSCI.5459-13.2014

Herting, M. M., Maxwell, E. C., Irvine, C., and Nagel, B. J. (2012). The impact of sex, puberty, and hormones on white matter microstructure in adolescents. *Cereb. Cortex* 22, 1979–1992. doi: 10.1093/cercor/bhr246

Hsu, J.-L., Van Hecke, W., Bai, C.-H., Lee, C.-H., Tsai, Y.-F., Chiu, H.-C., et al. (2010). Microstructural white matter changes in normal aging: a diffusion tensor imaging study with higher-order polynomial regression models. *Neuroimage* 49, 32–43. doi: 10.1016/j.neuroimage.2009.08.031

Iannone, R., Cheng, J., and Schloerke, B. (2020). *gt: Easily Create Presentation-Ready Display Tables*.

Inano, S., Takao, H., Hayashi, N., Abe, O., and Ohtomo, K. (2011). Effects of age and gender on white matter integrity. *AJNR Am. J. Neuroradiol.* 32, 2103–2109. doi: 10.3174/ajnr.A2785

Jelescu, I. O., Palombo, M., Bagnato, F., and Schilling, K. G. (2020). Challenges for biophysical modeling of microstructure. *J. Neurosci. Methods* 344:108861. doi: 10.1016/j.jneumeth.2020.108861

Jelescu, I. O., Veraart, J., Adisetiyo, V., Milla, S. S., Novikov, D. S., and Fieremans, E. (2015). One diffusion acquisition and different white matter models: how does microstructure change in human early development based on WMTI and NODDI? *Neuroimage* 107, 242–256. doi: 10.1016/j.neuroimage.2014.12.009

Jensen, J. H., and Helpern, J. A. (2010). MRI quantification of non-Gaussian water diffusion by kurtosis analysis. *NMR Biomed.* 23, 698–710. doi: 10.1002/nbm.1518

Jones, D. K., Knösche, T. R., and Turner, R. (2013). White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion MRI. *Neuroimage* 73, 239–254. doi: 10.1016/j.neuroimage.2012.06.081

Jones, D. K., Travis, A. R., Eden, G., Pierpaoli, C., and Basser, P. J. (2005). PASTA: pointwise assessment of streamline tractography attributes. *Magn. Reson. Med.* 53, 1462–1467. doi: 10.1002/mrm.20484

Kamiya, K., Hori, M., and Aoki, S. (2020). NODDI in clinical research. *J. Neurosci. Methods* 346:108908. doi: 10.1016/j.jneumeth.2020.108908

Kantarci, K., Senjem, M. L., Avula, R., Zhang, B., Samikoglu, A. R., Weigand, S. D., et al. (2011). Diffusion tensor imaging and cognitive function in older adults with no dementia. *Neurology* 77, 26–34. doi: 10.1212/WNL.0b013e31822313dc

Kessler, R. C., Amminger, G. P., Aguilar-Gaxiola, S., Alonso, J., Lee, S., and Ustün, T. B. (2007). Age of onset of mental disorders: a review of recent literature. *Curr. Opin. Psychiatry* 20, 359–364. doi: 10.1097/YCO.0b013e32816ebc8c

Klein, A., Andersson, J., Ardekani, B. A., Ashburner, J., Avants, B., Chiang, M.-C., et al. (2009). Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage* 46, 786–802. doi: 10.1016/j.neuroimage.2008.12.037

Kodiweera, C., Alexander, A. L., Harezlak, J., McAllister, T. W., and Wu, Y.-C. (2016). Age effects and sex differences in human brain white matter of young to middle-aged adults: a DTI, NODDI, and q-space study. *Neuroimage* 128, 180–192. doi: 10.1016/j.neuroimage.2015.12.033

Lakhani, B., Borich, M. R., Jackson, J. N., Wadden, K. P., Peters, S., Villamayor, A., et al. (2016). Motor skill acquisition promotes human brain myelin plasticity. *Neural Plast.* 2016:7526135. doi: 10.1155/2016/7526135

Lampinen, B., Szczepankiewicz, F., Mårtensson, J., van Westen, D., Sundgren, P. C., and Nilsson, M. (2017). Neurite density imaging versus imaging of microscopic anisotropy in diffusion MRI: a model comparison using spherical tensor encoding. *Neuroimage* 147, 517–531. doi: 10.1016/j.neuroimage.2016.11.053

Lebel, C., and Beaulieu, C. (2011). Longitudinal development of human brain wiring continues from childhood into adulthood. *J. Neurosci.* 31, 10937–10947. doi: 10.1523/JNEUROSCI.5302-10.2011

Lebel, C., and Deoni, S. (2018). The development of brain white matter microstructure. *Neuroimage* 182, 207–218. doi: 10.1016/j.neuroimage.2017.12.097

Lebel, C., Gee, M., Camicioli, R., Wieler, M., Martin, W., and Beaulieu, C. (2012). Diffusion tensor imaging of white matter tract evolution over the lifespan. *Neuroimage* 60, 340–352. doi: 10.1016/j.neuroimage.2011.11.094

Lebel, C., Treit, S., and Beaulieu, C. (2019). A review of diffusion MRI of typical white matter development from early childhood to young adulthood. *NMR Biomed.* 32:e3778. doi: 10.1002/nbm.3778

Lebel, C., Walker, L., Leemans, A., Phillips, L., and Beaulieu, C. (2008). Microstructural maturation of the human brain from childhood to adulthood. *Neuroimage* 40, 1044–1055. doi: 10.1016/j.neuroimage.2007.12.053

Lenroot, R. K., and Giedd, J. N. (2006). Brain development in children and adolescents: insights from anatomical magnetic resonance imaging. *Neurosci. Biobehav. Rev.* 30, 718–729. doi: 10.1016/j.neubiorev.2006.06.001

Lenth, R. V. (2021). *emmeans: Estimated Marginal Means, aka Least-Squares Means*.

Liu, X., Kinoshita, M., Shinohara, H., Hori, O., Ozaki, N., and Nakada, M. (2020). Does the superior fronto-occipital fascicle exist in the human brain? Fiber dissection and brain functional mapping in 90 patients with gliomas. *Neuroimage Clin.* 25:102192. doi: 10.1016/j.nicl.2020.102192

Lövdén, M., Bodammer, N. C., Kühn, S., Kaufmann, J., Schütze, H., Tempelmann, C., et al. (2010). Experience-dependent plasticity of white-matter microstructure extends into old age. *Neuropsychologia* 48, 3878–3883. doi: 10.1016/j.neuropsychologia.2010.08.026

Lynch, K. M., Cabeen, R. P., Toga, A. W., and Clark, K. A. (2020). Magnitude and timing of major white matter tract maturation from infancy through adolescence with NODDI. *Neuroimage* 212, 116672. doi: 10.1016/j.neuroimage.2020.116672

Mackey, A. P., Whitaker, K. J., and Bunge, S. A. (2012). Experience-dependent plasticity in white matter microstructure: reasoning training alters structural connectivity. *Front. Neuroanat.* 6:32. doi: 10.3389/fnana.2012.00032

Mah, A., Geeraert, B., and Lebel, C. (2017). Detailing neuroanatomical development in late childhood and early adolescence using NODDI. *PLoS One* 12:e0182340. doi: 10.1371/journal.pone.0182340

Maximov, I. I., Alnaes, D., and Westlye, L. T. (2019). Towards an optimised processing pipeline for diffusion magnetic resonance imaging data: effects of artefact corrections on diffusion metrics and their age associations in UK Biobank. *Hum. Brain Mapp.* 40, 4146–4162. doi: 10.1002/hbm.24691

Meola, A., Comert, A., Yeh, F.-C., Stefaneanu, L., and Fernandez-Miranda, J. C. (2015). The controversial existence of the human superior fronto-occipital fasciculus: connectome-based tractographic study with microdissection validation. *Hum. Brain Mapp.* 36, 4964–4971. doi: 10.1002/hbm.22990

Mori, S., Oishi, K., Jiang, H., Jiang, L., Li, X., Akhter, K., et al. (2008). Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *Neuroimage* 40, 570–582. doi: 10.1016/j.neuroimage.2007.12.035

Oishi, K., Zilles, K., Amunts, K., Faria, A., Jiang, H., Li, X., et al. (2008). Human brain white matter atlas: identification and assignment of common anatomical structures in superficial white matter. *Neuroimage* 43, 447–457. doi: 10.1016/j.neuroimage.2008.07.009

Olejnik, S., and Algina, J. (2003). Generalized eta and omega squared statistics: measures of effect size for some common research designs. *Psychol. Methods* 8, 434–447. doi: 10.1037/1082-989X.8.4.434

Pfefferbaum, A., and Sullivan, E. V. (2015). Cross-sectional versus longitudinal estimates of age-related changes in the adult brain: overlaps and discrepancies. *Neurobiol. Aging* 36, 2563–2567. doi: 10.1016/j.neurobiolaging.2015.05.005

Pines, A. R., Cieslak, M., Larsen, B., Baum, G. L., Cook, P. A., Adebimpe, A., et al. (2020). Leveraging multi-shell diffusion for studies of brain development in youth and young adulthood. *Dev. Cogn. Neurosci.* 43:100788. doi: 10.1016/j.dcn.2020.100788

Pohl, K. M., Sullivan, E. V., Rohlfing, T., Chu, W., Kwon, D., Nichols, B. N., et al. (2016). Harmonizing DTI measurements across scanners to examine the development of white matter microstructure in 803 adolescents of the NCANDA study. *Neuroimage* 130, 194–213. doi: 10.1016/j.neuroimage.2016.01.061

Powell, J. L., Parkes, L., Kemp, G. J., Sluming, V., Barrick, T. R., and García-Fiñana, M. (2012). The effect of sex and handedness on white matter anisotropy: a diffusion tensor magnetic resonance imaging study. *Neuroscience* 207, 227–242. doi: 10.1016/j.neuroscience.2012.01.016

Qian, W., Khattar, N., Cortina, L. E., Spencer, R. G., and Bouhrara, M. (2020). Nonlinear associations of neurite density and myelin content with age revealed using multicomponent diffusion and relaxometry magnetic resonance imaging. *Neuroimage* 223:117369. doi: 10.1016/j.neuroimage.2020.117369

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Raffelt, D., Tournier, J.-D., Fripp, J., Crozier, S., Connelly, A., and Salvado, O. (2011). Symmetric diffeomorphic registration of fibre orientation distributions. *Neuroimage* 56, 1171–1180. doi: 10.1016/j.neuroimage.2011.02.014

Reynolds, J. E., Grohs, M. N., Dewey, D., and Lebel, C. (2019). Global and regional white matter development in early childhood. *Neuroimage* 196, 49–58. doi: 10.1016/j.neuroimage.2019.04.004

Rheault, F., De Benedictis, A., Daducci, A., Maffei, C., Tax, C. M. W., Romascano, D., et al. (2020). Tractostorm: the what, why, and how of tractography dissection reproducibility. *Hum. Brain Mapp.* 41, 1859–1874. doi: 10.1002/hbm.24917

Rohlfing, T. (2013). Incorrect ICBM-DTI-81 atlas orientation and white matter labels. *Front. Neurosci.* 7:4. doi: 10.3389/fnins.2013.00004

Rosen, A. F. G., Roalf, D. R., Ruparel, K., Blake, J., Seelaus, K., Villa, L. P., et al. (2018). Quantitative assessment of structural image quality. *Neuroimage* 169, 407–418. doi: 10.1016/j.neuroimage.2017.12.059

Sampaio-Baptista, C., and Johansen-Berg, H. (2017). White matter plasticity in the adult brain. *Neuron* 96, 1239–1251. doi: 10.1016/j.neuron.2017.11.026

Schlegel, A. A., Rudelson, J. J., and Tse, P. U. (2012). White matter structure changes as adults learn a second language. *J. Cogn. Neurosci.* 24, 1664–1670. doi: 10.1162/jocn_a_00240

Schloerke, B., Cook, D., Larmarange, J., Briatte, F., Marbach, M., Thoen, E., et al. (2021). *GGally: Extension to "ggplot2"*.

Scholz, J., Klein, M. C., Behrens, T. E. J., and Johansen-Berg, H. (2009). Training induces changes in white-matter architecture. *Nat. Neurosci.* 12, 1370–1371. doi: 10.1038/nn.2412

Shiino, A., Chen, Y.-W., Tanigaki, K., Yamada, A., Vigers, P., Watanabe, T., et al. (2017). Sex-related difference in human white matter volumes studied: inspection of the corpus callosum and other white matter by VBM. *Sci. Rep.* 7:39818. doi: 10.1038/srep39818

Silveri, M. M., Dager, A. D., Cohen-Gilbert, J. E., and Sneider, J. T. (2016). Neurobiological signatures associated with alcohol and drug use in the human adolescent brain. *Neurosci. Biobehav. Rev.* 70, 244–259. doi: 10.1016/j.neubiorev.2016.06.042

Simmonds, D. J., Hallquist, M. N., Asato, M., and Luna, B. (2014). Developmental stages and sex differences of white matter and behavioral development through adolescence: a longitudinal diffusion tensor imaging (DTI) study. *Neuroimage* 92, 356–368. doi: 10.1016/j.neuroimage.2013.12.044

Singmann, H., Bolker, B., Westfall, J., Aust, F., and Ben-Shachar, M. S. (2021). *afex: Analysis of Factorial Experiments*.

Slater, D. A., Melie-Garcia, L., Preisig, M., Kherif, F., Lutti, A., and Draganski, B. (2019). Evolution of white matter tract microstructure across the life span. *Hum. Brain Mapp.* 40, 2252–2268. doi: 10.1002/hbm.24522

Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., et al. (2006). Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage* 31, 1487–1505. doi: 10.1016/j.neuroimage.2006.02.024

Song, S.-K., Yoshino, J., Le, T. Q., Lin, S.-J., Sun, S.-W., Cross, A. H., et al. (2005). Demyelination increases radial diffusivity in corpus callosum of mouse brain. *Neuroimage* 26, 132–140. doi: 10.1016/j.neuroimage.2005.01.028

Takao, H., Abe, O., Yamasue, H., Aoki, S., Sasaki, H., Kasai, K., et al. (2011). Gray and white matter asymmetries in healthy individuals aged 21-29 years: a voxel-based morphometry and diffusion tensor imaging study. *Hum. Brain Mapp.* 32, 1762–1773. doi: 10.1002/hbm.21145

Tamnes, C. K., Ostby, Y., Fjell, A. M., Westlye, L. T., Due-Tønnessen, P., and Walhovd, K. B. (2010). Brain maturation in adolescence and young adulthood: regional age-related changes in cortical thickness and white matter volume and microstructure. *Cereb. Cortex* 20, 534–548. doi: 10.1093/cercor/bhp118

Tamnes, C. K., Roalf, D. R., Goddings, A.-L., and Lebel, C. (2018). Diffusion MRI of white matter microstructure development in childhood and adolescence: methods, challenges and progress. *Dev. Cogn. Neurosci.* 33, 161–175. doi: 10.1016/j.dcn.2017.12.002

Telzer, E. H., Goldenberg, D., Fuligni, A. J., Lieberman, M. D., and Gálvan, A. (2015). Sleep variability in adolescence is associated with altered brain development. *Dev. Cogn. Neurosci.* 14, 16–22. doi: 10.1016/j.dcn.2015.05.007

Thiebaut de Schotten, M., Ffytche, D. H., Bizzi, A., Dell'Acqua, F., Allin, M., Walshe, M., et al. (2011). Atlasing location, asymmetry and inter-subject variability of white matter tracts in the human brain with MR diffusion tractography. *Neuroimage* 54, 49–59. doi: 10.1016/j.neuroimage.2010.07.055

Tournier, J.-D., Mori, S., and Leemans, A. (2011). Diffusion tensor imaging and beyond. *Magn. Reson. Med.* 65, 1532–1556. doi: 10.1002/mrm.22924

Tsuchida, A., Laurent, A., Crivello, F., Petit, L., Joliot, M., Pepe, A., et al. (2020). The MRi-Share database: brain imaging in a cross-sectional cohort of 1,870 university students. *bioRxiv* [Preprint]. doi: 10.1101/2020.06.17.154666

Tukey, J. W. (1977). *Exploratory Data Analysis*, 1st Edn. Reading, MA: Pearson.

Türe, U., Yaşargil, M. G., and Pait, T. G. (1997). Is there a superior occipitofrontal fasciculus? A microsurgical anatomic study. *Neurosurgery* 40, 1226–1232. doi: 10.1097/00006123-199706000-00022

Vos, S. B., Jones, D. K., Jeurissen, B., Viergever, M. A., and Leemans, A. (2012). The influence of complex white matter architecture on the mean diffusivity in diffusion tensor MRI of the human brain. *Neuroimage* 59, 2208–2216. doi: 10.1016/j.neuroimage.2011.09.086

Walhovd, K. B., Westlye, L. T., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., et al. (2011). Consistent neuroanatomical age-related volume differences across multiple samples. *Neurobiol. Aging* 32, 916–932. doi: 10.1016/j.neurobiolaging.2009.05.013

Warrington, S., Bryant, K. L., Khrapitchev, A. A., Sallet, J., Charquero-Ballester, M., Douaud, G., et al. (2020). XTRACT - Standardised protocols for automated tractography in the human and macaque brain. *Neuroimage* 217:116923. doi: 10.1016/j.neuroimage.2020.116923

Wassermann, D., Makris, N., Rathi, Y., Shenton, M., Kikinis, R., Kubicki, M., et al. (2016). The white matter query language: a novel approach for describing human white matter anatomy. *Brain Struct. Funct.* 221, 4705–4721. doi: 10.1007/s00429-015-1179-4

Wasserthal, J., Neher, P., and Maier-Hein, K. H. (2018). TractSeg–Fast and accurate white matter tract segmentation. *Neuroimage* 183, 239–253. doi: 10.1016/j.neuroimage.2018.07.070

Westlye, L. T., Walhovd, K. B., Dale, A. M., Bjørnerud, A., Due-Tønnessen, P., Engvig, A., et al. (2010). Life-span changes of the human brain white matter: diffusion tensor imaging (DTI) and volumetry. *Cereb. Cortex* 20, 2055–2068. doi: 10.1093/cercor/bhp280

Wickham, H. (2016). *ggplot2–Elegant Graphics for Data Analysis*, 2nd Edn. Cham: Springer International Publishing, doi: 10.1007/978-3-319-24277-4

Yeatman, J. D., Dougherty, R. F., Myall, N. J., Wandell, B. A., and Feldman, H. M. (2012). Tract profiles of white matter properties: automating fiber-tract quantification. *PLoS One* 7:e49790. doi: 10.1371/journal.pone.0049790

Yendiki, A., Panneck, P., Srinivasan, P., Stevens, A., Zöllei, L., Augustinack, J., et al. (2011). Automated probabilistic reconstruction of white-matter pathways in health and disease using an atlas of the underlying anatomy. *Front. Neuroinformatics* 5:23. doi: 10.3389/fninf.2011.00023

Zhang, H., Schneider, T., Wheeler-Kingshott, C. A., and Alexander, D. C. (2012). NODDI: practical *in vivo* neurite orientation dispersion and density imaging of the human brain. *Neuroimage* 61, 1000–1016. doi: 10.1016/j.neuroimage.2012.03.072

Zhang, H., Yushkevich, P. A., Alexander, D. C., and Gee, J. C. (2006). Deformable registration of diffusion tensor MR images with explicit orientation optimization. *Med. Image Anal.* 10, 764–785. doi: 10.1016/j.media.2006.06.004

Zhang, W., Olivi, A., Hertig, S. J., van Zijl, P., and Mori, S. (2008). Automated fiber tracking of human brain white matter using diffusion tensor imaging. *Neuroimage* 42, 771–777. doi: 10.1016/j.neuroimage.2008.04.241

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Systems Neuroscience

ORIGINAL RESEARCH
published: 30 September 2021
doi: 10.3389/fnsys.2021.725413

Immune-Related Genetic Overlap Between Regional Gray Matter Reductions and Psychiatric Symptoms in Adolescents, and Gene-Set Validation in a Translational Model

Lukas Penninck[1†], El Chérif Ibrahim[2†], Eric Artiges[1,3], Victor Gorgievski[4], Sylvane Desrivières[5], Severine Farley[4], Irina Filippi[1], Carlos E. A. de Macedo[4], Raoul Belzeaux[2,6], Tobias Banaschewski[7], Arun L. W. Bokde[8], Erin Burke Quinlan[5], Herta Flor[9,10], Antoine Grigis[11], Hugh Garavan[12], Penny Gowland[13], Andreas Heinz[14], Rüdiger Brühl[15], Frauke Nees[7,9,16], Dimitri Papadopoulos Orfanos[11], Tomáš Paus[17,18], Luise Poustka[19†], Juliane H. Fröhner[20], Michael N. Smolka[20], Henrik Walter[14], Robert Whelan[21], Julien Grenier[22], Gunter Schumann[5,23,24], Marie-Laure Paillère Martinot[1,25†], Eleni T. Tzavara[4,6,26†] and Jean-Luc Martinot[1†] for the IMAGEN Consortium*

[1] Institut National de la Santé et de la Recherche Médicale, INSERM U1299 "Trajectoires Développementales en Psychiatrie", Université Paris-Saclay, Ecole Normale Supérieure Paris-Saclay, CNRS, Centre Borelli, Gif-sur-Yvette, France, [2] Aix Marseille Univ, CNRS, INT, Inst Neurosci Timone, Marseille, France, [3] EPS Barthelemy Durand, Etampes, France, [4] University of Paris, CNRS, INCC, Paris, France, [5] Centre for Population Neuroscience and Precision Medicine (PONS), Institute of Psychiatry, Psychology & Neuroscience, SGDP Centre, King's College London, London, United Kingdom, [6] AP-HM, Hôpital Sainte Marguerite, Pôle de Psychiatrie Universitaire Solaris, Marseille, France, [7] Department of Child and Adolescent Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany, [8] Discipline of Psychiatry, School of Medicine and Trinity College Institute of Neuroscience, Trinity College Dublin, Dublin, Ireland, [9] Institute of Cognitive and Clinical Neuroscience, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany, [10] Department of Psychology, School of Social Sciences, University of Mannheim, Mannheim, Germany, [11] NeuroSpin, CEA, Université Paris-Saclay, Gif-sur-Yvette, France, [12] Department of Psychiatry and Psychology, University of Vermont, Burlington, VT, United States, [13] Sir Peter Mansfield Imaging Centre, School of Physics and Astronomy, University of Nottingham, Nottingham, United Kingdom, [14] Department of Psychiatry and Psychotherapy, CCM, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany, [15] Physikalisch-Technische Bundesanstalt (PTB), Braunschweig, Germany, [16] Institute of Medical Psychology and Medical Sociology, University Medical Center Schleswig Holstein, Kiel University, Kiel, Germany, [17] Department of Psychiatry, Faculty of Medicine and Centre Hospitalier Universitaire Sainte-Justine, University of Montreal, Montreal, QC, Canada, [18] Department of Psychology and Psychiatry, University of Toronto, Toronto, ON, Canada, [19] Department of Child and Adolescent Psychiatry and Psychotherapy, University Medical Centre Göttingen, Göttingen, Germany, [20] Department of Psychiatry and Neuroimaging Center, Technische Universität Dresden, Dresden, Germany, [21] School of Psychology and Global Brain Health Institute, Trinity College Dublin, Dublin, Ireland, [22] Université de Paris, INSERM UMRS 1124, Paris, France, [23] PONS Research Group, Department of Psychiatry and Psychotherapy, Humboldt University, Berlin and Leibniz Institute for Neurobiology, Magdeburg, Germany, [24] Institute for Science and Technology of Brain-inspired Intelligence (ISTBI), Fudan University, Shanghai, China, [25] AP-HP.Sorbonne Université, Department of Child and Adolescent Psychiatry, Pitié-Salpêtrière Hospital, Paris, France, [26] Fondation Fondamental, Créteil, France

OPEN ACCESS

Edited by:
Wen-Jun Gao, Drexel University, United States

Reviewed by:
Wang Jiesi, University of Chinese Academy of Sciences, China
Adriana Caballero, University of Illinois at Chicago, United States

***Correspondence:**
Jean-Luc Martinot
jean-luc.martinot@inserm.fr

[†] These authors have contributed equally to this work

Received: *15 June 2021*
Accepted: *31 August 2021*
Published: *30 September 2021*

Citation:
Penninck L, Ibrahim EC, Artiges E, Gorgievski V, Desrivières S, Farley S, Filippi I, de Macedo CEA, Belzeaux R, Banaschewski T, Bokde ALW, Quinlan EB, Flor H, Grigis A, Garavan H, Gowland P, Heinz A, Brühl R, Nees F, Papadopoulos Orfanos D, Paus T, Poustka L, Fröhner JH, Smolka MN, Walter H, Whelan R, Grenier J, Schumann G, Paillère Martinot M-L, Tzavara ET and Martinot J-L (2021) Immune-Related Genetic Overlap Between Regional Gray Matter Reductions and Psychiatric Symptoms in Adolescents, and Gene-Set Validation in a Translational Model. Front. Syst. Neurosci. 15:725413. doi: 10.3389/fnsys.2021.725413

Adolescence is a period of vulnerability for the maturation of gray matter (GM) and also for the onset of psychiatric disorders such as major depressive disorder (MDD), bipolar disorder and schizophrenia. Chronic neuroinflammation is considered to play a role in the etiology of these illnesses. However, the involvement of neuroinflammation in the observed link between regional GM volume reductions and psychiatric symptoms is not

established yet. Here, we investigated a possible common immune-related genetic link between these two phenomena in european adolescents recruited from the community. Hippocampal and medial prefrontal cortex (mPFC) were defined *a priori* as regions of interest (ROIs). Their GM volumes were extracted in 1,563 14-year-olds from the IMAGEN database. We found a set of 26 SNPs that correlated with the hippocampal volumes and 29 with the mPFC volumes at age 14. We formed two ROI-Related Immune-gene scores (RRI) with the inflammation SNPs that correlated to hippocampal GM volume and to mPFC GM volume. The predictive ability of both RRIs with regards to the presence of psychiatric symptoms at age 18 was investigated by correlating the RRIs with psychometric questionnaires obtained at age 18. The RRIs (but not control scores constructed with random SNPs) correlated with the presence of depressive symptoms, positive psychotic symptoms, and externalizing symptoms in later adolescence. In addition, the effect of childhood maltreatment, one of the major environmental risk factors for depression and other mental disorders, interacted with the RRI effect. We next sought to validate this finding by investigating our set of inflammatory genes in a translational animal model of early life adversity. Mice were subjected to a protocol of maternal separation at an early post-natal age. We evaluated depressive behaviors in separated and non-separated mice at adolescence and their correlations with the concomitant expression of our genes in whole blood samples. We show that in mice, early life adversity affected the expression of our set of genes in peripheral blood, and that levels of expression correlated with symptoms of negative affect in adolescence. Overall, our translational findings in adolescent mice and humans provide a novel validated gene-set of immune-related genes for further research in the early stages of mood disorders.

Keywords: immunity genes, psychiatric symptoms, adolescence, MRI, childhood maltreatment

INTRODUCTION

Some large-scale studies combining genetic and brain structural data hypothesize the existence of shared neurobiological mechanisms underlying prevalent psychiatric disorders, such as major depressive disorder (MDD), attention-deficit/hyperactivity disorder, and schizophrenia (Parker et al., 2020; Patel et al., 2021). The central findings supporting this hypothesis are the associations found between disorder-specific regional differences in brain structure (e.g., cortical thickness or regional volumes) and common clusters of genes involved in brain development or maturation. Although they do not establish causality, these observations point to the interplay of genetic and brain structural underpinnings in the pathophysiology of psychiatric illnesses. Here, we aim to further contribute to this endeavor by applying a targeted approach, i.e., by focusing on a limited number of genes and brain regions. The main advantage of such an approach is the improvement of statistical power to identify associations in smaller samples. Specifically, we will investigate the possible association between neuroinflammatory-related genes and regional gray matter (GM) volumes in the hippocampus and medial prefrontal cortex (mPFC) in the development of mood disorders (MDD and bipolar disorder) and schizophrenia.

Although a wide range of structural abnormalities has been associated with psychiatric disorders (e.g., Patel et al., 2021), herein we focused on volumetric GM measurements in the hippocampus and mPFC as regions of interest (ROI). Indeed, reduced hippocampal and PFC volume are among the most replicated findings in MRI studies of depression (MacQueen et al., 2003; Schmaal et al., 2020). Lower hippocampal volumes have been associated with adolescent onset MDD (Chen et al., 2010; Schmaal et al., 2016), and ventral medial PFC maturation has been related to negative affect in the developing brain (Ducharme et al., 2014). GM reductions in these two regions have been put forward as indicators of the severity and stage of MDD (Belleau et al., 2019). Still, in their review, Belleau et al. (2019) speculate that, although hippocampal and mPFC GM reductions have been associated with MDD, these reductions are neither necessary nor sufficient for inducing a depressive episode. Instead, these structural abnormalities should be regarded as intermediary effectors driving the progression and recurrence of depression. Consistently, our group has reported lower volumes in both regions as variables of interest for tackling irregular sleep habits paving the way to psychiatric symptoms in adolescents (Lapidaire et al., 2021).

Our second reduction in scope, i.e., focusing on a carefully selected set of immune-related genes rather than a genome-wide paradigm, is founded on research putting forward chronic neuroinflammation as a neurobiological characteristic of MDD driving GM loss (Kubera et al., 2011; Barnes et al., 2017). This theory, originally called the "inflammatory and neurodegenerative hypothesis of depression" by Maes et al. (2009), is based on multiple pieces of evidence. The first paper demonstrating a close connection between depression and the immune system was published in 1990 and found that MDD was often associated with a significantly higher number of activated T-cells (Maes et al., 1990). Since then, there have been many consistent findings of increased levels of proinflammatory cytokines in the cerebrospinal fluid (CSF) of patients with depression, the most prominent being interleukin-1 (IL-1), IL-2, IL-6, IL-8, IL-12, interferon-γ and tumor necrosis factor-α. In addition, elevations in peripheral blood concentrations of chemokines, adhesion molecules, acute phase proteins and inflammatory mediators such as prostaglandins have been observed. Lastly, depression could be induced by administrating cytokines (see the comprehensive reviews by Raison et al., 2006 and Miller and Raison, 2015).

To our knowledge, there is no report investigating the putative association between immune-related genetic variation and MDD-related hippocampal and mPFC GM reductions. In order to do so, we will combine the benefits of a candidate gene approach and a polygenic approach. More specifically, we will construct two "polygenic" scores using only inflammation-related common genetic polymorphisms: one in association with hippocampal GM volume and another in association with mPFC GM volume. Hence, these scores will be referred to as ROI-related immune-gene scores (RRI-scores). The IMAGEN database will be used to access genetic and T1 imaging data in 14-year-olds, an age of particular interest as confounding effects due to the use of medication can be considered minimal. A second notable asset of the IMAGEN database is the availability of follow-up data at the age of 18, including a wide range of psychometric data. We will use this follow-up to investigate both RRI-scores with regards to the participants' psychiatric symptoms later in adolescence. Thus, we will examine whether the putative association between neuroinflammatory genes and regional GM reductions plays a role in the development of psychiatric symptoms.

Moreover, we will test the translational hypothesis that genetic predisposition influences the capacity of an environmental risk factor to induce a psychiatric disorder (Caspi and Moffitt, 2006; Bagot et al., 2014). First, we will assess whether there is an interaction between the RRIs and the degree of childhood maltreatment (CM) explaining negative affects at adolescence, in the IMAGEN database. Second, we will use a translational approach employing an animal model of early life adversity. We will assess whether (i) mice subjected to early life adversity display depressive-like behaviors at adolescence; (ii) the expression of the constructed hippocampal RRI gene-set is altered in peripheral blood in these mice; (iii) transcript levels correlate with the severity of depression-related behavioral scores in adolescent mice. The advantage of this combined approach is that we use transcriptional profiling, which measures the expression of genes and is sensitive to both genotype and environment, to gain insight toward the (patho)physiological link between inflammatory pathways, childhood trauma, and depression symptoms in adolescence.

MATERIALS AND METHODS

Participants

Participants' datasets were drawn from the IMAGEN project, a European multi-center collaboration combining genetic, neuro-imaging and neuropsychological data from 2223 adolescents at baseline (BL; 14 years old). Participants were followed up 2 years (follow-up 1; FU1) and 4 years later (FU2). An initial sample of 1563 14-year-old adolescents was defined, for which genetic information, T1-weighted MRI images passing the different quality control procedures and the multiple necessary variables were available. In order to perform correlational analyses with psychometric measurements, subgroups of the initial sample were constructed with participants for whom the necessary psychometric data were available. Recruitment procedures have been previously described (Schumann et al., 2010). Written informed consent was obtained from all participants and their legal guardians and verbal assent was obtained from the adolescents.

Neuro-Imaging Data

T1-Weighted MRI

High-resolution T1-weighted anatomical MR images were obtained by means of three Tesla scanners (Philips, Siemens, and General Electric), using a standardized 3D T1-weighted magnetization prepared rapid acquisition gradient echo (MPRAGE) sequence based on the ADNI protocol.[1] The full details of the MRI acquisition protocols and quality checks have been described previously (Schumann et al., 2010). Image preprocessing was performed with Statistical Parametric Mapping 12 software (SPM12) and its toolbox extension Computational Anatomy Toolbox 12 software (CAT12). In summary, T1-weighted images were segmented and normalized using customized tissue probability maps. Then, the normalized, segmented, and modulated gray matter (GM) and white matter (WM) images were smoothed using a 8-mm full-width at half-maximum Gaussian kernel. Total GM, WM, and cerebrospinal fluid (CSF) volumes were computed for each participant. Total intracranial volume (TIV) was defined as the sum of GM, WM, and CSF volumes. Correct segmentation by CAT12 was verified through visual evaluation of the outliers determined by the automated quality control step "Check Sample Homogeneity" available in CAT12.

Extraction of Regional Gray Matter Volumes

The matlab-script "get_totals.m"[2] was used to extract the hippocampal and mPFC GM volumes from the baseline GM-segmented T1 MRI images. In order to do so, two masks were

[1] http://adni.loni.usc.edu/methods/mri-analysis/mri-acquisition/

[2] http://www0.cs.ucl.ac.uk/staff/g.ridgway/vbm/get_totals.m

designed (see **Supplementary Figure 1**) using WFU Pickatlas software (a SPM12 toolbox extension): a bilateral hippocampal mask, available in Pickatlas, and a mPFC mask, composed of the Brodmann areas (BA) 10, 11, 12, 14, 24, 25, 32, and 33. Here the medial prefrontal cortex was defined in its widest sense. For instance, BA11 was added because it pertains to the orbito-frontal cortex in its most medial part; BA10 pertains to the anterior prefrontal pole but it also includes a medial part.

Genetic Data

SNP Genotyping

The DNA purification and genotyping procedures implemented in the IMAGEN study have been previously described (Desrivières et al., 2015). Population homogeneity was verified with the Structure software using HapMap populations as reference groups (Pritchard et al., 2000). Further correction for population stratification through principal component analysis was deemed unnecessary. After the quality control measures, genotypic data for a total of 466 125 SNPs were considered. The software PLINK was used to extract the SNP genotypes.

Construction of RRI-Scores

Based on an extensive literature screening, using the keywords "chronic," "neuroinflammation," and "review" in PubMed, genes encoding direct and indirect contributors to neuroinflammation were characterized. To this end, every gene (or protein) that was found to be related to neuroinflammation according to at least three reviews was added to the list. Albeit not systematic, we did consider this procedure to be appropriate for this exploratory study. Including more genes by loosening the constraints might be worth exploring in future research, but will not necessarily lead to RRIs with higher predictive power. The SNPs in and around (±5 kb) the listed immune-related genes were obtained by means of the UCSC Genome Browser[3] and only those genotyped in the IMAGEN database were selected, a total of 674 SNPs. The methodology used in this study to construct the two RRIs (one associated with hippocampal GM volume, the other with mPFC GM volume) was based on the recent guide put forward by Choi et al. (2020). The effect of every SNP on either the hippocampal or the mPFC GM volume was assessed by performing linear regression analyses in R, using data from the initial sample and the standard *lm* function in R. The dependent variable was the BL GM volume of either the hippocampus or the mPFC, the independent variable was the major allele count for the SNP of interest (0 = homozygous for the minor allele, 1 = heterozygous, 2 = homozygous for the major allele). The regression was controlled for the covariates sex, Puberty Developmental Scale (PDS) score as a proxy of age, TIV and scanner type. Next, the SNPs that correlated ($p < 0.1$ without correction for multiple comparisons) with the hippocampal or mPFC GM volume were selected. In order to control for linkage disequilibrium (LD) and avoid redundancy in the SNPs included in the score, a manual procedure analogous to SNP pruning was carried out. More specifically, the online application LDmatrix developed by the National Institute of Health[4] was used to study the LD between the SNPs. Groups of SNPs that were found in high LD ($r^2 > 0.5$) were replaced by the most significantly correlated SNP of that group as representative, eliminating the other SNP(s) of that group. Lastly, for every participant, a hippocampal RRI (HRRI) and mPFC RRI (MRRI) were calculated based on the participant's genotype for the group of independent SNPs that were correlated with the hippocampal and mPFC GM volume, respectively. More precisely, the score was defined by the sum of minor alleles for the included SNPs, weighted by the effect size of those alleles individually. A normalization of the HRRI and MRRI values was performed in order to obtain two scores ranging from 0 to 10 that could easily be compared and combined.

Construction of Control Scores

A collection of 674 SNPs available in the IMAGEN database was randomly selected. Using an identical methodology as described above, a control score explaining hippocampal GM volume and a control score explaining mPFC GM volume were designed using these 674 random SNPs.

Questionnaire Data

Questionnaires

Five questionnaire measurements were extracted from the IMAGEN database. First, the algorithmically calculated scores (ranging from 0 to 5) representing the probability of depression according to the DSM-IV (referred to as DepBand) were obtained for participants at FU2 through the Development and Well-Being Assessment (DAWBA), a self-administered diagnostic questionnaire.[5] Second, the Community Assessment Psychic Experiences-42 (CAPE-42) questionnaire was used to obtain three scores evaluating the presence of depressive symptoms (referred to as the Depressive Dimension Score; DDS) as well as psychotic experiences, both positive (Positive Dimension Score; PDS) and negative (Negative Dimension Score; NDS) at FU2. Third, the self-reported Strengths and Difficulties Questionnaire (SDQ) at FU2 was used to construct two scores representing externalizing and internalizing behaviors; the Externalizing Score (ES) by summing the Conduct Problems Score and Hyperactivity Score, the Internalizing Score (IS) by summing the Emotional Symptoms Score and the Peer Problems Score. Fourth, a score representing childhood maltreatment (CM) was constructed based on information from the Childhood Trauma Questionnaire (CTQ). This retrospective recall-based questionnaire was administered to the participants at BL and, as described in the manual, produces a score (ranging from 0 to 4) representing the endured stress with regards to six categories. Participants were subsequently categorized in five groups based on the highest score in the six subscales. Ultimately, the three highest groups were merged, resulting in a CM score ranging from 0 to 2. Fifth, the Alcohol Use Disorders Identification Test (AUDIT) Score at FU2 was extracted from the database in order to be included as a covariate in linear regression analyses modeling FU2 GM volumes.

[3]https://genome.ucsc.edu

[4]https://ldlink.nci.nih.gov/?tab=ldmatrix

[5]http://www.dawba.info

Correlational Analyses

The above-mentioned psychometric measurements (DepBand, DDS, PDS, NDS, ES, and IS) were modeled separately in function of the HRRI, the MRRI and the sum of both scores (HMRRI), as well as the control scores. These regression analyses were performed in R. Since the dependent variables DepBand, ES and IS displayed probability distributions similar to a Poisson distribution, Poisson regression (the log-linear type of the generalized linear model) was opted for. The dependent variables DDS, PDS, and NDS were found normally distributed. However, a log-transformation of the dependent variable was implemented in order to correct for the positive skewness. All regression analyses were controlled for the covariates gender and CM. Also, the interaction between CM and the score was evaluated. *P*-values were corrected for multiple comparisons through the Benjamini-Hochberg false discovery rate (FDR) procedure, using the *p.adjust* function in R.

Animals

All experiments on mice were carried out according to policies on the care and use of laboratory animals of European Community legislation 2010/63/EU. The local Ethics Committee (Comité d'éthique en expérimentation animale Charles Darwin N°5) approved the protocols used in this study (protocol number 01486).

The mice were kept under standard conditions at 22 ± 1°C, and a 12-h light-dark cycle with food and water available *ad libitum*.

Maternal Separation/Maternal Stress Protocol

Pregnant dams (BALB/c Jico) were purchased from Centre d'Elevage Janvier, (Le Genest St Isle, France) to arrive in our facility 5 days before expected delivery. Dams and their respective litters were divided into two groups. The first group (MS, *Maternaly separated; n = 3 dams*) was subjected to maternal separation/maternal stress procedure; the second group (NS, *No Separation; n = 2 dams*) of dams and respective litters was kept in standard housing conditions as controls.

The maternal separation/maternal stress protocol was adapted from Franklin et al. (2010). The protocol combined (i) physical separation of the pups from the mother and among them; (ii) a short maternal stress at the end of the separation period; (iii) unpredictability regarding the timing of the separation and the maternal stressors.

For maternal separation the pups were placed in separate clean compartments inside a temperature- and humidity-controlled terrarium, to avoid any physical distress of the pups; the mother was placed in a clean novel cage with bedding, food, and water. Maternal separation lasted for 3 h and was applied once daily from post-natal day 1 (P1) to P14; the timing was unpredictable. Maternal stress was applied to the dam at the end of the 3 h separation period and consisted of one of the following: 20 min contention in a plastic perforated tube; 10 min forced swimming stress; 10 min tail-suspension stress. Stressors were alternated pseudorandomly. Both MS and NS groups were left undisturbed between P14 and P21 (weaning). At weaning the sex of the pups was determined (for the present cohort: 6 male and 17 female) and they were subsequently assigned to social groups of 3–4 mice per cage, composed of animals of same sex, and subjected to the same protocol (MS or NS), but from more than one dams to avoid litter effect. The sex ratio per group was, for MS: 4M/9F; and for NS: 3M/7F. A Fisher's exact test applied to these sex ratios is not statistically significant ($p > 0.999$).

Behavioral Characterization of the Pups at Late Adolescence

We evaluated behaviors associated with depression (anhedonia, anxiety) in the separated ($n = 13$) and non-separated ($n = 10$) pups at late adolescence (P52–59). Behavioral dimensions were assessed with the Sucrose preference (anhedonia; P52) and Dark-light tests (anxiety; P59).

Sucrose Preference

For the sucrose preference test mice were first habituated to drink from two graduated pipettes one filled with water, and the other with sucrose solution for 3 days, the side of the sucrose pipette being alternated each day. On day 4 and after an overnight (15 h) deprivation of water, the two pipettes were presented again; one was filled with water and the other with 4% sucrose. The water and sucrose solution consumed over a 3 h-period, were measured. The sucrose preference index is defined as (sucrose consumed)/(sucrose consumed + water consumed) × 100 (percentage index).

Dark-Light

The apparatus consisted of one box divided in two compartments, an illuminated one (30 × 20 × 20 cm), which is open, and a dark one (15 × 20 × 20 cm), which is covered with a lid. A small aperture (width of 5.5 cm and height of 7 cm) allows the mouse to freely move between the dark compartment and the illuminated one. At the beginning of the experiment, the mouse was placed in the illuminated box, facing the aperture. Time spent in the lit box was measured during a 10 min period.

Next, we performed a *Z*-normalization. For this an individual *z*-score was calculated for each test and for each animal as follows: *z*-score = [(individual data for observed parameter) - (mean of control group)]/(standard deviation of control group). For both sucrose preference and dark/light *z*-scores were multiplied by −1, as decreased sucrose preference and decreased time in lit compartment measure depressive/anxiety-like behaviors. For the computation of means and standard deviation of control groups, control groups were defined as NS mice; note that for control groups the mean of *z*-score by behavioral dimension is equal to zero. Subsequently, a global "depression-index" for each animal was calculated by averaging the *z*-scores of the two individual tests as previously described (Apazoglou et al., 2018).

Blood Collection

At the end of the behavioral evaluations (P60), 0.25 ml of blood was collected from the submandibular vein and stabilized with 1.3 ml RNAlater® solution (Life Technologies, Ambion,

Austin, TX). Mice were euthanized several months later by pentobarbital injection.

RNA Isolation

Total RNAs were purified from the blood using the Mouse RiboPure-Blood RNA isolation kit (Invitrogen), according to manufacturer's recommendations. After washings, total RNAs were eluted with 0.1 mM EDTA and were subsequently submitted to DNase treatment (DNA-free^{TM} kit, Life Technologies, Ambion, Austin, TX). RNA concentration was determined using a nanodrop ND-1000 spectrophotometer (Thermo Scientific, Waltham, MA). Three samples (one female MS/two male NS) did not provide enough quantities and qualities and were further excluded from the gene expression analysis.

Candidate mRNA Expression Quantification

850 ng of total RNA was reverse transcribed using the High-Capacity cDNA Reverse Transcription kit (Life Technologies, Applied Biosystems, Foster City, CA). 800 ng of the resulting cDNA were combined with a TaqMan® Fast advanced Master Mix (Thermo Fisher Scientific) and real-time PCR reactions were simultaneously run in triplicate in a thermocycler under the following conditions: 2 min at 50°C, 10 min at 92°C, 45 cycles of 1 s at 95°C, and 20 s at 60°C, in custom array microfluidic cards (Applied Biosystems, Pleasanton, CA) using a QuantStudio^{TM} 7 system and data collected using QuantStudio^{TM} Real-Time PCR Software v1.1 (Applied Biosystems). For each of the 16 candidate genes tested (related to the SNPs included in the HGPS), primer-probe sets were selected using the web portal of the manufacturer (Applied Biosystems, see **Supplementary Table 5**). In addition, *Rab5a* was universally used as a reference gene (Hervé et al., 2017). Raw *Ct* values were obtained with manual baseline settings on the ThermoFisher cloud RQ software (Applied Biosystems), and then the relative expression level of each mRNA was quantified by using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). In this method, each candidate gene is quantified relative to the expression of *Rab5a* and each amplification is also compared to a calibrator sample (the mean of the samples from the NS mice).

Statistics

Behavioral Evaluation

The *z*-scores for anhedonia, dark light, and the global depression-like index were analyzed separately. For each parameter, the comparison of two independent groups (MS vs. NS) involved Student's *t*-test that were performed using the STATISTICA software. The results are expressed as mean ± SEM (standard error).

Gene Expression

For gene expression comparisons, after observing non-homogeneous variances for each candidate gene in each subgroup (through the [R] function *levene_test*) and absence of normality of residuals (through the [R] function *shapiro_test*) of a parametric model (through the [R] function *lm*), a non-parametric (by permutation) equivalent of a two-way factorial ANOVA was performed through the [R] function *aovp* in the *lmPerm* library. When ANOVA effects were significant, multiple group comparisons for each gene were performed through the [R] function *pairwise.perm.t.test* in the *RVAideMemoire* library to provide FDR *p*-value adjustment.

Behavior and Gene Expression Correlations

Correlograms, allowing visualization of behavior and gene expression data into correlation matrices were implemented through the [R] function *corrplot* in the *corrplot* library and variables were ordered according to first principal components. Linear regressions with 95% confidence intervals were plotted through the [R] functions *ggplot, geom_point, and geom_smooth* (with "lm" method) in the *ggplot2* library.

RESULTS

Selection of Genes and SNPs

Supplementary Table 1 lists the 90 immune-related genes selected for this study and classifies them in 6 categories: "Cytokines and Cytokine Receptors" (43 genes), "Oxidative Stress Effectors" (4), "Monocytosis and Granulopoiesis" (14), "Inflammatory Signaling Pathway" (21), "Kynurenine Pathway" (5), and "Phospholipases" (3). For these 90 genes, 674 related common SNPs were found genotyped in the IMAGEN database.

The Single Effects of Immune-Related SNPs on Brain Structure

In our initial sample of 1563 14-year-old participants, the mean bilateral hippocampal GM volume was 1.20 ± 0.109 ml and the mean mPFC GM volume was 30.1 ± 3.42 ml. Other volumetric data from the 14-year-olds (BL) and 18-year-olds (FU2) can be found in **Supplementary Table 2**. Sex, PDS, TIV and scanner type were all found to be significantly reacted with GM volume ($p < 0.001$), regardless of the region. The individual effects of the selected 674 SNPs on BL hippocampal and mPFC GM volume were assessed through linear regression analyses, controlled for sex, PDS, TIV, and scanner type. No correlation surpassed the significance threshold of $p < 0.05$ after Bonferroni correction for multiple comparisons.

Construction of the RRIs

As described in the section "Materials and Methods," we constructed two scores, one explaining the hippocampal GM volume at BL (HRRI), the other explaining the mPFC GM volume at BL (MRRI), using only immune-related SNPs. We found 26 "independent" SNPs that were considerably correlated with the hippocampal volume and thus incorporated in the HRRI; 29 SNPs were combined in the MRRI (**Supplementary Tables 3, 4**).

Correlation of the RRIs With Psychometric Data

The first psychometric measurement of interest was the DepBand, representing depression probability at FU2 (**Table 1**). Controlling for the covariates sex and CM, DepBand, could

TABLE 1 | Correlations between depression probability (DAWBA score) at age 18 and the genetic scores.

	Independent variable	χ^2 (df)	*B*	*z*	*p*	*p (corr)*
DepBand at FU2	HRRI	1192 (857)	0.0207	0.770	0.441	0.721
	MRRI	1185 (857)	0.0764	2.811	0.00494	0.0296*
	HMRRI	1186 (857)	0.0477	2.518	0.0119	0.0536
	CM:HRRI	1180 (856)	0.122	3.57	0.000364	0.00327*
	CM:MRRI	1184 (856)	0.0271	0.735	0.463	0.556
	CM:HMRRI	1179 (856)	0.0679	2.80	0.00516	0.0310*

The DepBand score (DAWBA questionnaire) obtained at FU2 was correlated with the three RRIs (ROI-Related-Immune-gene-score) as well as the three interactions between the scores and childhood maltreatment (CM) by means of a Poisson regression (N = 861).
$\chi 2$ (df) = residual deviance (degrees of freedom), used for Chi-Squared goodness-of-fit test; B = unstandardized regression coefficient.
** $p < 0.05$ (FDR-corrected).*

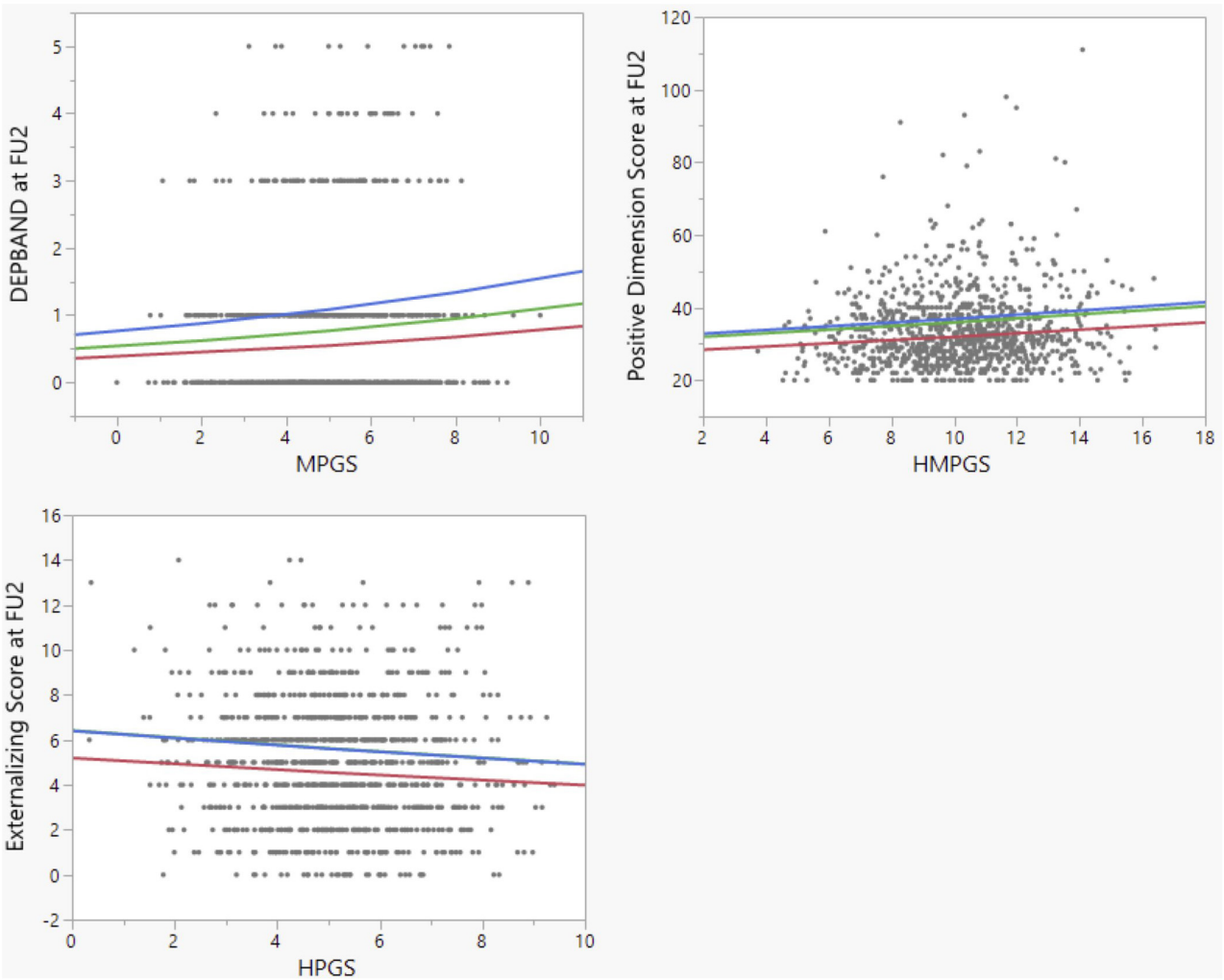

FIGURE 1 | Correlations between RRIs (determined from their link with GM regional volumes at age 14 as described in the text) and psychometric measurements at age 18. MPGS = MRRI, HPGS = HRRI, HMPGS = HMRRI. The colored lines are proposed regression lines for the different levels of childhood maltreatment (CM), after controlling for CM and sex. Red: CM = 0; green: CM = 1; blue: CM = 2. The blue and green line overlap in the bottom left graph.

not be correlated with the HRRI [*p(corr)* = 0.721]. The MRRI, however, positively and significantly covaried with DepBand [*p(corr)* = 0.0296; **Figure 1**]. Since both scores were not found redundant ($p = 0.436$), we also created the HMRRI by adding up the HRRI and MRRI. The HMRRI positively covaried with DepBand [*p(corr)* = 0.0536] but did not survive FDR-correction. Next, the interactions between CM and the scores in relation to DepBand were evaluated. The number of participants with

a CM score of 0 was 588, 283 had a score of 1, and 108 had a score of 2. A positive and significant interaction was observed between the HRRI and childhood maltreatment score CM [$p(corr) = 0.00327$]. This suggests that, as the level of endured trauma increases, the association between MRRI and presence of depressive symptoms at age 18 increases as well. It is important to note that all goodness-of-fit Chi-Squared tests for the Poisson regressions were found significant, suggesting that the data do not fit the model perfectly well. Alternative models were considered but not found better.

The second psychometric tool of interest was the CAPE-42, which consists out of three subscores (**Table 2**). A positive and significant correlation was observed between the log-transformed Positive Dimension Score, representing the amount of positive psychotic symptoms at FU2, and the HMRRI [$p(corr) = 0.0296$; **Figure 1**]. As no significant interactions were found between CM and the RRIs in explaining one of the three CAPE-42 subscores, these results are not shown in **Table 2**.

Lastly, the Externalizing Score (ES) and Internalizing Score (IS) at FU2 were modeled in function of the scores, again using Poisson regression (**Table 3**). A significant inverse correlation was observed between the ES and the HRRI [$p(corr) = 0.0294$; **Figure 1**]. We also found a significant positive interaction between HRRI and CM in explaining the ES [$p(corr) = 0.000857$].

Correlation of the Control Scores With Psychometric Data

No significant correlations were found between the control scores and the different psychiatric symptom measurements. In addition, we did not observe any significant interactions between childhood maltreatment and the control scores in function of the psychiatric symptoms.

Behavioral Effects of Early Life Stress in Adolescent Animals

Supplementary Figures 2, 3 as well as **Figure 2** show that early life stress profoundly affects anhedonia and anxiety at late adolescence in mice. At late adolescence, mice subjected to early life stress between P1 and P14 show increased anhedonia and anxiety as compared to NS mice (raw data are presented in **Supplementary Figures 2, 3** and z-score in **Figure 2**). The global depression-index (**Figure 2**) is also significantly elevated in MS mice indicating that early life adversity (P1–P14) induces long-lasting negative affects still present in late adolescence (P52–P59).

Effect of Early Life Stress on Candidate Gene Expression in Adolescent Animals

We next sought to investigate the effect of early life stress on the transcriptional expression of HRRI set of genes. Among the 17 candidate genes, 3 were not analyzed further (*IKBKG*, *IL12B*, *IL13*) because of low quality results. The remaining 14 candidate genes were well expressed in mouse blood and subsequently constituted the focus of our analysis as "mouse HRRI." **Figure 3** shows that maternal separation has a significant effect on the global transcript level of mouse HRRI ($p < 2.0E\text{-}16$). Individually, *post-hoc* analysis demonstrated significant dysregulation for *Ikbkb*, *Il10ra*, *Il10rb*, *Il18*, *Pla2g6,* and *Ptgs1*. Among those *Ptgs1* was increased while all the others were decreased in MS mice.

Correlations Between Gene Expression and Behavior in Adolescent Mice

Figure 4 shows that among the 14 mouse HRRI genes profiled, 6 were correlated with the global depression-index (a composite score of anhedonia and anxiety subscores). We noted that all the genes that correlated with behavior, also highly correlated with each other. **Figure 5** illustrates the tight link between one of the most significant altered genes, *Il10rb*, and depressive-like behavior. Regression analyses for behavior and *Ikbkb*, *Il10rb*, *Il18*, *Pla2g6,* and *Ptgs1* are shown in **Supplementary Figure 4**.

DISCUSSION

In this translational study, we explored the link between inflammation-related genes and brain structure, along with early life adversity and emergence of psychiatric symptoms. We investigated imaging genetics in a large database of community adolescents. We formed scores aggregating inflammation genes related with hippocampal or mPFC volumes at age 14. We found

TABLE 2 | Correlations between the clinical dimensions (CAPE-42 subscores) and the genetic scores.

	Independent variable	R^2	*B*	*t*	*p*	*p (corr)*
Positive Dimension (log)	HRRI	0.0449	0.0115	2.01	0.0444	0.114
	MRRI	0.0466	0.0134	2.38	0.0175	0.0630
	HMRRI	0.0506	0.0124	3.11	0.00192	0.0296*
Negative Dimension (log)	HRRI	0.0431	−0.00887	−1.06	0.289	0.520
	MRRI	0.0420	0.0087	0.226	0.821	0.921
	HMRRI	0.0423	−0.00342	−0.583	0.560	0.840
Depressive Dimension (log)	HRRI	0.091	0.000741	0.091	0.928	0.928
	MRRI	0.0933	0.0124	1.54	0.123	0.277
	HMRRI	0.0952	0.00667	1.161	0.246	0.492

The three subscores of the CAPE-42 questionnaire obtained at FU2 were correlated with the three RRIs after log-transformation of the dependent variables (N = 930).
R^2 = coefficient of determination; B = unstandardized regression coefficient.
** $p < 0.05$ (FDR-corrected).*

TABLE 3 | Correlations between Externalizing and Internalizing Score (SDQ) at age 18 and the genetic scores.

	Independent variable	χ^2 (df)	*B*	*z*	*p*	*p (corr)*
Externalizing at FU2	HRRI	1563 (951)	−0.0275	−2.90	0.00368	0.0294*
	MRRI	1572 (951)	−0.00262	−0.282	0.778	0.921
	HMRRI	1567 (951)	−0.0146	−2.22	0.0263	0.079
	CM:HRRI	1547 (950)	0.0538	4.07	4.76×10^{-5}	0.000857*
	CM:MRRI	1572 (950)	−0.00832	−0.610	0.542	0.588
	CM:HMRRI	1561 (950)	0.0228	2.44	0.01473	0.0529
Internalizing at FU2	HRRI	1824 (951)	−0.00465	−0.471	0.637	0.844
	MRRI	1825 (951)	−0.00158	−0.164	0.870	0.921
	HMRRI	1824 (951)	−0.00305	−0.445	0.657	0.845
	CM:HRRI	1819 (950)	0.0298	2.21	0.0269	0.0807
	CM:MRRI	1824 (950)	−0.00118	−0.0850	0.932	0.932
	CM:HMRRI	1822 (950)	0.0143	1.51	0.132	0.237

The Externalizing and Internalizing Score (SDQ) obtained at FU2 were correlated with the three RRIs as well as the three interactions between the RRIs and childhood maltreatment (CM) by means of a Poisson regression (N = 955).
χ2 (df) = residual deviance (degrees of freedom), used for Chi-Squared goodness-of-fit test; B = unstandardized regression coefficient.
** $p < 0.05$ (FDR-corrected).*

that these RRIs related with psychiatric symptoms at age 18. Also, we found a set of inflammation genes that were related to gray matter volume of hippocampal regions, and to childhood maltreatment score in these adolescents. Expression levels of inflammation genes associated with psychiatric symptoms (genes from HRRI) were subsequently examined in a murine model of early life adversity. Interestingly, expression of these genes was significantly altered after maternal separation in mice.

Excessive activation of the immune system as well as abnormalities in brain structure have been associated with depression. Furthermore, depression as well as its associated structural abnormalities are relatively heritable. Meta-analyses have found a heritability of 37% for depression (Sullivan et al., 2000) and a heritability between 40 and 70% for hippocampal volume (Gu and Kanai, 2014). Yet, characterization of the contributing individual genetic factors has proven to be very difficult. In a large genome wide association (GWA) study that investigated how common genetic variants affect the structure of subcortical regions, only a couple of genetic loci could be significantly correlated (Hibar et al., 2015). Similarly, despite considerable success within other illnesses such as diabetes and rheumatoid arthritis, GWA analyses of MDD have overall failed to produce results at the SNP level (Bogdan et al., 2017). Explanations for this could be found in the phenotypic heterogeneity, the lack of very large sample sizes and the complex functional architecture of the genetic polymorphisms. This last issue has been addressed by exploring gene-environment interactions and applying polygenic approaches. For example, in a recent study, structural abnormalities in schizophrenia were explored by creating a polygenic risk score (PRS) based on the weighted effects of SNPs found associated with schizophrenia in a prior GWA study (Alnæs et al., 2019). Also, PRSs for bipolar disorder and PRSs for schizophrenia both were found to have certain predictive power with regards to depression, corroborating prior evidence that these disorders share some common genetic overlap (Musliner et al., 2019).

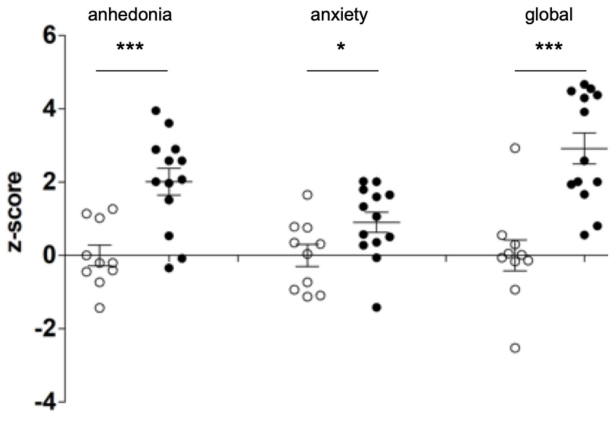

FIGURE 2 | Early life stress induces depressive-like behaviors in adolescent mice. Newborn mice were either subjected to a maternal separation paradigm between P1 (post-natal day 1) and P14 (MS, black dots) or were left undisturbed (controls, NS, white dots). NS and MS mice were evaluated for anhedonia and anxiety (measured in the sucrose preference and the dark-light tests, respectively) in late adolescence, between P52 and P59. Anhedonia and anxiety scores were *z*-transformed and a composite depression-index (global *z*-score) was averaged. Two-tailed Student's *t*-test shows increased levels of anhedonia ($df = 21$; $t = 4.155$) and anxiety ($df = 21$; $t = 2.202$) as well as increased depression index ($df = 21$, $t = 4.797$) in MS mice as compared to NS. * $p < 0.05$; *** $p < 0.0001$.

Since psychiatric disorders as well as their associated structural abnormalities seem to involve a genetic contribution, we hypothesized the existence of an immune-related genetic overlap between GM structural reductions and psychiatric symptoms. This was explored by constructing two RRIs based solely on SNPs related to inflammatory genes: one predicting hippocampal GM volume in 14-year-olds (HRRI), the other predicting mPFC GM volume (MRRI). We found that both scores were correlated with the presence of different symptoms later in adolescence.

We observed that the RRIs describing the genetic variation in less than 30 inflammatory SNPs had small yet significant

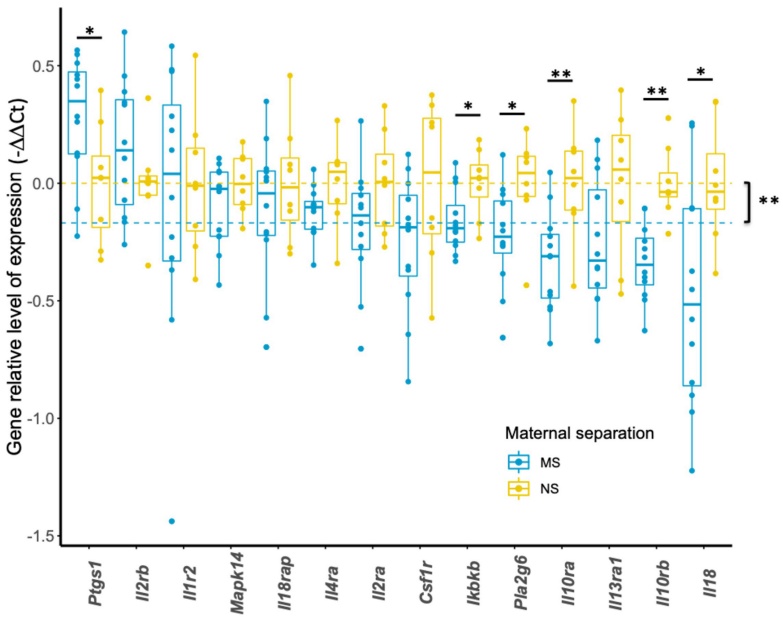

FIGURE 3 | Transcriptional expression of HRRI candidate genes in adolescent mice. Total RNA from blood of NS and MS mice at P60 was profiled by RT-qPCR for 14 HRRI candidate gene transcripts. The expression of each transcript was quantified relative to the expression of a reference gene, *Rab5a*, whereas the mean of NS mice was used as a calibrator. Statistical analysis was realized using a permutation-based non-parametric factorial ANOVA. Yellow and blue dashed lines indicate the mean values of all NS and MS, respectively, and highlight a significant "gene" effect between NS and MS animals ($p = 0.00295$). Significant multigroup comparisons for each gene were performed by pairwise permutation *t*-tests and are as follows, *Ptgs1* ($p = 0.022$), *Ikbkb* ($p = 0.028$), *Pla2g6* ($p = 0.022$), *Il10ra* ($p = 0.008$), *Il10rb* ($p = 0.002$), *Il18* ($p = 0.032$). * $p < 0.05$; ** $p < 0.01$.

predictive power regarding certain psychometric measurements obtained at age 18. As expected, effect sizes were consistently relatively small. The MRRI was found correlated with the presence of depressive symptoms. Secondly, both HRRI and MRRI were correlated with positive psychotic symptoms. Schizophrenia, the main disorder linked with these symptoms, has been consistently associated with neuroanatomical abnormalities such as reductions in GM volume (Weinberger, 1987; Keshavan et al., 2005; Bakhshi and Chance, 2015). Thirdly, we found a negative relationship between the HRRI and externalizing symptoms at age 18. This seems unexpected at first sight. A recent review addressing neuro-imaging findings in two of the major externalizing disorders, conduct disorder and oppositional defiant disorder, did not describe any studies reporting increases of hippocampal GM volume (Noordermeer et al., 2016). However, functional deficiencies in the amygdala, common in externalizing disorders, could be explained by abnormalities in the neighboring hippocampal complex (Yang and Wang, 2017). Lastly, we observed positive interactions between the scores and childhood maltreatment. This means that the probability of developing certain psychiatric symptoms due to a history of childhood maltreatment will be larger in the context of a specific genetic predisposition, in this case a high RRI. Gene-environment interactions in psychiatric diseases have been described repeatedly. For example, a polymorphism in the promoter region of the serotonin transporter gene was reported to moderate the influence of stressful life events on depression (Caspi et al., 2003).

The ability of the RRIs to predict to a limited extent the presence of psychiatric symptoms suggests the existence of the proposed genetic overlap. Indeed, the same variation in immune-related genes was found to explain both GM volume reductions in the hippocampus or mPFC and the degree of certain psychiatric symptoms. However, it could be argued that the RRIs predictive ability is solely due to the fact that the RRIs are constructed in such a way that they represent a portion of the GM volume variance. As the link between GM volumes and psychiatric illnesses is already established, the ability of an alternative score representing those structural abnormalities to predict psychiatric symptoms would not be surprising. In order to investigate this, we performed a control study by constructing two scores on the basis of random SNPs. These control scores did not significantly correlate with any psychometric measurement, nor did they display any interactions with childhood maltreatment. This higher predictive power of the non-random RRIs points at the involvement of the immune system. We thus not only corroborate prior evidence for the link between structural GM reductions and psychiatric illnesses, but also provide pioneering evidence strongly suggesting an immune-related genetic overlap between regional GM volumes and psychiatric symptoms, and define a novel combination of genes involved in this link. In order to further investigate this suggested causality, it would be interesting to perform a longitudinal study in which brain structural changes during adolescence are associated with the development of psychiatric symptoms.

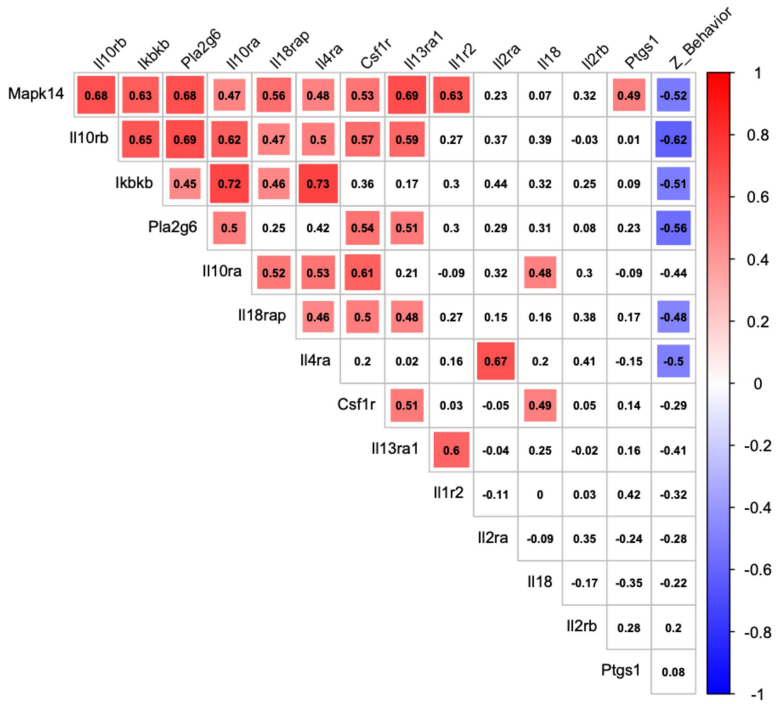

FIGURE 4 | Correlation between transcriptional expression of HRRI candidate genes and depressive-like behavior in adolescent mice. Correlation matrix graph, correlogram, highlighting correlation between qualitative (Z_Behavior, a depression-index reflecting a composite score of anhedonia and anxiety subscores) and quantitative (transcriptional relative level of HRRI candidate genes obtained by RT-qPCR from blood samples) variables in adolescent mice. The variables are ordered according to first principal components. Positive correlations are displayed in red and negative correlations in blue color. The intensity of the color and the size of the squares are proportional to the Pearson correlation coefficients. Only significant correlations are indicated by a colored square.

To further investigate the functional importance of the novel gene-set that we defined we employed a translational approach linking genotype and gene expression analyses. Indeed, the phenotype during adolescence is likely to be modulated by both genotype and environment, so that genotype analyses alone probably cannot account for their interaction (Kent et al., 2012). In contrast, transcriptional profiling that measures the expression of genes is sensitive to both genotype and environment and therefore may offer insights in pathophysiology. We focused on blood transcriptomics, since blood signature demonstrated that it could represent a surrogate for brain gene expression and may predict stress-induced behaviors (Sullivan et al., 2006; van Heerden et al., 2009; Rollins et al., 2010; Tylee et al., 2013; Witt et al., 2013; Luykx et al., 2016; Hervé et al., 2017).

We implemented an animal model of early life adversity and measured depression-like behaviors in adolescence. Based on the above reported effects in human subjects we hypothesized that early life adversity would affect not only behavior but also the expression of our set of genes, and that expression would correlate with symptoms of negative affects at adolescence.

Mice were subjected to a protocol of early life adversity (maternal separation) at an early post-natal age (P1–P15). We evaluated behaviors associated with depression (anhedonia: sucrose preference; anxiety: dark-light box) in MS and NS mice at adolescence (P52–P59). We measured the expression of our genes in whole blood samples collected at the same time-point (P60). We specifically focused on the hippocampal gene-set since the hippocampus is a region consistently implicated in depression and depression-like phenotypes in humans and mice (Vythilingam et al., 2002; Turecki et al., 2012; Apazoglou et al., 2018).

Our results showed that mice subjected to early life adversity displayed negative affects at adolescence (**Figure 2**). The expression of our mouse HRRI gene-set was altered in mice subjected to early life adversity (**Figure 3**), and transcript levels inside this gene-set correlated with depression-related behavioral score at adolescence (**Figure 4**).

Within the examined gene-set, **Figure 3** shows a significant decrease for *Ikbkb, Il10ra, Il10rb, Il18, Pla2g6,* and an increase for *Ptgs1* transcripts.

Figure 4 shows that among these genes, three (*Ikbkb, Il10rb, Pla2g6)* were significantly correlated with the global depression-index in MS mice. Notably the three also highly correlated with each other.

Interestingly, *Ikbkb, Il10rb, and Pla2g6*, which were decreased in MS mice are implicated in inflammatory homeostasis. *Ikbkb* is a regulator of the canonical NF-Kappa-B pathway a key-pathway in immunity/inflammation (Schmid and Birbach, 2008); *Il10rb* encodes for the anti-inflammatory cytokine IL10 (Shouval et al., 2014), and mice invalidated for IL10 show increased depressive-like behaviors (Mesquita et al., 2008); *Pla2g6* encodes for the iPLA2β protein, which regulates an overall anti-inflammatory response and whose dysregulation is associated with neurogenerative disorders (Guo et al., 2018).

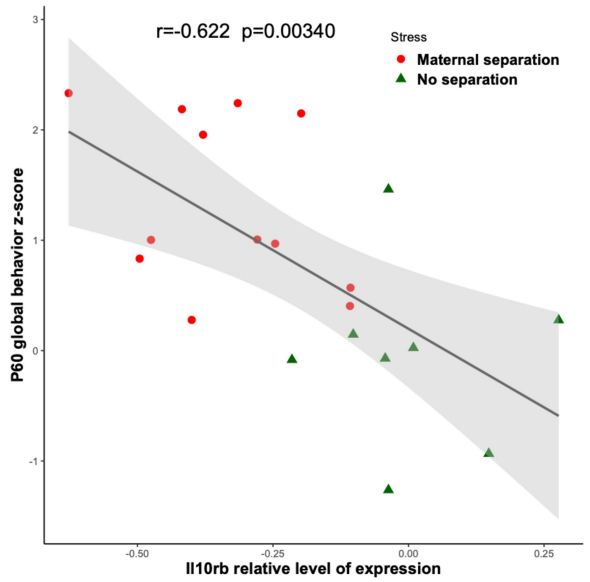

FIGURE 5 | Link between *Il10rb* relative level of transcriptional expression and depressive-like behavior. Linear regressions with 95% confidence intervals (in gray) are plotted between the depression-index and the *Il10rb* transcriptional relative level obtained by RT-qPCR from blood in MS (red circles) and NS (green triangles) adolescent mice. The Pearson correlation coefficient and the associated *p*-value are indicated.

The present findings suggest an inflammatory network of genes that most likely is involved in "depression-associated" neuroinflammatory adaptations in the periphery and CNS. We propose that early stressors like adversity can trigger an imbalance between anti-inflammatory and pro-inflammatory transcripts that may be at the origin of psychiatric symptoms in adolescence. These transcripts might provide both clinical biomarkers and novel targets in understanding and preventing individual developmental trajectories of psychiatric vulnerability.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: The datasets used for analysis can be acceded through requests to the last two authors. Requests to access these datasets should be directed to J-LM and ET, Jean-luc.martinot@ inserm.fr and eleni.tzavara@inserm.fr.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics committee CPP IDF 7, le Kremlin Bicètre, France. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. The animal study was reviewed and approved by the Comité d'éthique en expérimentation animale Charles Darwin N°5, Institut du Cerveau et de la Moelle Epinière, Paris.

AUTHOR CONTRIBUTIONS

LPe, J-LM, ET, EI, and SD: study conception. J-LM, ET, RBe, EQ, and AH: study administration. IF, VG, TB, AB, HG, RBr, PG, FN, TP, MS, HW, RW, and JG: human study data acquisition. DP, AG, LPo, and JF: human study data QC. VG, EI, SF, CdM, and JG: mouse data acquisition. LPe, EI, EA, and ET: data analysis. LPe, EI, M-LP, ET, and J-LM: manuscript writing. M-LP, ET, EI, and J-LM: manuscript reviewing and editing. All authors contributed to the article and approved the submitted version.

FUNDING

This work received support from the following sources: the European Union-funded FP6 Integrated Project IMAGEN (Reinforcement-related behaviour in normal brain function and psychopathology) (LSHM-CT- 2007-037286), the Horizon 2020 funded ERC Advanced Grant "STRATIFY" (Brain network based stratification of reinforcement-related disorders) (695313), Human Brain Project (HBP SGA 2, 785907, and HBP SGA 3, 945539), the Medical Research Council Grant "c-VEDA" (Consortium on Vulnerability to Externalizing Disorders and Addictions) (MR/N000390/1), the National Institute of Health (NIH) (R01DA049238, A decentralized macro and micro gene-by-environment interaction analysis of substance use behavior and its brain biomarkers), the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London, the Bundesministeriumfür Bildung und Forschung (BMBF grants 01GS08152; 01EV0711; Forschungsnetz AERIAL 01EE1406A, 01EE1406B; Forschungsnetz IMAC-Mind 01GL1745B), the Deutsche Forschungsgemeinschaft (DFG grants SM 80/7-2, SFB 940, TRR 265, NE 1383/14-1), the Medical Research Foundation and Medical Research Council (grants MR/R00465X/1 and MR/S020306/1), the National Institutes of Health (NIH) funded ENIGMA (grants 5U54EB020403-05 and 1R56AG058854-01). Further support was provided by grants from: – the ANR (ANR-12-SAMA-0004, AAPG2019 – GeBra, ANR-19-NEUR-ANTaRES), the Eranet Neuron (AF12-NEUR0008-01 – WM2NA; and ANR-18-NEUR00002-01 – ADORe), the Fondation de France (00081242), the Fondation pour la Recherche Médicale (DPA20140629802), the Mission Interministérielle de Lutte-contre-les-Drogues-et-les-Conduites-Addictives (MILDECA), the Assistance-Publique-Hôpitaux-de-Paris and INSERM (interface grant), Paris Sud University IDEX 2012, the Fondation de l'Avenir (grant AP-RM-17-013), the Fédération pour la Recherche sur le Cerveau; the National Institutes of Health, Science Foundation Ireland (16/ERCD/3797), United States (Axon, Testosterone and Mental Health during Adolescence; RO1 MH085772-01A1), and by NIH Consortium grant U54 EB020403, supported by a cross-NIH alliance that funds Big Data to Knowledge Centres of Excellence.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnsys.2021.725413/full#supplementary-material

REFERENCES

Alnæs, D., Kaufmann, T., van der Meer, D., Coìrdova-Palomera, A., Rokicki, J., Moberget, T., et al. (2019). Brain heterogeneity in schizophrenia and its association with polygenic risk. *JAMA Psychiatry* 76, 739–748.

Apazoglou, K., Farley, S., Gorgievski, V., Belzeaux, R., Lopez, J. P., Grenier, J., et al. (2018). Antidepressive effects of targeting ELK-1 signal transduction. *Nat. Med.* 24, 591–597.

Bagot, R. C., Labontè, B., Penþa, C. J., and Nestler, E. J. (2014). Epigenetic signaling in psychiatric disorders: stress and depression. *Dialogues Clin. Neurosci.* 16, 281–295. doi: 10.31887/dcns.2014.16.3/rbagot

Bakhshi, K., and Chance, S. A. (2015). The neuropathology of schizophrenia: a selective review of past studies and emerging themes in brain structure and cytoarchitecture. *Neuroscience* 303, 82–102. doi: 10.1016/j.neuroscience.2015.06.028

Barnes, J., Mondelli, V., and Pariante, C. M. (2017). Genetic contributions of inflammation to depression. *Neuropsychopharmacology* 42, 81–98. doi: 10.1038/npp.2016.169

Belleau, E. L., Treadway, M. T., and Pizzagalli, D. A. (2019). The impact of stress and major depressive disorder on hippocampal and medial prefrontal cortex morphology. *Biol. Psychiatry* 85, 443–453. doi: 10.1016/j.biopsych.2018.09.031

Bogdan, R., Salmeron, B. J., Carey, C. E., Agrawal, A., Calhoun, V. D., Garavan, H., et al. (2017). Imaging genetics and genomics in psychiatry: a critical review of progress and potential. *Biol. Psychiatry* 82, 165–175. doi: 10.1016/j.biopsych.2016.12.030

Caspi, A., and Moffitt, T. E. (2006). Gene-environment interactions in psychiatry: joining forces with neuroscience. *Nat. Rev. Neurosci.* 7, 583–590. doi: 10.1038/nrn1925

Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H., et al. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 301, 386–389. doi: 10.1126/science.1083968

Chen, M. C., Hamilton, J. P., and Gotlib, I. H. (2010). Decreased hippocampal volume in healthy girls at risk of depression. *Arch. Gen. Psychiatry* 67, 270–276. doi: 10.1001/archgenpsychiatry.2009.202

Choi, S. W., Mak, T. S. H., and O'Reilly, P. F. (2020). Tutorial: a guide to performing polygenic risk score analyses. *Nat. Protoc.* 15, 2759–2772. doi: 10.1038/s41596-020-0353-1

Desrivières, S., Lourdusamy, A., Tao, C., Toro, R., Jia, T., Loth, E., et al. (2015). Single nucleotide polymorphism in the neuroplastin locus associates with cortical thickness and intellectual ability in adolescents. *Mol. Psychiatry* 20, 263–274. doi: 10.1038/mp.2013.197

Ducharme, S., Albaugh, M. D., Hudziak, J. J., Botteron, K. N., Nguyen, T. V., Truong, C., et al. (2014). Anxious/depressed symptoms are linked to right ventromedial prefrontal cortical thickness maturation in healthy children and young adults. *Cereb. Cortex* 11, 2941–2950.

Franklin, T. B., Russig, H., Weiss, I. C., Gräff, J., Linder, N., Michalon, A., et al. (2010). Epigenetic transmission of the impact of early stress across generations. *Biol. Psychiatry* 68, 408–415. doi: 10.1016/j.biopsych.2010.05.036

Gu, J., and Kanai, R. (2014). What contributes to individual differences in brain structure? *Front. Hum. Neurosci.* 8:262. doi: 10.3389/fnhum.2014.00262

Guo, Y. P., Tang, B. S., and Guo, J. F. (2018). PLA2G6-associated neurodegeneration (PLAN): review of clinical phenotypes and genotypes. *Front. Neurol.* 9:1100. doi: 10.3389/fneur.2018.01100

Hervé, M., Bergon, A., Le Guisquet, A. M., Leman, S., Consoloni, J. L., Fernandez-Nunez, N., et al. (2017). Translational identification of transcriptional signatures of major depression and antidepressant response. *Front. Mol. Neurosci.* 10:248. doi: 10.3389/fnmol.2017.00248

Hibar, D. P., Stein, J. L., Renteria, M. E., Arias-Vasquez, A., Desrivières, S., Jahanshad, N., et al. (2015). Common genetic variants influence human subcortical brain structures. *Nature* 520, 224–229.

Kent, J. W. Jr., Göring, H. H., Charlesworth, J. C., Drigalenko, E., Diego, V. P., Curran, J. E., et al. (2012). Genotype×age interaction in human transcriptional ageing. *Mech. Ageing Dev.* 133, 581–590.

Keshavan, M. S., Berger, G., Zipursky, R. B., Wood, S. J., and Pantelis, C. (2005). Neurobiology of early psychosis. *Br. J. Psychiatry (Supp.)* 48, s8–s18.

Kubera, M., Obuchowicz, E., Goehler, L., Brzeszcz, J., and Maes, M. (2011). In animal models, psychosocial stress-induced (neuro)inflammation, apoptosis and reduced neurogenesis are associated to the onset of depression. *Prog. Neuropsychopharmacol. Biol. Psychiatry.* 35, 744–759. doi: 10.1016/j.pnpbp.2010.08.026

Lapidaire, W., Urrila, A. S., Artiges, E., Miranda, R., Vulser, H., Bézivin-Frere, P., et al. (2021). IMAGEN consortium. Irregular sleep habits, regional grey matter volumes, and psychological functioning in adolescents. *PLoS One* 16:e0243720. doi: 10.1371/journal.pone.0243720

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(−Delta Delta C(T)) method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Luykx, J. J., Olde Loohuis, L. M., Neeleman, M., Strengman, E., Bakker, S. C., Lentjes, E., et al. (2016). Peripheral blood gene expression profiles linked to monoamine metabolite levels in cerebrospinal fluid. *Transl. Psychiatry* 6:e983. doi: 10.1038/tp.2016.245

MacQueen, G. M., Campbell, S., McEwen, B. S., Macdonald, K., Amano, S., Joffe, R. T., et al. (2003). Course of illness, hippocampal function, and hippocampal volume in major depression. *Proc. Natl. Acad. Sci. U.S.A.* 100, 1387–1392. doi: 10.1073/pnas.0337481100

Maes, M., Bosmans, E., Suy, E., Vandervorst, C., De Jonckheere, C., and Raus, J. (1990). Immune disturbances during major depression: upregulated expression of interleukin-2 receptors. *Neuropsychobiology* 24, 115–120. doi: 10.1159/000119472

Maes, M., Yirmyia, R., Noraberg, J., Brene, S., Hibbeln, J., Perini, G., et al. (2009). The inflammatory & neurodegenerative (I&ND) hypothesis of depression: leads for future research and new drug developments in depression. *Metab. Brain Dis.* 24, 27–53.

Mesquita, A. R., Correia-Neves, M., Roque, S., Castro, A. G., Vieira, P., Pedrosa, J., et al. (2008). IL-10 modulates depressive-like behavior. *J. Psychiatr. Res.* 43, 89–97. doi: 10.1016/j.jpsychires.2008.02.004

Miller, A., and Raison, C. (2015). The role of inflammation in depression: from evolutionary imperative to modern treatment target. *Nat. Rev. Immunol.* 16, 22–34. doi: 10.1038/nri.2015.5

Musliner, K. L., Mortensen, P. B., McGrath, J. J., Suppli, N. P., Hougaard, D. M., Bybjerg-Grauholm, J., et al. (2019). Association of polygenic liabilities for major depression, bipolar disorder, and schizophrenia with risk for depression in the Danish population. *JAMA Psychiatry* 76, 516–525. doi: 10.1001/jamapsychiatry.2018.4166

Noordermeer, S. D., Luman, M., and Oosterlaan, J. (2016). A systematic review and meta-analysis of neuroimaging in oppositional defiant disorder (ODD) and conduct disorder (CD) taking attention-deficit hyperactivity disorder (ADHD) Into Account. *Neuropsychol. Rev.* 26, 44–72. doi: 10.1007/s11065-015-9315-8

Parker, N., Patel, Y., Jackowski, A. P., Pan, P. M., Salum, G. A., Pausova, Z., et al. (2020). Assessment of neurobiological mechanisms of cortical thinning during childhood and adolescence and their implications for psychiatric disorders. *JAMA Psychiatry* 77, 1127–1136. doi: 10.1001/jamapsychiatry.2020.1495

Patel, Y., Parker, N., Shin, J., Howard, D., French, L., Thomopoulos, S. I., et al. (2021). Virtual histology of cortical thickness and shared neurobiology in 6 psychiatric disorders. *JAMA Psychiatry* 78, 47–63.

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959. doi: 10.1093/genetics/155.2.945

Raison, C. L., Capuron, L., and Miller, A. H. (2006). Cytokines sing the blues: inflammation and the pathogenesis of depression. *Trends Immunol.* 27, 24–31. doi: 10.1016/j.it.2005.11.006

Rollins, B., Martin, M. V., Morgan, L., and Vawter, M. P. (2010). Analysis of whole genome biomarker expression in blood and brain. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 153b, 919–936.

Schmaal, L., Pozzi, E., CHo, T., van Velzen, L. S., Veer, I. M., Opel, N., et al. (2020). ENIGMA MDD: seven years of global neuroimaging studies of major depression through worldwide data sharing. *Transl. Psychiatry* 10:172.

Schmaal, L., Veltman, D. J., van Erp, T. G., Samann, P. G., Frodl, T., Jahanshad, N., et al. (2016). Subcortical brain alterations in major depressive disorder: findings from the ENIGMA major depressive disorder working group. *Mol. Psychiatry.* 21, 806–812.

Schmid, J. A., and Birbach, A. (2008). IkappaB kinase beta (IKKbeta/IKK2/IKBKB)–a key molecule in signaling to the transcription

factor NF-kappaB. *Cytokine Growth Factor Rev.* 19, 157–165. doi: 10.1016/j.cytogfr.2008.01.006

Schumann, G., Loth, E., Banaschewski, T., Barbot, A., Barker, G., Buüchel, C., et al. (2010). The IMAGEN study: reinforcement-related behaviour in normal brain function and psychopathology. *Mol. Psychiatry* 15, 1128–1139. doi: 10.1038/mp.2010.4

Shouval, D. S., Biswas, A., Goettel, J. A., Mccann, K., Conaway, E., Redhu, N. S., et al. (2014). Interleukin-10 receptor signaling in innate immune cells regulates mucosal immune tolerance and anti-inflammatory macrophage function. *Immunity* 40, 706–719. doi: 10.1016/j.immuni.2014.03.011

Sullivan, P. F., Fan, C., and Perou, C. M. (2006). Evaluating the comparability of gene expression in blood and brain. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 141b, 261–268. doi: 10.1002/ajmg.b.30272

Sullivan, P. F., Neale, M. C., and Kendler, K. S. (2000). Genetic epidemiology of major depression: review and meta-analysis. *Am. J. Psychiatry* 157, 1552–1562. doi: 10.1176/appi.ajp.157.10.1552

Turecki, G., Ernst, C., Jollant, F., Labonté, B., and Mechawar, N. (2012). The neurodevelopmental origins of suicidal behavior. *Trends Neurosci.* 35, 14–23. doi: 10.1016/j.tins.2011.11.008

Tylee, D. S., Kawaguchi, D. M., and Glatt, S. J. (2013). On the outside, looking in: a review and evaluation of the comparability of blood and brain "-omes". *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 162b, 595–603. doi: 10.1002/ajmg.b.32150

van Heerden, J. H., Conesa, A., Stein, D. J., Montaner, D., Russell, V., and Illing, N. (2009). Parallel changes in gene expression in peripheral blood mononuclear cells and the brain after maternal separation in the mouse. *BMC Res. Notes* 2:195. doi: 10.1186/1756-0500-2-195

Vythilingam, M., Heim, C., Newport, J., Miller, A. H., Anderson, E., Bronen, R., et al. (2002). Childhood trauma associated with smaller hippocampal volume in women with major depression. *Am. J. Psychiatry* 159, 2072–2080. doi: 10.1176/appi.ajp.159.12.2072

Weinberger, D. R. (1987). Implications of normal brain development for the pathogenesis of schizophrenia. *Arch. Gen. Psychiatry* 44, 660–669. doi: 10.1001/archpsyc.1987.01800190080012

Witt, S. H., Sommer, W. H., Hansson, A. C., Sticht, C., Rietschel, M., and Witt, C. C. (2013). Comparison of gene expression profiles in the blood, hippocampus and prefrontal cortex of rats. *Silico Pharmacol.* 1:15.

Yang, Y., and Wang, J. Z. (2017). From structure to behavior in basolateral amygdala-hippocampus circuits. *Front. Neural Circuits* 11:86. doi: 10.3389/fncir.2017.00086

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Neuroscience

TECHNOLOGY AND CODE
published: 16 November 2021
doi: 10.3389/fnins.2021.668293

A Practical Guide to Sparse *k*-Means Clustering for Studying Molecular Development of the Human Brain

*Justin L. Balsor[1], Keon Arbabi[1], Desmond Singh[2], Rachel Kwan[2], Jonathan Zaslavsky[2], Ewalina Jeyanesan[1] and Kathryn M. Murphy[1,2]**

[1] McMaster Neuroscience Graduate Program, McMaster University, Hamilton, ON, Canada, [2] Department of Psychology, Neuroscience and Behavior, McMaster University, Hamilton, ON, Canada

Studying the molecular development of the human brain presents unique challenges for selecting a data analysis approach. The rare and valuable nature of human postmortem brain tissue, especially for developmental studies, means the sample sizes are small (n), but the use of high throughput genomic and proteomic methods measure the expression levels for hundreds or thousands of variables [e.g., genes or proteins (p)] for each sample. This leads to a data structure that is high dimensional ($p \gg n$) and introduces the curse of dimensionality, which poses a challenge for traditional statistical approaches. In contrast, high dimensional analyses, especially cluster analyses developed for sparse data, have worked well for analyzing genomic datasets where $p \gg n$. Here we explore applying a lasso-based clustering method developed for high dimensional genomic data with small sample sizes. Using protein and gene data from the developing human visual cortex, we compared clustering methods. We identified an application of sparse *k*-means clustering [robust sparse *k*-means clustering (RSKC)] that partitioned samples into age-related clusters that reflect lifespan stages from birth to aging. RSKC adaptively selects a subset of the genes or proteins contributing to partitioning samples into age-related clusters that progress across the lifespan. This approach addresses a problem in current studies that could not identify multiple postnatal clusters. Moreover, clusters encompassed a range of ages like a series of overlapping waves illustrating that chronological- and brain-age have a complex relationship. In addition, a recently developed workflow to create plasticity phenotypes (Balsor et al., 2020) was applied to the clusters and revealed neurobiologically relevant features that identified how the human visual cortex changes across the lifespan. These methods can help address the growing demand for multimodal integration, from molecular machinery to brain imaging signals, to understand the human brain's development.

Keywords: human brain, development, clustering, synaptic proteins, transcriptomic data, high dimension and low sample size, sparsity-based algorithm

OPEN ACCESS

Edited by:
Tomáš Paus,
Université de Montréal, Canada

Reviewed by:
Leon French,
Centre for Addiction and Mental Health (CAMH), Canada
Shaojie Ma,
Yale University, United States

***Correspondence:**
Kathryn M. Murphy
kmurphy@mcmaster.ca

Specialty section:
This article was submitted to Neurogenomics, a section of the journal Frontiers in Neuroscience

Received: *15 February 2021*
Accepted: *30 September 2021*
Published: *16 November 2021*

Citation:
Balsor JL, Arbabi K, Singh D, Kwan R, Zaslavsky J, Jeyanesan E and Murphy KM (2021) A Practical Guide to Sparse k-Means Clustering for Studying Molecular Development of the Human Brain. Front. Neurosci. 15:668293. doi: 10.3389/fnins.2021.668293

INTRODUCTION

As molecular tools have become integrated with human neuroscience, there has been a renewed interest in mapping human brain development. Many studies have compared molecular changes among age groups (Law et al., 2003; Duncan et al., 2010; Pinto et al., 2010; Kang et al., 2011; Siu et al., 2015, 2017; Zhu et al., 2018) using distinct life-span stages that developmentalists have described based on physical, cognitive, and psychosocial maturation (Sigelman and Rider, 2017). However, age-binning assumes that those stages are a good fit for molecular development of the brain. In contrast, other areas of human neuroscience are applying data-driven approaches such as principal component analysis (PCA) (Bray, 2017) or unsupervised clustering (Lebenberg et al., 2018) to identify age-related changes in brain development. Applying cluster analysis to studying the molecular development of the human brain is challenging because of the limited availability of developmental postmortem tissue samples. Nevertheless, clustering algorithms have been developed for high dimensional biological datasets that have a small sample size (n) but measurements from many molecular features (p) (e.g., genes or proteins). Here we apply one of those approaches, sparse k-means clustering (Witten and Tibshirani, 2010; Kondo et al., 2016), to illustrate a data-driven approach for studying brain development that uses the expression of many genes or proteins to partition samples into age-related clusters. Then we show that clustering can identify aspects of human visual cortex development that are not apparent in typical developmental ontologies.

Cellular and molecular findings from postmortem brain tissue are used as benchmarks for linking age-related changes in non-invasive brain imaging signals with the underlying neurobiology. For example, many imaging studies reference synaptic development measurements (Huttenlocher and Dabholkar, 1997) to account for rapid changes in cerebral cortex MRI signals during the first few years of life. More recently, gene expression databases have been used to identify candidate cellular and molecular features, such as those underlying cortical thinning throughout the life-span (Vidal-Pineiro et al., 2020) or testosterone-related structural properties of the adolescent cerebral cortex (Liao et al., 2021). However, the rare and valuable nature of human postmortem brain samples means that gene expression studies have small sample sizes, especially compared to modern MRI studies that use a population neuroscience approach and aggregate data from hundreds or thousands of subjects (Paus, 2016). The issue of sample size is especially critical for brain development, as even well-established tissue banks (e.g., NIH NeuroBioBank) have fewer than 250 samples for most age groups and fewer than 50 for key ages of child development. Finally, the labor-intensive nature of molecular techniques means that studies can only use a subset of the available samples [e.g., (Pinto et al., 2010) $n = 28$; (Kang et al., 2011) $n = 57$; (Siu et al., 2015, 2017) $n = 30$; (Zhu et al., 2018) $n = 26$]. Nevertheless, the high dimensional data collected by molecular studies provide a wealth of information about how the brain changes across the life-span.

Although MRI and postmortem studies of human brain development face different methodological challenges, they share many analytical approaches. Both rely on analyses from the high dimensional toolbox to uncover information relevant to the complexities of brain development. Differences in experimental design, however, place distinct constraints on those analyses. High throughput molecular tools have significantly increased the amount of information obtained from each postmortem sample, generating long lists of gene or protein expression values. Those values represent a vector that describes where each sample exists in a high dimensional space that captures the molecular complexity of human brain development. However, the large number of measurements but small number of samples means that the high dimensional space is sparse with points spread virtually equidistantly across the space. The challenge is to determine how samples cluster together in that sparse space and if those data-driven clusters reflect stages of human development.

Cluster analysis is not new in biology (Eisen et al., 1998; Tamayo et al., 1999; Hastie et al., 2000, 2001), but applying it to postmortem studies of human brain development presents unique problems because of the small sample sizes of those studies. When standard clustering techniques have been used to study gene expression changes in human brain development, clusters are found for regional and prenatal versus postnatal groups, but distinct postnatal clusters matching developmental stages have not been reported (Colantuoni et al., 2011; Kang et al., 2011; Carlyle et al., 2017; Li et al., 2018; Zhu et al., 2018; Disorder et al., 2021). Accordingly, it has been challenging to link cognitive, perceptual, or social-emotional stages and prolonged development found using brain imaging with the underlying maturation of molecular mechanisms in the human brain.

Here we provide a practical guide to sparse clustering that focuses on overcoming the small sample size problem to reveal postnatal patterns of molecular development in the human brain. We introduce sparsity-based clustering, and one approach in particular, sparse k-means clustering, developed to address the problem of datasets with a large number of observations from proteins or genes (p) but a small number of samples (n) resulting in a data structure that is $p \gg n$ (Witten and Tibshirani, 2010). Finally, we illustrate the value of applying clustering by interrogating the neurobiological features of the clusters to reveal new aspects of the developing human visual cortex.

Challenges Clustering Small Sample Sizes

Currently, transcriptomic, proteomic, and other omics datasets of human brain development include measurements of many molecular features from a small number of samples. The combinatorial nature of those data makes it challenging to use traditional statistical comparisons to understand the many molecular changes that occur in the developing brain. Instead, high dimensional analyses that use all of the data are needed to classify the biological features that differentiate the human brain across the lifespan. However, even when clustering is used, the complexity of the findings can still be challenging to interpret, and studies may need to group the data into predefined

age categories to describe the spatiotemporal dynamics of the developing brain (Li et al., 2018).

In the mathematical notation used for clustering algorithms, the genes or proteins are called features or observations and are represented by p, while the number of samples is represented by n. Most human brain development datasets are either $p \approx n$ or $p > n$ and are best described as high dimensional datasets with more features than samples. When clustering those data, algorithms can *borrow strength* from the large number of features that represent each sample in high dimensional space. However, if only a subset of the features contributes to partitioning the samples into clusters, then the analyses may run into the *curse of dimensionality* (Bellman, 1983). For brain development, this means that developmentally relevant features may become obscured as more and more genes or proteins that do not contribute to developmental changes are included in the dataset. A central problem in analyzing these $p > n$ datasets is to identify the molecular features associated with age-related clusters from a very large set of candidate genes. Two approaches for focusing on relevant features include either preprocessing the data using dimension reduction methods (e.g., PCA, tSNE) or using sparsity-based clustering algorithms that retain all of the features but subset or reweight them during clustering (see **Supplementary Material**).

Some of the common approaches to unsupervised dimension reduction and clustering often used in neuroscience, like PCA and tSNE, can effectively separate data points into clusters in low-dimensional space, especially if there are large differences in features that fall on orthogonal sets of dimensions. For example, tSNE analysis of transcriptomic data identified separate clusters for cortical and cerebellar development (Kang et al., 2011; Carlyle et al., 2017), and PCA has shown that age can explain a large fraction of the variation in protein expression during cortical development (Pinto et al., 2015; Breen et al., 2018). Some of these approaches represent linear combinations of genes or proteins, and focus on reducing dimensionality by identifying correlated features. Problems arise when the features that differentiate clusters are not orthogonal, which may cause linear methods like PCA breakdown and reduce the data onto inappropriate dimensions (Chang, 1983). Thus, traditional dimension reduction and clustering methods are prone to pruning off too much information and, thereby, may miss subtle but significant changes in the human brain's molecular development. In contrast, sparsity-based clustering methods follow a different approach that keeps all of the features and reweights them in a dissimilarity matrix.

Approaches to Sparsity-Based Clustering

Because traditional dimension reduction methods may prune off too much information or miss more subtle changes in the human brain's molecular development, we tested a set of sparsity-based clustering algorithms. Here, sparsity refers to the idea that not all 30,000 genes play a role in brain development and irrelevant dimensions may mask clusters. Furthermore, as more and more features are included, observations become increasingly spread out until they are virtually equidistant. Sparsity-based clustering is a useful approach for analyzing those high dimensional data because the algorithms are not distance-based and can identify a smaller number of molecular features that reflect the spatiotemporal dynamics of neurodevelopment.

In this section, we introduce and compare four clustering methods designed to handle data sparsity but it is not an exhaustive review of sparsity-based clustering.

The agglomerative approach of CLIQUE (Agrawal et al., 1998) finds grids or subspaces in high dimensional data by assigning the desired number of equal length intervals (*xi*) to the grid and a global density value (*tau*) as input parameters. Notably, CLIQUE does not specify the number of clusters in the arguments, but instead compares how many points are in each rectangle of the grid with the overall density parameter and continues to partition the subspaces until the density is less than *tau*. A rectangle in the grid is considered to be dense if the proportion of points in it exceeds the *tau* parameter. CLIQUE then identifies a cluster as the maximal set of dense units in a subspace. For example, using an interval (*xi*) of 2, each dimension of the data is partitioned into two non-overlapping rectangles (units) and dense units are identified for further partitioning if they contain a greater proportion of the total number of points than the input value for *tau*. This approach does not strictly partition points into unique clusters and usually results in data points being assigned to more than one cluster. CLIQUE is also prone to classifying points as outliers and excluding them from the analysis.

The divisive clustering of PROCLUS (Aggarwal et al., 1999) is based on medoids and uses a three-step top-down approach to projected clustering. The steps involve (1) initializing the number of clusters (k) and the number of dimensions to consider in the subspace search, (2) iteratively assigning medoids to find the best clusters for the local dimensions, and (3) a final pass to refine the clusters. Typically, PROCLUS has better accuracy than CLIQUE in partitioning points into clusters, but the *a priori* selection of cluster size (k) is not easy and demands an iterative approach to finding clusters. Furthermore, by restricting the subspace search size, some essential features may be omitted from the analysis.

Both CLIQUE and PROCLUS were developed for datasets with many more samples (n), often 2–3 orders of magnitude larger than most datasets of human brain development. Although those algorithms are accurate for large datasets with thousands of samples, they are less well suited for discovering clusters in small sample sizes. So we needed to test sparsity-based clustering designed for small datasets, and this criteria led us to select two more approaches to sparse hierarchical clustering, SPARCL and robust and sparse k-means clustering (RSKC) (Kondo, 2016; Witten and Tibshirani, 2018).

SPARCL was developed by Witten and Tibshirani (2010) to adaptively select and reweigh the subset of features during clustering thus eliminating the need for data reduction preprocessing. The algorithm uses a lasso-type penalty to address the challenge of clustering samples that differ on a small number of features. The reweighted variables then become the input to k-means hierarchical clustering. The adaptive feature selection of SPARCL focuses on the subset of genes or proteins that underlie differences among clusters, so this process is similar to removing noise from the data. Thus, SPARCL simultaneously clusters the samples and identifies the

dominant features thereby making it easier to determine the subset of proteins or genes responsible for partitioning samples into different clusters.

SPARCL has many strengths for analyzing datasets with $p \approx n$ or $p > n$; however, it can form clusters containing just one observation (Witten and Tibshirani, 2010). A more recent extension of the algorithm, RSKC, addresses small clusters by assuming that outlier observations cause this problem. RSKC uses the same clustering framework as SPARCL, except that it is "*robust*" to outliers (Kondo et al., 2016). RSKC iteratively identifies clusters in the data, then identifies clusters with a small number of data points (e.g., $n = 1$) and flags those data points as potential outliers. The outliers are temporarily removed from the analysis, and clustering proceeds as outlined above for SPARCL. Once all clusters have been identified, the outliers are re-inserted in the high-dimensional space and grouped with the nearest neighbor cluster. Thus, RSKC identifies clusters in the data and includes all of the data points.

MATERIALS AND METHODS

Datasets

Our lab has been studying development of human visual cortex (V1) by quantifying expression of synaptic and other neural proteins using a library of postmortem tissue samples ($n = 31$, age range 21 days – 79 years, male/female = 18/13) (**Supplementary Table 1**). In addition, genome-wide exon-level transcriptomic data that was collected by Kang et al. (2011) was used and the postnatal V1 data were extracted ($n = 48$, age range 4 months – 82 years, male/female = 27/21) (**Supplementary Table 2**). The transcriptomic data were used to test the reproducibility and scalability of the sparsity-based clustering. The preprocessed exon array data from Kang et al. (2011) were downloaded from the Gene Expression Omnibus (GSE25219). The exon-summarized expression data for 17,656 probes were extracted, and probe identifiers were matched to genes. If a gene was matched by two or more probes and the probes were highly correlated as determined by Kang et al. (2011) (Pearson correlation, $r \geq 0.9$), then the expression values were averaged for a total of 17,237 genes.

The clustering methods were tested using three groups of protein or gene data. The first group of protein data was from a series of studies using the Murphy lab postmortem samples to examine the development of molecular mechanisms that regulate experience-dependent plasticity in human V1 (Murphy et al., 2005; Pinto et al., 2010, 2015; Williams et al., 2010; Siu et al., 2015, 2017). Western blotting was run using each sample (2–5 times) to probe for 23 different proteins (**Supplementary Table 3**). The tissue preparation and Western blotting methods have been described in detail previously (Siu et al., 2017, 2018). The initial clustering tests used a subset of seven proteins (GluN1, GluN2A, GluN2B, GluA2, $GABA_A\alpha1$, $GABA_A\alpha3$, and Synapsin) to explore age-related clustering with AGNES, PROCLUS, CLIQUE, SPARCL, and RSKC.

Next, the sparsity-based clustering using RSKC was explored using all 23 proteins to determine how adding more features changed the age-based clustering. Then the reliability of the age-related clustering was explored by running 100 iterations of RSKC with the 23 proteins. A heatmap illustrating the number of times each sample was partitioned into a cluster was made to visualize the reliability.

The scalability of RSKC was tested using a larger protein database and the much larger gene database. These tests included clustering a matrix with 95 proteins collected from the Murphy lab postmortem tissue samples. This time the samples were probed with a high density ELISA array (RayBiotech Quantibody Human Cytokine Array 4000) and an additional 72 proteins were measured for a total of 95 proteins ($p = 95$) (**Supplementary Table 4**). Finally, RSKC clustering was done using the genes in the Kang database by selecting those listed in the SynGO ontology ($n = 988$) (Koopmans et al., 2019) and also the full set of genes ($n = 17{,}237$).

The Basic Steps to Sparsity-Based Clustering Using R

Here we describe four sparsity-based high-dimensional clustering approaches (PROCLUS, CLIQUE, SPARCL, and RSKC) for analyzing the development of human V1 using 7 or 23 proteins. Then we explore the scalability of the RSKC method using two larger datasets with 95, 988, or 17,237 proteins or genes.

All of the analyses were done in the R programming language using the integrated development environment RStudio (version 1.3.1093). The basic steps in the workflow used to examine each of the clustering methods are illustrated in **Figure 1**. The text refers to the R packages that were used and R Markdowns with code and figures are included in **Supplementary Material**.

Figure 1 illustrates the steps that were used for testing various sparsity-based clustering methods to examine if they produce an age-related progression in the median age of clusters. The data were prepared in an *nxp* matrix with each sample forming a row and the features, either proteins or genes, arranged in columns. Those data were used as the input to the clustering algorithms. Here the sparsity-based algorithms tested were PROCLUS and CLIQUE from the *subspace* package (Hassani, 2015), SPARCL (Witten and Tibshirani, 2018), and RSKC (Kondo, 2016). For the algorithms a range of k or xi values from 2 to 9 were tested to explore the types of clusters produced.

The results of a tSNE dimension reduction was used to visualize clusters for all of the methods tested. However, clustering was not done on the tSNE data itself even though that is a commonly used approach. We used tSNE strictly as a visualization tool because it does a good job of projecting points from high dimensional space onto 2D so that neighboring points reflect their similarity.

The Elbow method was used to determine the number of clusters. Finally, the quality of the age-related clustering of the samples was evaluated by making a boxplot to visualize the progression of the median ages.

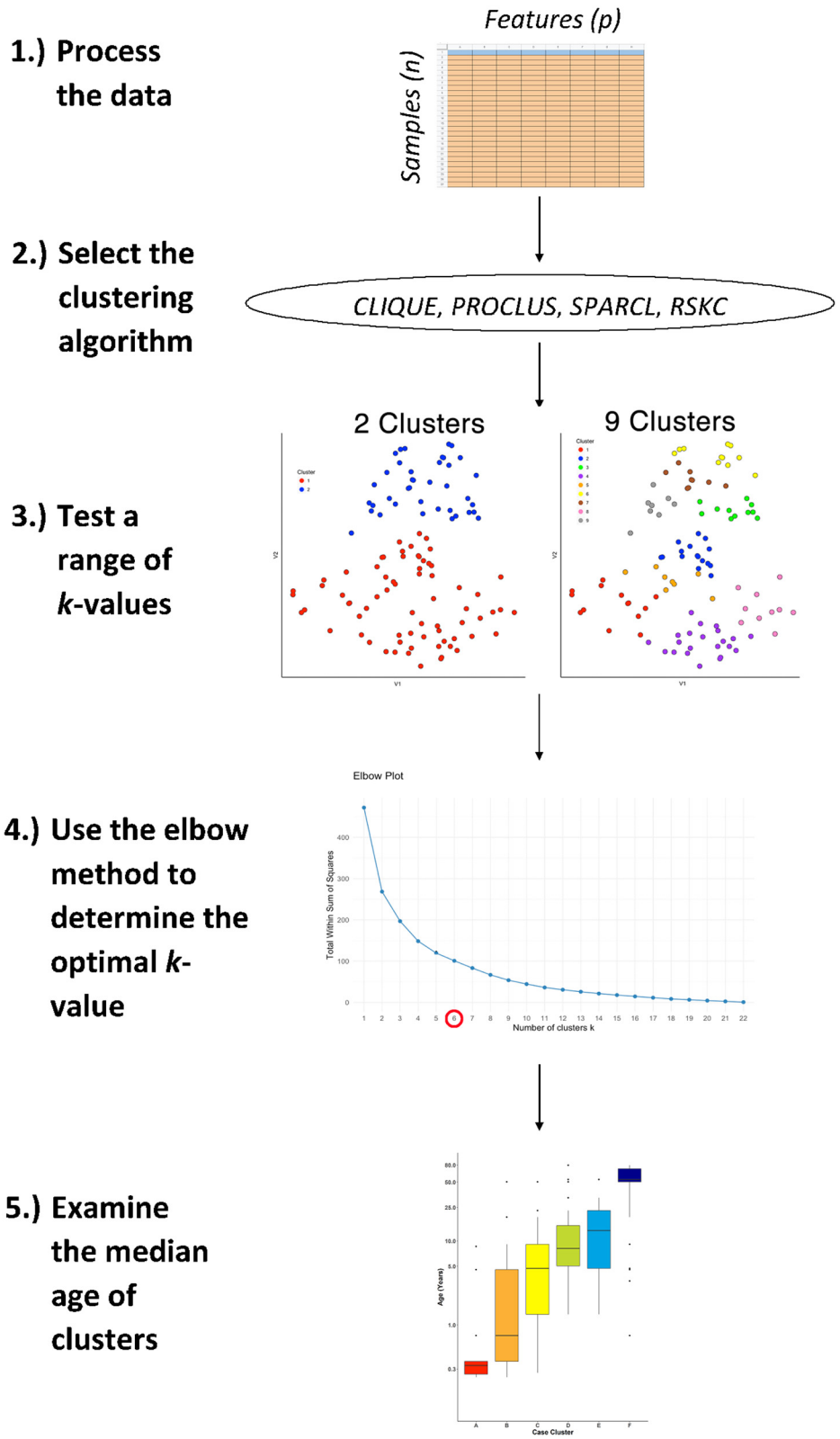

FIGURE 1 | The workflow for studying age-related molecular development of the brain. First, arrange the data into an $n \times p$ matrix, where features (p) are represented as columns and samples (n) as rows. Then, select the desired sparse clustering algorithm (e.g., CLIQUE, PROCLUS, SPARCL, and RSKC) and test its performance along a range of clusters (k). Lastly, determine the optimal k value using the elbow method and compare the median age of clusters with boxplots.

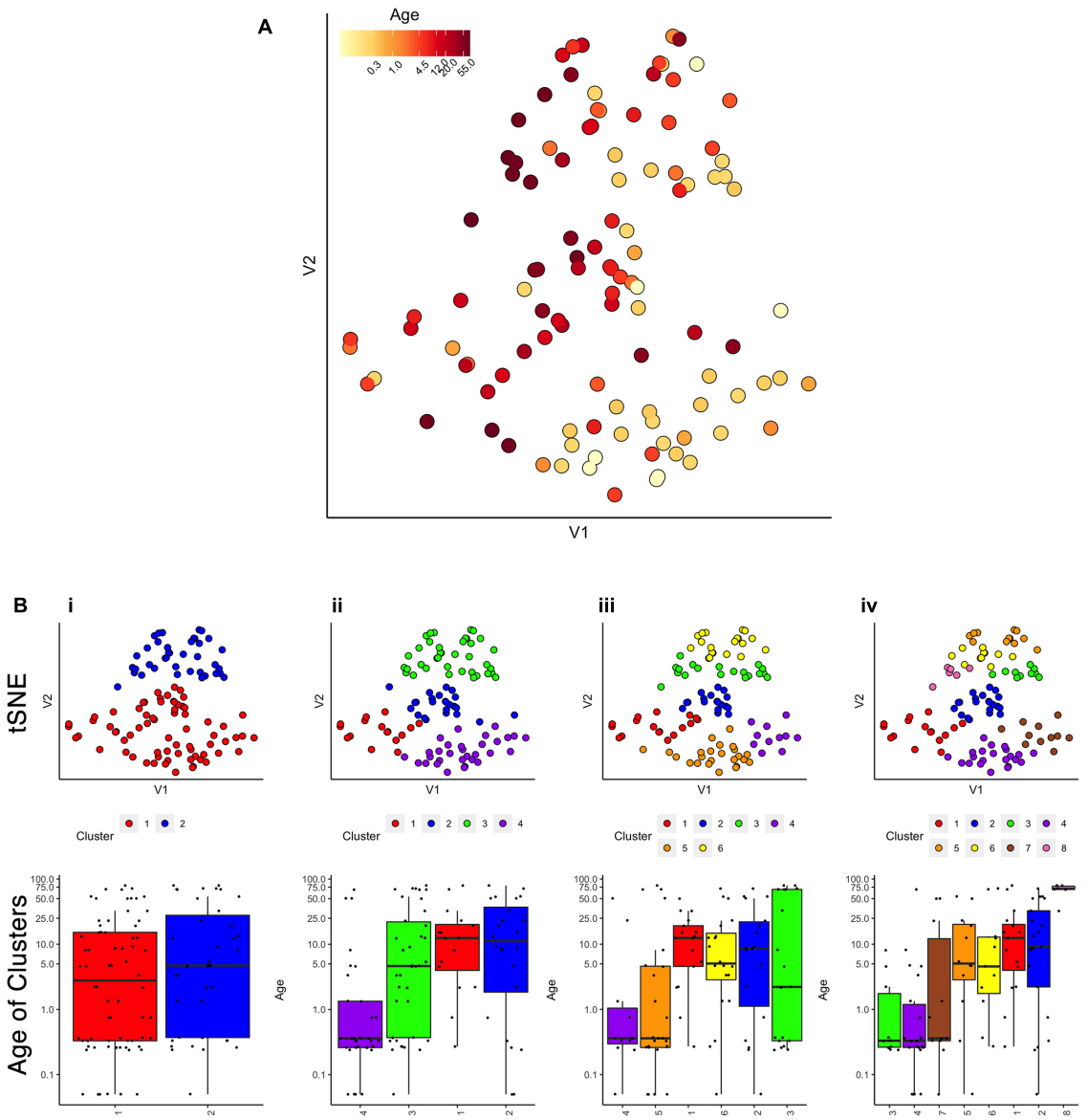

FIGURE 2 | Age-related organization of cases and initial clustering results. **(A)** A 2D tSNE scatter plot color-coded according to individual cases' age. **(B)** tSNE 2D scatter plots and box plots showing the results of *agnes* for k = 2, 4, 6, and 8 case clusters. The tSNE plots display individual samples as points. Points are color-coded according to their designated cluster determined by *agnes*. Boxplots denote the median and interquartile range of ages in each cluster, and points denote outliers.

This workflow was used for all of the clustering methods described in the next section and an example R Markdown of the analysis is included in **Supplementary Material**.

RESULTS

Evaluating Sparsity-Based Clustering for Finding Age-Related Clusters

First, we evaluated the data by exploring if simply visualizing the samples using tSNE produced an age-related organization. The human V1 samples with seven proteins and all of the WB runs were used as the input to the *tsne* package (Donaldson and Donaldson, 2010; **Figure 2A**). Color-coding the samples by their age showed a global progression in the ages with younger samples mapped to the bottom right and older to the top left in the 2D tSNE space. Next, we applied a commonly used agglomerative hierarchical clustering algorithm, AGNES in the *cluster* package (Maechler, 2019), to test if this clustering approach would reveal age-related groupings of the samples. This algorithm uses the dissimilarity matrix to merge nodes in the tree and it partitioned these data into clusters that suggest an age-related progression (**Figure 2B**). However, groups of 2 or 3 adjacent clusters had very similar median ages indicating poor age-related

separation of the samples. A major weakness of this hierarchical clustering approach is that incorrect branching can never be undone. Nevertheless, these findings show that even distance-based hierarchical clustering of human V1 postnatal samples can find some age-related progression of postnatal samples.

Next, we tested the two density projection sparsity-based clustering methods that use either top-down (PROCLUS) or bottom-up (CLIQUE) clustering with all of the observations (n = 31) and seven of the proteins from the human visual cortex development dataset. The outputs were visualized in 2D using tSNE, and the data points were color-coded according to the clusters identified by each method. Finally, to determine if the clusters represented developmental changes in the dataset, we plotted boxplots showing the median age of the samples in the cluster.

PROCLUS

The PROCLUS clustering method was implemented in RStudio using the *ProClus* function in the *subspace* package version 1.0.4 (Hassani, 2015). We explored clusters between k = 2–9, and **Figure 3** shows the results for 2, 4, 6, and 8 clusters for the human V1 data with seven proteins and all runs included.

Visualizing the clusters found with PROCLUS (**Figure 3A**) showed a mixing of the samples, but the boxplots illustrating the ages of the samples in the clusters suggested an age-related progression, especially for 4 or 6 clusters (**Figure 3B**). The PROCLUS clusters' age progression was somewhat better than the hierarchical clusters but still had clusters with very similar median ages. More importantly, some clusters had only one or two data points, and many samples were tagged as outliers (small gray dots) and excluded from the clusters. Thus, PROCLUS's iterative top-down feature identification and cluster border adjustments performed poorly for identifying age-related clusters of human V1 development.

CLIQUE

The bottom-up clustering method CLIQUE was tested to determine how well this iterative approach to building clusters performed using seven proteins to group the human V1 samples into age-related clusters.

The CLIQUE function from the *subspace* package (Hassani, 2015) was used to test clustering. CLIQUE requires an input value for the interval setting because the intervals divide each dimension into equal-width bins that are searched for dense regions of data points. Here we tested a range of input interval values (xi = 2–8) and those resulted in 4–9 clusters (**Figures 3C,D**).

CLIQUE allows data points to be in more than one cluster, so to visualize the multi-cluster identities, we plotted the data points using concentric color-coded rings. CLIQUE placed all of the data into multiple overlapping clusters, which was true for all interval settings (xi = 2–8). The poor partitioning of samples resulted in no progression in the clusters' median age (**Figure 3D**). Thus the iterative bottom-up clustering of CLIQUE performed poorly for clustering the samples into age-related groups.

Comparing these top-down PROCLUS and bottom-up CLIQUE density methods for sparsity clustering showed that neither algorithm was a good fit for producing age-related clustering of the samples. PROCLUS performed somewhat better because some of the parameters resulted in clusters with a progression in the median cluster age; but, the number of data points treated as outliers was unacceptably high.

SPARCL

Next, we tested a sparsity-based clustering algorithm, sparse k-means clustering, optimized for small sample sizes (Witten and Tibshirani, 2010). The SPARCL package (version 1.0.4) (Witten and Tibshirani, 2018) was used to cluster the human V1 samples with data from 7 proteins. This approach adaptively finds subsets of variables that capture the different dimensions and includes all samples in the clusters. SPARCL searches across multiple dimensions in the data and adjusts each variable's weight based on the contribution to the clustering. Thus, the term "sparse" in this method refers to selecting different subsets of proteins to define each cluster.

To implement sparse k-means clustering, we used the *Kmeans.sparsecluster* function in the SPARCL package (Witten and Tibshirani, 2018). We explored a range of k clusters between k = 2–9. The SPARCL package also includes a function to help determine other input variables, such as the boundaries for reweighting the variables (*wbounds*) to produce optimal clustering.

Visualizing the clusters created by SPARCL showed useful partitioning of the samples into clusters (**Figure 3E**) that moved from the bottom right to the top left in the tSNE plot. Also, the boxplots illustrate a good progression of the median cluster age for 4 and 6 clusters. However, SPARCL is prone to making clusters with only 1 sample, and that was the case in this example for k = 4–9 clusters. To address that problem, we tested another sparse k-means cluster algorithm that is robust to making clusters of n = 1.

Robust Sparse *k*-Means Clustering

Finally, we tested a modified version of the SPARCL algorithm called RSKC (Kondo et al., 2016). The RSKC algorithm was designed to be robust to the influence of outliers that can drive other algorithms to create clusters of n = 1. RSKC operates by iteratively omitting outliers from cluster analysis, assigning all remaining samples to clusters, and then reinserting outliers to the analysis by grouping them into the nearest-neighboring cluster.

Using the *RSKC* package in R (Kondo, 2016) we explored clustering for a range of k values (k = 2–9) using the human V1 dataset with 7 proteins and all runs (**Figure 4**). The visualization of the clusters on the tSNE plot showed good grouping of the samples into spatially separated clusters. The boxplots illustrate good progression in the median ages of the clusters, especially for 4 or 6 clusters (**Figure 4B**). In addition, the algorithm adaptively reweighted the proteins to identify the most robust clusters and we plotted the weights for each of the 7 proteins (**Figure 4C**). This component of RSKC identified the lifespan variations in GluN2B, Synapsin, and GluN2A as having the greatest impact on the clustering of the samples.

Next, the scalability of RSKC was explored using the full dataset of 23 proteins measured for the human V1 samples

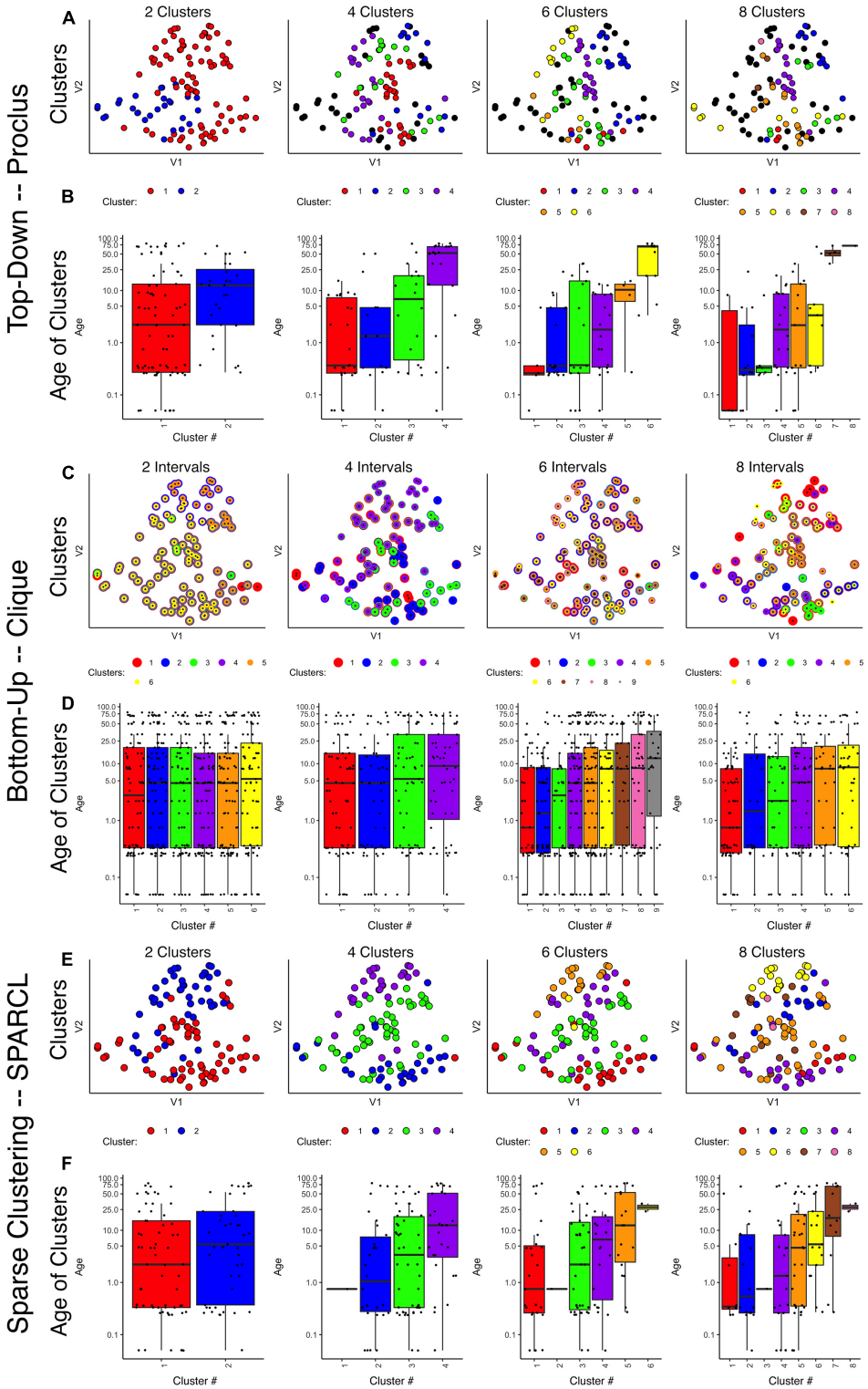

FIGURE 3 | Comparison of various sparse clustering methods. Top-down PROCLUS subspace method across range of cluster numbers (2, 4, 6, and 8). The clusters are visualized in tSNE 2D scatter plots of the data by color-coding each data point with its cluster identity **(A)** and in boxplots showing the median age of the samples in each cluster **(B)**. **(C,D)** Bottom-up CLIQUE subspace clustering method for a range of "intervals." Different clusters are visualized as colored dots in a tSNE representation of the data **(C)** and as box plots depicting the mean age of the samples **(D)**. **(E,F)** Sparse clustering after varying the inputted *k* cluster number (2, 4, 6, and 8). Different clusters are visualized as colored dots in a tSNE representation of the data **(E)** and as box plots depicting the mean age of the samples **(F)**. The colors in scatter plots and boxplots represent the cluster designation for all plots.

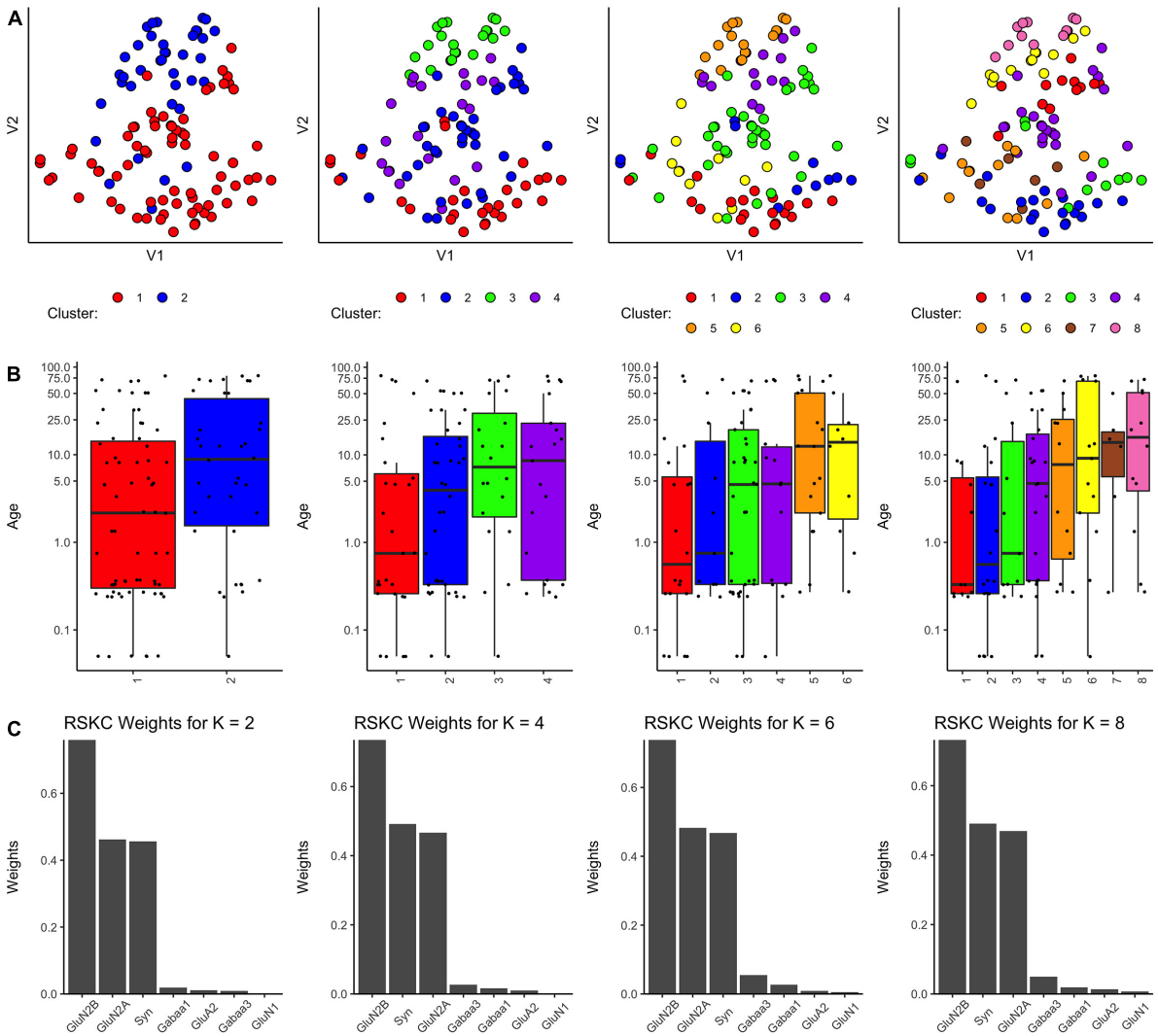

FIGURE 4 | Age-related clustering of seven synaptic proteins for a single iteration. Expression data from seven synaptic proteins were input into RSKC and used to identify k = 2, 4, 6, and 8 case clusters. For each k value, three plots were constructed: **(A)** 2D tSNE scatter plots showing samples color-coded by their cluster designations, **(B)** box plots displaying the distribution of ages for each cluster, and **(C)** a bar graph representing the RSKC weights for all seven proteins.

(Murphy lab) (**Figure 5**). In this example, the average expression value for each protein was used and the elbow plot method identified six clusters. **Figure 5** shows the results of three separate runs of RSKC on the 23 protein dataset. All three runs resulted in similar clustering (**Figures 5A–C**) with a tight progression of age-related clustering from cluster A with the youngest median age to cluster F with the oldest age. The addition of more proteins to the RSKC clustering provided greater precision for identifying the subtle changes that represent the temporal dynamics of human V1 development. The weights for the 23 proteins (**Figures 5D–F**) showed that all of the proteins contributed to this high dimensional clustering. Comparing the feature weights among the three runs showed some reordering in the weight of individual proteins suggesting that care is needed when using weights from a single run. These weights were used to improve the visualization of the clusters in a tSNE plot. The protein expression values for each sample were transformed by multiplying with the corresponding weight and those transformed data were visualized using tSNE (**Figures 5G–I**). Those plots showed the separation of the clusters in the 2D tSNE space.

Since the starting conditions for clustering can affect which samples end up in a cluster, we tested how robust RSKC clusters were by running the algorithm 100 times with different starting conditions. We then plotted the results of 100 iterations in a boxplot showing the age-related clusters and a heatmap showing the number of times each sample fell into the different clusters (**Figures 6A,B**). This analysis showed that the progression in the age of the clusters was robust to the starting conditions (**Supplementary Table 5**). Furthermore, the heatmap showed that clusters B and C were the least stable, but the other clusters had strong consistency for which samples were partitioned

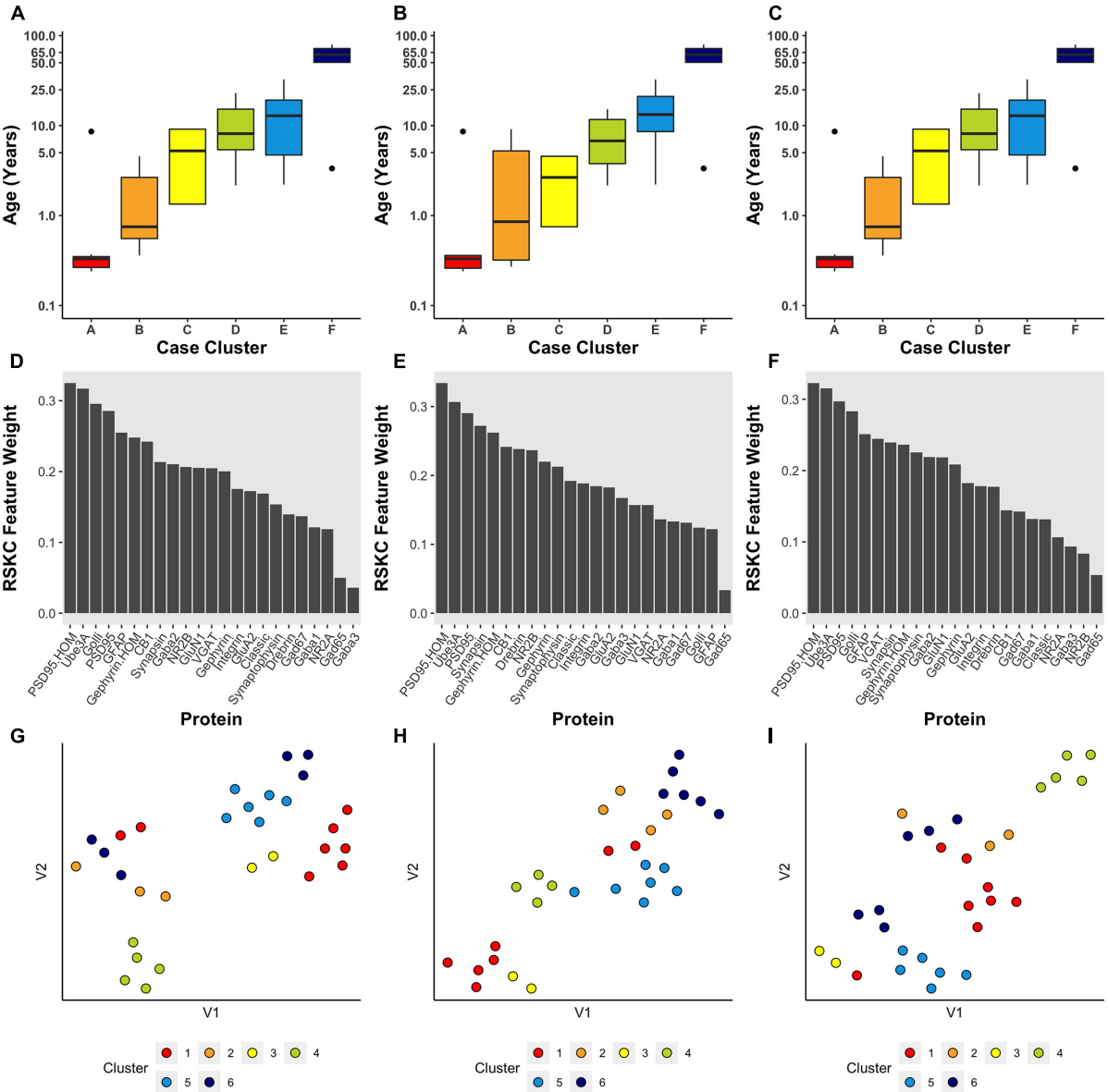

FIGURE 5 | Age-related clustering of 23 synaptic proteins for three single iterations. **(A–C)** Expression data from 23 synaptic proteins was used to identify six case clusters. Boxplots of cluster age ordered from youngest (red) to oldest (dark blue) median age. **(D–F)** Bar plot visualizing RSKC feature weights for proteins. **(G–I)** tSNE plot of the protein data scaled by RSKC weights and color-coded by RSKC cluster. In both **(A–C)** and **(G–I)**, sample ages were reduced to sample averages to reduce crowding.

into those clusters. The Jaccard similarity was calculated for all cluster pairs to determine the proportion of samples shared between the clusters. Cases were counted as shared when the cases were partitioned to the cluster 10 or more times because the metric is sensitive to small samples sizes. The similarity indices ranged from 0 to 22% (adjacent pairs: A–B 12%, B–C 22%, C–D 11%, and D–E 20%), with cluster C having the most cases shared with other clusters. In addition, the average feature weight for each of the 23 proteins was calculated from the 100 runs (**Figure 6C**) and illustrated the gradual progression of feature weights.

Testing Robust Sparse *k*-Means Clustering With Larger Numbers of Proteins or Genes

So far, we have shown that RSKC does a good job of partitioning samples into age-related clusters with datasets that have fewer than 25 proteins. Here we examine if RSKC scales to larger datasets with 2–3 orders of magnitude more features.

We ran the RSKC clustering using data collected from the Murphy lab human V1 samples with measurements for 95 proteins (**Supplementary Table 4**; **Figure 7A**). Once again,

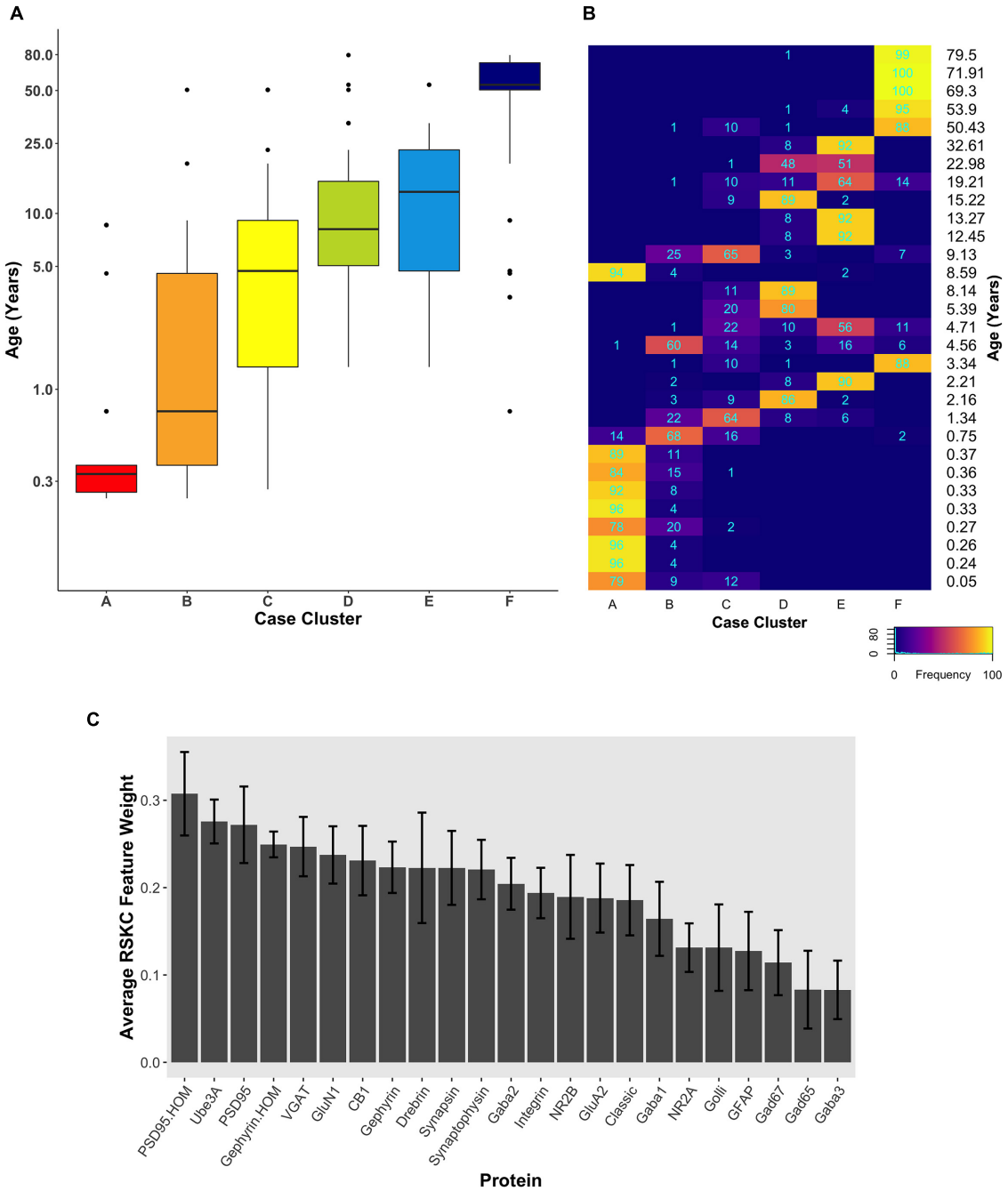

FIGURE 6 | Robust, age-related clustering of 23 synaptic proteins for 100 iterations. **(A)** Expression data from 23 synaptic proteins was used to identify six case clusters. The cluster designation of each case over 100 iterations of RSKC was used to visualize the distribution of case ages (in years). Boxplots denote the median and interquartile range of ages in each cluster, and points denote outliers. **(B)** Heatmap visualizing the number of times each case was assigned to each cluster over 100 iterations of RSKC. **(C)** The average RSKC feature weight for each of the proteins from the 100 iterations.

100 iterations of RSKC clustering was used to ensure that the clusters were robust to the starting condition. This analysis found strong age-related clustering of the samples showing six well-defined clusters that stepped across the lifespan. We tested if the progression of cluster ages could arise by chance by rerunning the clustering but on each iteration the age of the sample was randomized. As expected, randomizing the ages resulted in clusters with a very broad range of ages and no progression in the mean cluster age (**Supplementary Figure 1**).

Next, RSKC clustering was extended to the transcriptomic dataset from Kang et al. (2011; **Supplementary Table 2**). First, RSKC was run using the 88 genes that matched the proteins in

Figure 7A. Even though the two datasets used different samples it was possible to compare the ages of the clusters because the range of ages and number of samples were similar. The progression of age-related clusters for the gene data (**Figure 7B**) was similar to the protein clusters and there was a strong correlation ($r = 0.81$) between the median ages of the six cluster pairs.

The strong correlation between the protein- and gene-cluster ages was particularly interesting because previous studies have shown that the correlation between large sets of protein and gene expression values is notoriously low (e.g., $r \sim 0.2$) (Gry et al., 2009). To assess if the datasets used here simply had an unusually strong similarity between the lifespan changes in the expression values for each protein and gene pair we calculated those correlations. To facilitate this analysis the protein and gene expression values were normalized by calculating *z*-scores and the normalized values were partitioned into six age-bins (<1, 1–5, 5–12, 12–20, 20–55, and >55 years) (**Supplementary Figure 2**). The correlation coefficient was then calculated for the 88 protein-gene pairs using the mean gene and protein expression values from the six age bins (**Supplementary Figure 3A**). The mean correlation between the 88 protein-gene pairs was $r = 0.15$ and the median correlation was only slightly higher (median $r = 0.21$, 95% CI 0.04–0.25) (**Supplementary Figure 3B**). Thus, it is unlikely that the strong correlation found between the ages of the protein- and gene-clusters arose from a simple linear relationship between those two types of molecular measurements. Instead, the common cluster ages for these different omics datasets suggest similar high dimensional patterns that RSKC uses to partition the samples into the series of age-related clusters.

Finally, we examined how well RSKC performed on datasets with measurements of hundreds to thousands of genes using 988 genes that overlap with the SynGO database of synaptic genes (Koopmans et al., 2019) and then with all 17,237 genes in the Kang dataset. The SynGO genes were analyzed to assess if a large set of functionally genes might reveal a different pattern of clusters from the full set of genes. The analysis of synaptic genes showed an age-related progression of the median age of the clusters (**Figure 7C**). Compared with the protein clusters (**Figure 7A**), the median age of the SynGO clusters jumped between clusters B and C (**Figure 7C**) and a very similar pattern of age-related clusters was found when all 17,237 genes in the Kang dataset were used (**Figure 7D**). Thus, RSKC cluster analysis of 95 proteins revealed the tightest age-related clusters, but the gene data also resulted in the partitioning of samples into age-related clusters. This finding contrasts with hierarchical clustering used by Kang et al. (2011), (Supplementary Figure 8) that did not partition postnatal samples into age-related clusters. Thus, the optimization of sparse *k*-means cluster analysis (RSKC) for small sample sizes provides another approach for analyzing the human brain's molecular development that is sensitive to the subtle molecular changes that occur across the postnatal lifespan.

A Note About Selecting the Number of Clusters

An essential step in *k*-means clustering is selecting *k*, which denotes the number of groups to classify observation into. The correct choice of *k* is often ambiguous, as there are many different approaches for making this decision. Intuitively, an optimal *k* lies in between maximum generalization of the data using a single cluster and maximum accuracy by assigning each observation to its own cluster. One of the most common heuristics for determining *k* is the elbow plot method, where the sum of squared distances of observations to the nearest cluster center is plotted for various values of *k*. As *k* increases, the sum of squared distances tends toward zero. The "elbow" occurs at the point of diminishing returns for minimizing the sum of squared distances, and the *k* value at this point is selected as the optimal number of clusters.

To tailor the selection of *k* to RSKC, we applied the elbow method to the Weighted Within Sum of Squares (WWSS), the objective function maximized by the algorithm. WWSS was calculated for various values of *k* and averaged over 100 iterations. The elbow can be identified using the elbowPoint function in the *akmedoids* package (version 0.1.5) (Adepeju et al., 2020), which uses a Savitzky–Golay filter to smooth the curve and identify the *x* where the curvature is maximized. This method found that $k = 6$ was the optimal number of clusters for all of the applications of RSKC used in this paper.

There are more than 30 methods to determine the optimal values for *k* and a large number of journal papers (e.g., Tibshirani et al., 2001) and web resources (e.g., Cluster Validation Essentials) that can be used to learn more. The R packages *NbClust* (Charrad et al., 2015) and optCluster (Sekula, 2020) are particularly helpful tools for choosing the number of clusters because they test various methods for selecting *k* (Charrad et al., 2014).

APPLICATION OF ROBUST SPARSE *k*-MEANS CLUSTERING CLUSTERS TO STUDY HUMAN VISUAL CORTEX DEVELOPMENT

Previous studies using the datasets analyzed here (Murphy et al., 2005; Pinto et al., 2010, 2015; Williams et al., 2010; Kang et al., 2011; Siu et al., 2015, 2017) have examined molecular development by assigning samples into age-bins that approximate the lifespan stages defined by developmentalists. In contrast, the previous section describes a data-driven approach to partitioning samples into age-related clusters using sparse *k*-means clustering (RSKC). This use of unsupervised clustering raises the possibility that it might reveal aspects of human visual cortex molecular development that have escaped previous analyses. This section explores some of the information about human visual cortex development that can be revealed by examining the content of age-related clusters.

First, we compare partitioning of the samples into pre-defined age-bins versus data-driven clustering of the 23 proteins for post-mortem intervals (PMIs), the proportion of cases, and the biological sex of the cases (**Supplementary Figure 4**). The distribution of PMIs was similar between the two methods of partitioning the lifespan as was the

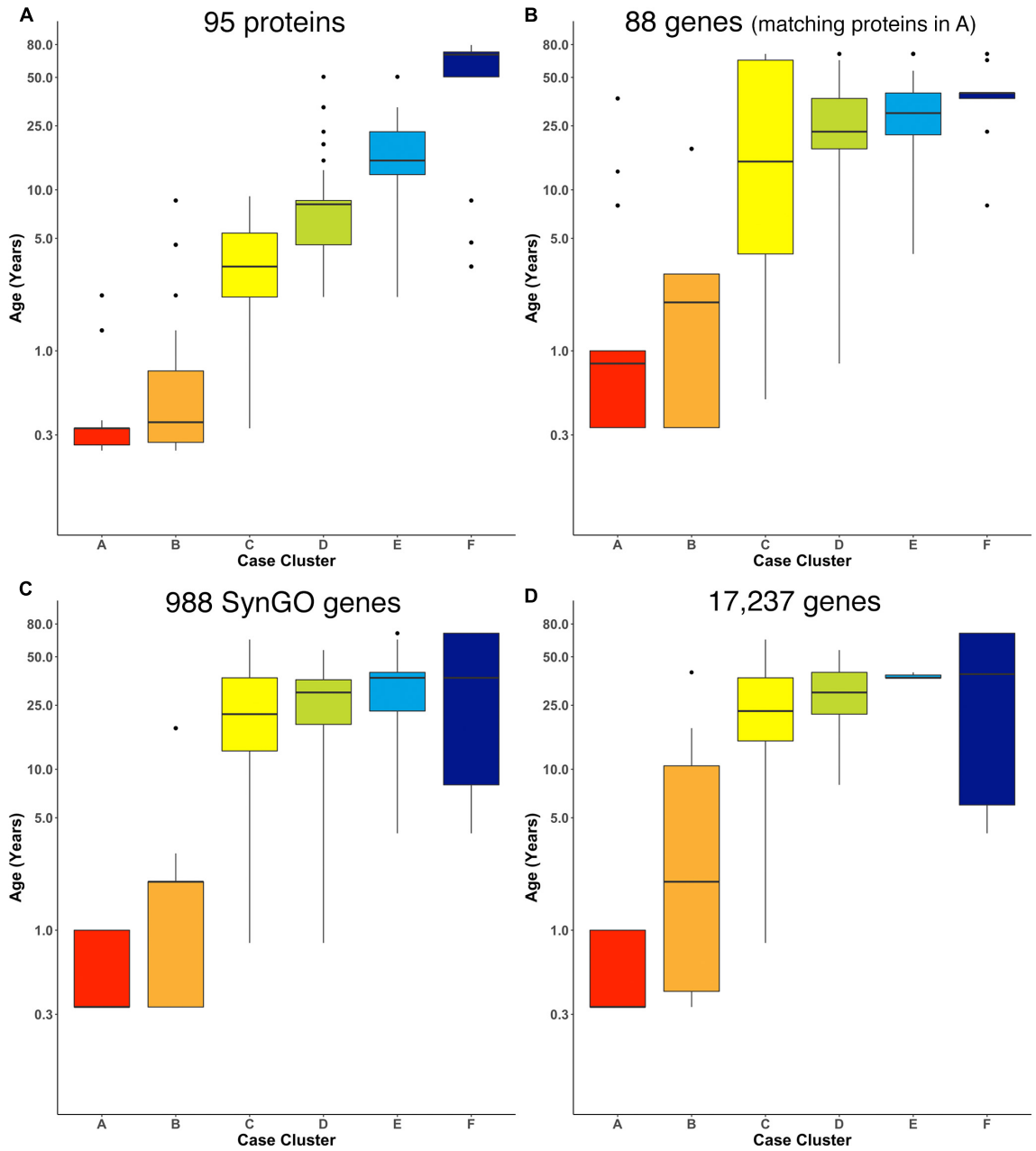

FIGURE 7 | Age-related clusters for large numbers of proteins and genes. Expression data from **(A)** 95 synaptic and immune-related proteins, **(B)** 88 genes that correspond with the protein in **(A)**, **(C)** 988 synaptic genes that correspond with the SynGO gene list, and **(D)** 17,237 protein-coding genes was used to identify six age-related clusters. The cluster designation of each sample over 100 iterations of RSKC was used to visualize the distribution of sample ages. Boxplots denote the median and interquartile range of ages in each cluster, and points denote outliers.

proportion of samples and the balance of females and males in the bins. The progression of cluster ages was apparent when the age bins were color-coded to reflect the cluster identity (**Supplementary Figure 4G**). That histogram illustrated an interesting aspect of cortical development during young childhood (1–4 years) where samples in that age-bin were partitioned into five different clusters. Similar to previous studies that observed heightened childhood heterogeneity with waves of inter-individual variability that peak between 1 and 3 years (Pinto et al., 2015; Siu et al., 2017). The findings here suggest that the relationship between chronological and brain age varies across the lifespan.

The developmental trajectories of the 23 proteins were plotted using LOESS fits (95% CI) to the expression values (normalized to control), and each sample was color-coded by their cluster assignment. The LOESS curves were ordered based

on similar trajectories to illustrate the range of developmental patterns with some increasing (e.g., $GABA_A\alpha1$) or decreasing (e.g., $GABA_A\alpha2$) monotonically across the lifespan while others followed an inverted-U (e.g., gephyrin), an undulating pattern (e.g., VGAT), or remained relatively unchanged (e.g., $GABA_A\alpha3$) (**Figure 8A**). The range of trajectories highlights the need for high-dimensional analyses to capture the complexity of this development. To help describe when the expression level of a protein in a cluster was above or below the overall mean, we implemented the over-representation analysis (ORA_phenotype function) described previously (Balsor et al., 2020; **Figure 8B**). Briefly, for each protein, a normal distribution was simulated using the mean and standard deviation of the expression values for all samples. Then the boxplots were color-coded by comparing the expression values for each cluster with the simulated distribution. Here, the box for a cluster was coded as over-represented (red) if the 25th percentile was above 95% of the simulated distribution and under-represented if the 75th percentile was below 5% of the simulated distribution. Of course, other cutoff values for the ORA can be implemented to be more stringent or lenient for the color-coding (e.g., **Supplementary Figure 5**), or other methods such as estimation statistics (Bernard, 2019) can be used for this step depending on the nature of the question.

Here, the ORA identified a range of over- or under-represented proteins in each cluster from a high of 12 proteins in clusters C to 5 proteins in cluster F (A – 6 proteins, B – 11 proteins, C – 12 proteins, D – 8 proteins, E – 7 protein, and F – 5 proteins). These LOESS curves and boxplots for the expression of each protein help to describe development, but it is challenging to synthesize an overall pattern for human V1 development when confronted with making hundreds of pairwise comparisons. To address that problem we implemented a series of visualizations and analyses aimed at representing the high-dimensional nature of these data.

The first step in addressing the high-dimensional patterns of protein expression captured by the age-related clusters was to plot a bubble chart illustrating the expression levels of all 23 proteins for the 6 clusters. That visualization ordered the proteins by their RSKC weight and color-coded each bubble with the normalized mean protein expression with blue representing low and red high expression levels (**Figure 9**). The visualization helped identify that cluster D has high expression levels for many proteins. That cluster represents older children and the transition to adolescence (mean cluster age = 10.3 years, CI 9.6–11.1 years) when rapid changes in cortical microstructure have been found (Norbom et al., 2021). In addition, groups of proteins with either high or low expression can be identified in a cluster, such as the higher expression of Golli-MBP, GFAP, CB1, and NR2B in cluster B. Thus, this visualization shows the mean expression for the 23 proteins in the 6 clusters, but it is still challenging to derive what differentiates the clusters. To address this, we applied our recently developed workflow (Balsor et al., 2019, 2020) that includes dimension reduction, identification of features and the construct of a plasticity phenotype visualization to characterize the development of the human visual cortex. This workflow is described in detail in a previous publication (Balsor et al., 2020).

Adding Principal Component Analysis for Dimension Reduction

This part of the workflow aims to reduce the dimensionality of the data by identifying combinations of functionally related proteins that we call features and using those features to capture the high dimensional pattern of brain development. The first step involves using PCA, a standard approach for reducing dimensionality when studying brain development (Jones et al., 2007; Beston et al., 2010; Bray, 2017). The scree plot showed that the first three dimensions capture ∼60% of the variance in the data (**Supplementary Figure 6**), and the correlation matrix identified the strength of the relationship between each protein and the 23 dimensions. For example, the expression of Gephyrin and PSD95 was strongly correlated with Dim 1 while VGAT, $GABA_A\alpha2$, CB1, and GluN1 were strongly correlated with Dim 2. In addition, the quality of the representation for each protein on the first three dimensions was assessed using the $\cos^2$ metric. The $\cos^2$ (cosine square, coordinates square) conveys the quality of the representation of that variable using the projection angle onto each PC dimension. The closer that $\cos^2$ is to 1, the better the quality of that variable's projection onto the dimension. The biplots illustrate the quality of the representation of each protein on Dim 1, 2, and 3 (**Figures 10A,B**) and show that some aspects of the RSKC-defined clusters are apparent when the samples are plotted in the PC space (**Figures 10C,D**). However, clustering of the samples in PC space was less distinct than illustrated in **Figure 5** where tSNE plots were used to visualize clusters in the RSKC-weight transformed data.

We examined which proteins were well represented by the first three dimensions by plotting the $\cos^2$ values for Dim 1, 2, and 3 (**Figures 10E–G**) and the sum of the $\cos^2$ for those dimensions (**Figure 10H**). The matrix of $\cos^2$ values illustrated that only two of the proteins ($GABA_A\alpha3$ and Drebrin) were weakly represented by the first three dimensions (**Figure 10I**). The remaining steps focus on Dim 1, 2, and 3 because they captured a large amount of the variance and had high-quality representations for most proteins.

Comparing Principal Component Analysis and Robust Sparse *k*-Means Clustering

We compared RSKC and PCA by assessing the similarity of RSKC weights and PCA $\cos^2$ values for each protein (sum of Dim 1, 2, and 3) for the 23 proteins (**Figure 11A**). There was a strong correlation (ρ = 0.72) between the two approaches; however, some proteins fell away from the line of best fit. Next, the differences between RSKC weights and PCA $\cos^2$ of the proteins were assessed using a Bland–Altman analysis (Giavarina, 2015). This used the calculated differences between the normalized measures and plotted those as the difference score for each protein. Also, interval bands were plotted to represent no difference between the RSKC and PCA measurements (blue band), when RSKC measurements were greater (positive red

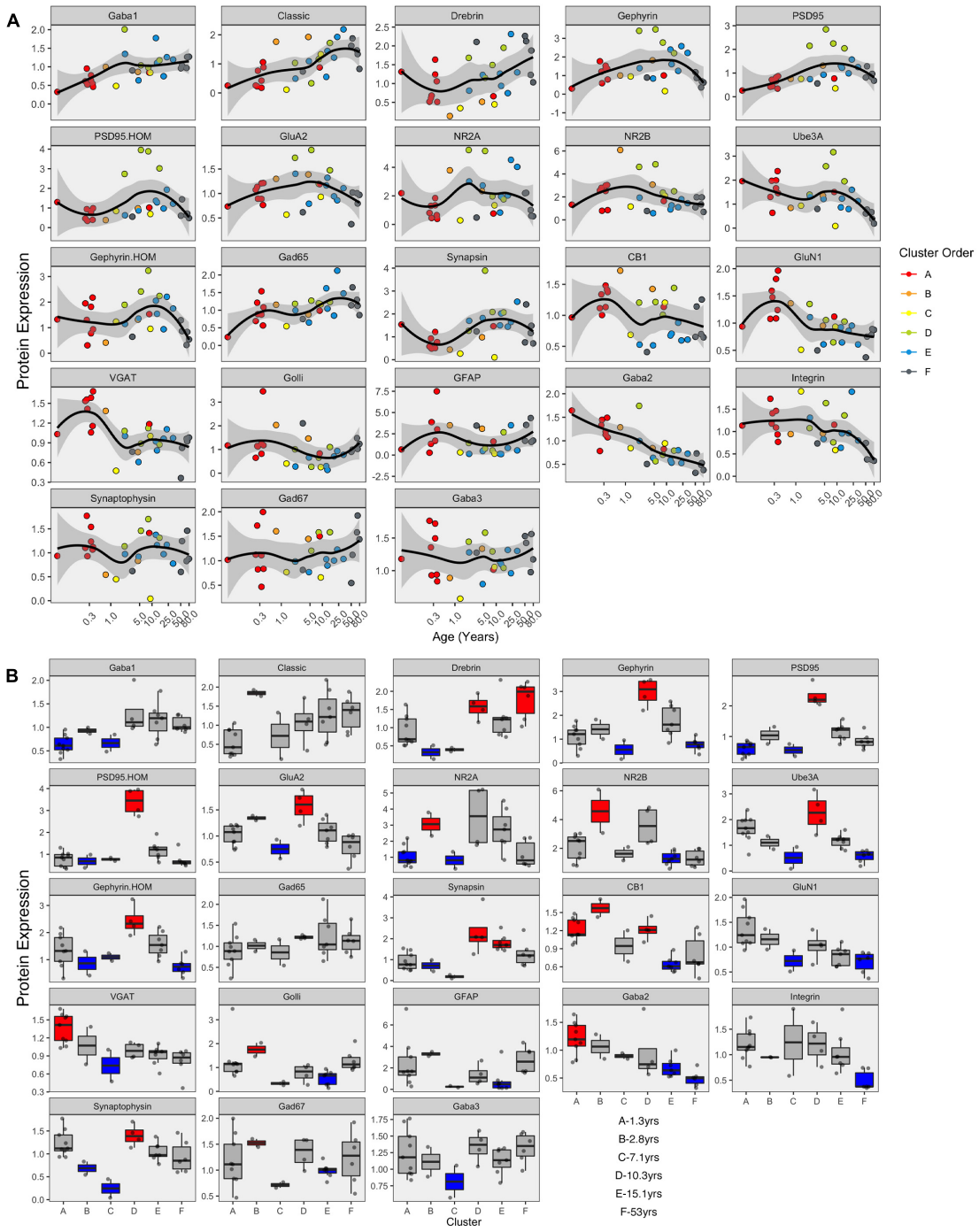

FIGURE 8 | Development of proteins by age and by cluster. Expression profiles for each of the 23 proteins **(A)**. Individual samples are represented by a single dot, and colored according to the corresponding RSKC cluster assignment. LOESS curves for each profile are shown in black, with 95% upper and lower confidence intervals shown bounding gray outline. Protein profiles are ordered according to similar developmental trajectories. **(B)** Overrepresentation analysis showing protein expression as boxplots representing each cluster. Over-represented clusters were colored red if the 25th percentile of the RSKC cluster was greater than the 95th percentile of a simulated normal distribution. Under-represented clusters were colored blue if the 75th percentile of the RSKC cluster was less than the 5th percentile of a simulated normal distribution. Boxes that fell within the middle 90% of the simulated normal distribution were left gray.

band) and when PCA measurements were greater (negative red band) (**Figure 11B**). The blue band was slightly offset from zero, indicating a bias for the normalized PCA $\cos^2$ values to be greater than the RSKC weights. The plot identified key proteins, such as the Gephyrin and PSD95 homogenates, and Ube3A which were more strongly represented by the RSKC weights. All three of

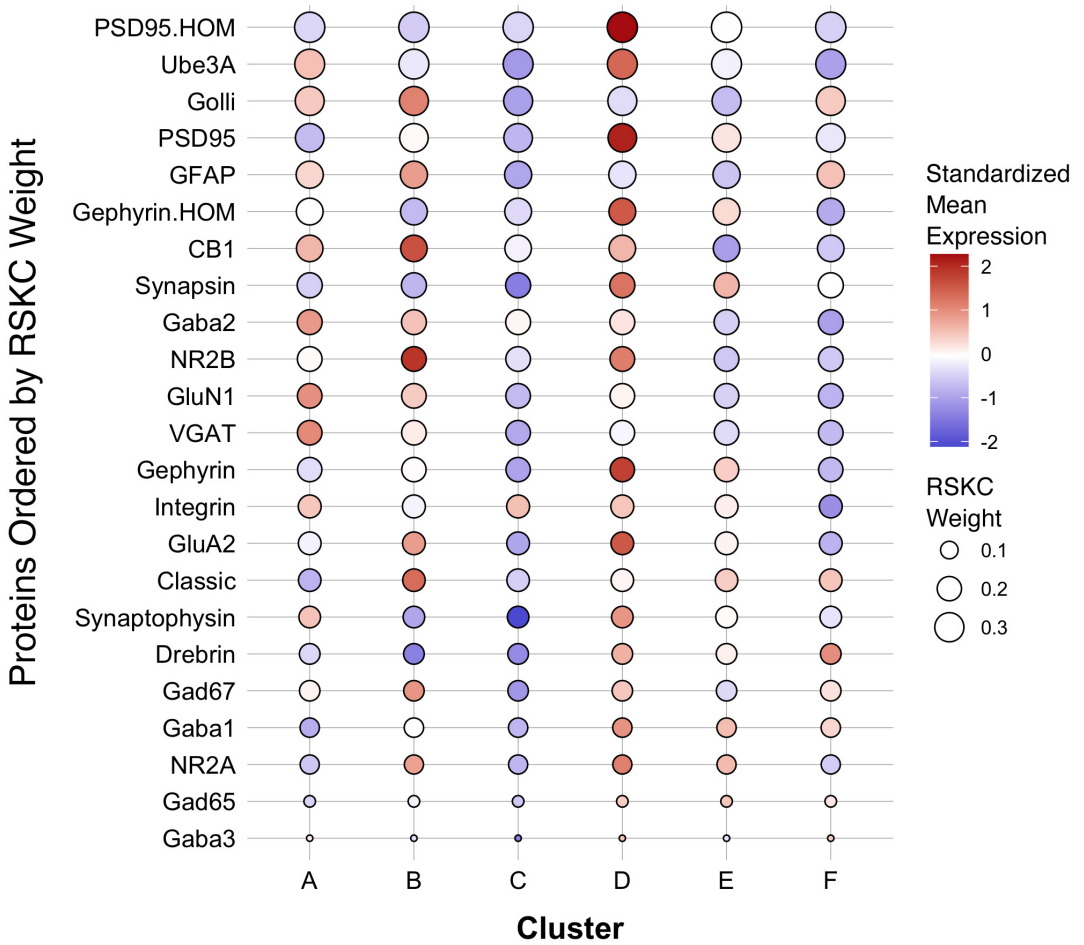

FIGURE 9 | Bubble plot of mean protein expression across each cluster. Proteins are ordered by their corresponding RSKC weight with the highest weighted protein arranged at the top, and the lowest weighted protein at the bottom and ordered by developmental cluster from left to right. The color of the dot represents standardized protein expression for each cluster, while the size of the dot represents the RSKC weight (see legend).

those proteins are essential molecular components that regulate the experience-dependent development of the visual cortex. For example, Ube3A is involved in the experience-dependent cycling of AMPA receptors (Greer et al., 2010), is required for ocular dominance plasticity (Yashiro et al., 2009; Sato and Stryker, 2010) and is selectively lost during aging of the human visual cortex (Williams et al., 2010).

Using Principal Component Analysis Basis Vectors to Identify Candidate Plasticity Features

The proteins in the dataset are known to regulate experience-dependent plasticity in the visual cortex (e.g. Quinlan et al., 1999a,b; Fagiolini et al., 2003, 2004; Hensch, 2004, 2005; Hensch and Fagiolini, 2005; McGee et al., 2005; Philpot et al., 2007; Yashiro and Philpot, 2008; Cho et al., 2009; Gainey et al., 2009; Smith et al., 2009; Kubota and Kitajima, 2010; Larsen et al., 2010; Levelt and Hübener, 2012; Lambo and Turrigiano, 2013; Cooke and Bear, 2014; Guo et al., 2017; Turrigiano, 2017; Hensch and Quinlan, 2018). We took advantage of that *a priori* knowledge and the output from PCA to identify a new set of features that could be used to probe the neurobiology of the RSKC clusters. Although the RSKC weights reflect the contribution of individual proteins for partitioning the samples into clusters, the weights do not provide insights into combinations and balances of proteins that regulate plasticity. Thus, it is necessary to add another analysis that can help to identify those networks and balances of proteins that regulate experience-dependent plasticity.

This step is a semi-supervised approach to select combinations of proteins using the PCA output ($\cos^2$ values and the basis vectors) and the known functions of the proteins. These steps generate a list of candidate plasticity features that are combined to construct an extended phenotype (Dawkins, 1982). We call the collection of features a plasticity phenotype and it can be used to infer the plasticity state of the visual cortex. The approach is described in detail in Balsor et al. (2020) and briefly outlined here.

Two heuristics were applied to identify combinations and balances among the proteins, using proteins that met the $\cos^2$ cutoff shown in **Figure 10H**. First, the *a priori* knowledge

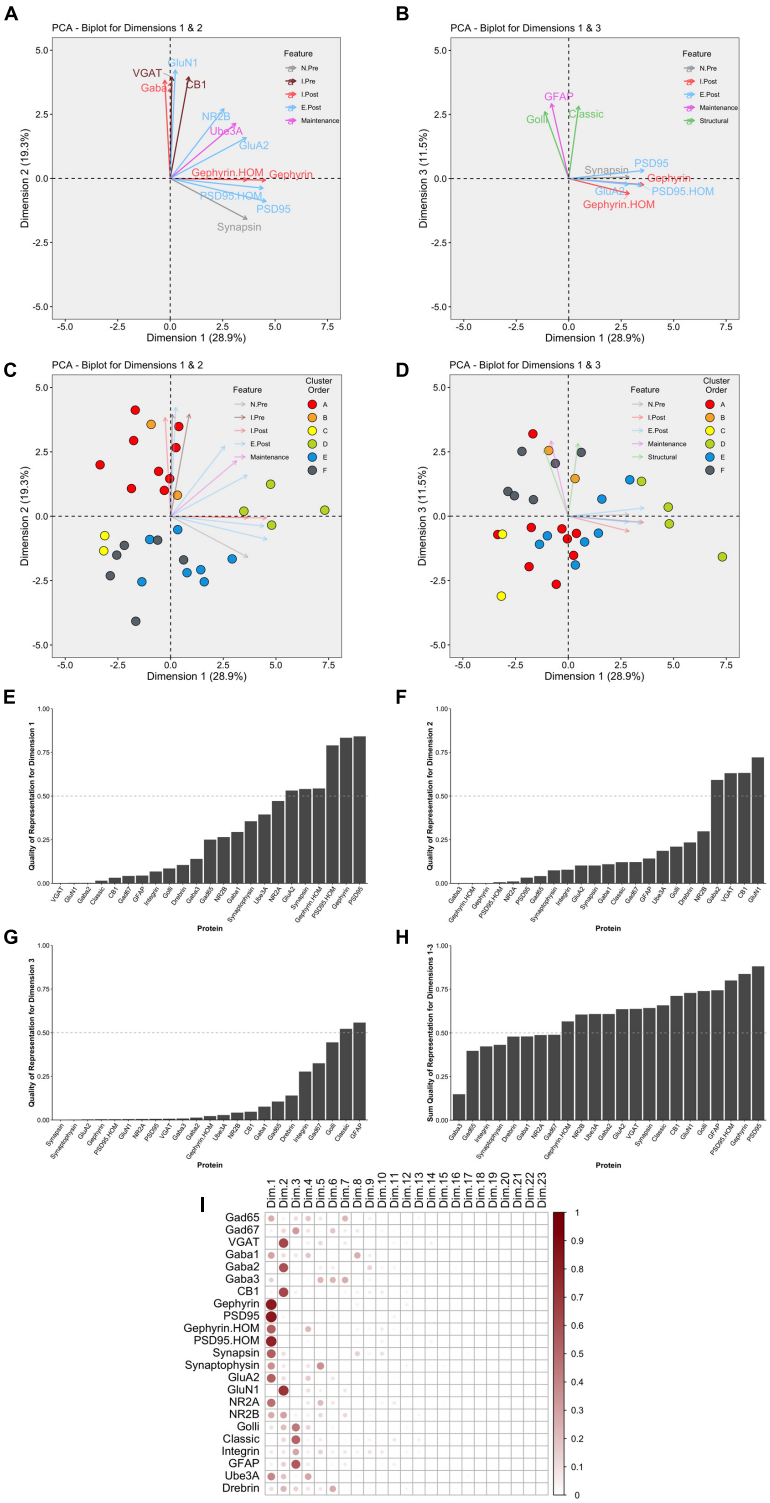

FIGURE 10 | Examination of relevant PCA identified dimensions. PCA biplots **(A–D)** show protein features as vectors (arrows) and individual samples as dots on pairings of PCA dimensions 1 and 2 **(A,C)** and dimensions 1 and 3 **(B,D)**. The strength of the representation (cos^2) for a protein on the given set of dimensions is reflected by the length of the vector, and only proteins with $cos^2 > 0.5$ are shown. The color of each point corresponds to their cluster, matching the original cluster colors in **Figure 6**. Bar plots represent the quality of representation of each protein with each dimension **(E–G)**, as well as the summed quality of representation across all three dimensions **(H)**. The dashed line represents cos^2 cutoff value for representation of 0.5. **(I)** Matrix illustrating the quality of representation for each protein with each PCA dimension, representing the strength (circle size) and direction (zero = white, positive = red) of cos^2.

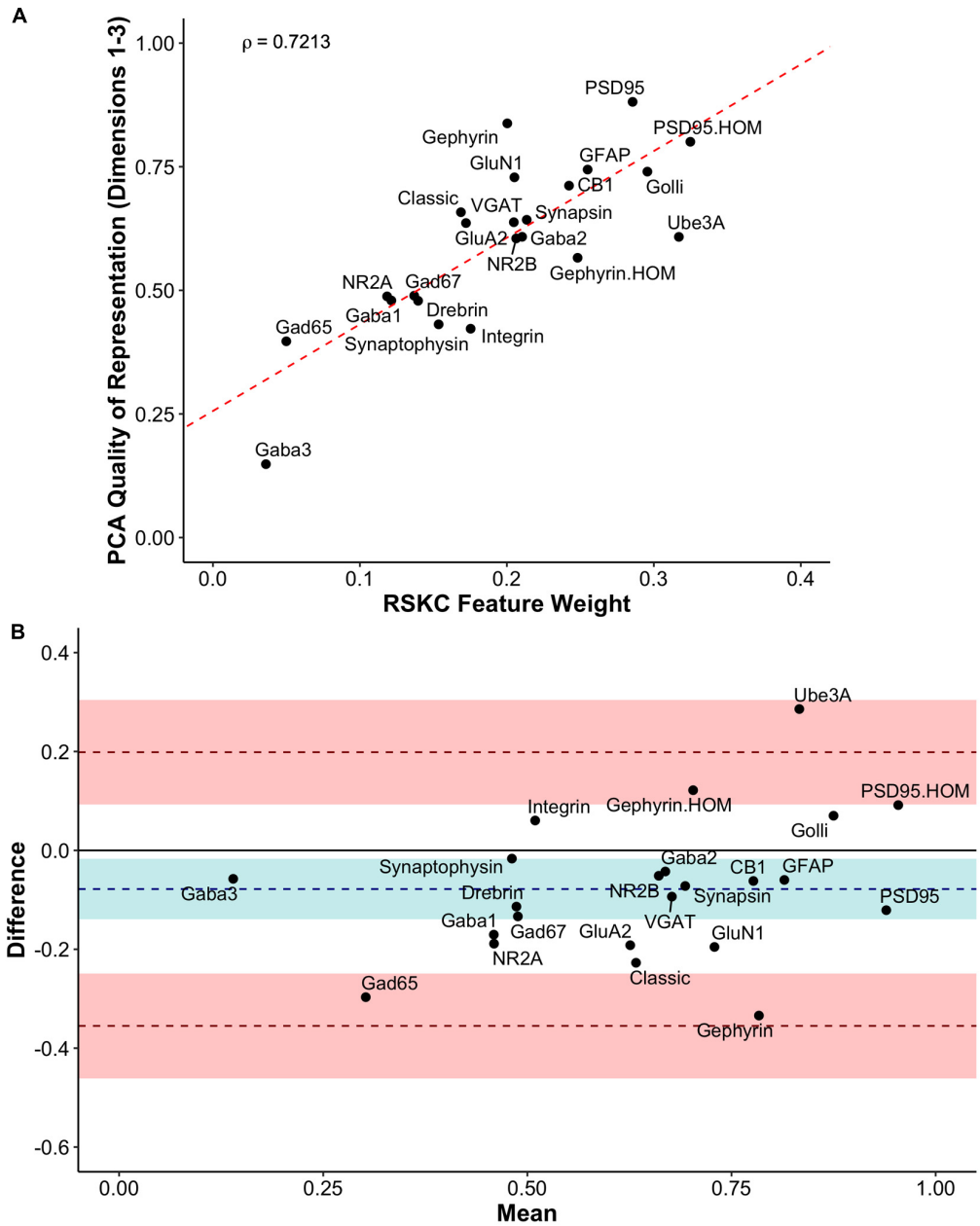

FIGURE 11 | Exploring the relationship between PCA and RSKC feature identification. **(A)** Scatter plot showing the PCA quality of representation (cos^2) for the first three dimensions and the RSKC weights. The dashed line represents the line of best fit, and rho is Spearman's rank correlation coefficient. **(B)** Bland–Altman plot comparing PCA cos^2 for the first three dimensions and RSKC weights for 23 proteins. The cos^2 and RSKC weights were each computed as proportions of their respective maximum values. The dashed blue line represents the mean difference, with the 95% confidence intervals shown as the blue shaded area. The top dashed red line represents the upper limit of agreement (+1.96 SD) and the bottom dashed red line is the lower limit of agreement (–1.96 SD), with corresponding 95% confidence intervals shown as red shaded areas.

about the function of the proteins in plasticity and development of the visual cortex was used to guide the inspection of the three basis vectors (**Figures 12A–C**). Second, the amplitude and direction of each protein on the basis vector were used to select candidate features to sum or use in a relative difference index. For example, on PC1, we noted that four highly conserved synaptic markers (Pinto et al., 2015), the pre-synaptic proteins synapsin and synaptophysin and the post-synaptic proteins PSD95 and gephyrin had large positive amplitudes, so they were summed to create one of the candidate features (PGSS). On PC2, the receptor subunits $GABA_A\alpha1$ and $GABA_A\alpha2$ had opposite directions, so these were used for an index ($GABA_A\alpha1$:$GABA_A\alpha2$) because the balance between those subunits is developmentally regulated and governs the kinetics of

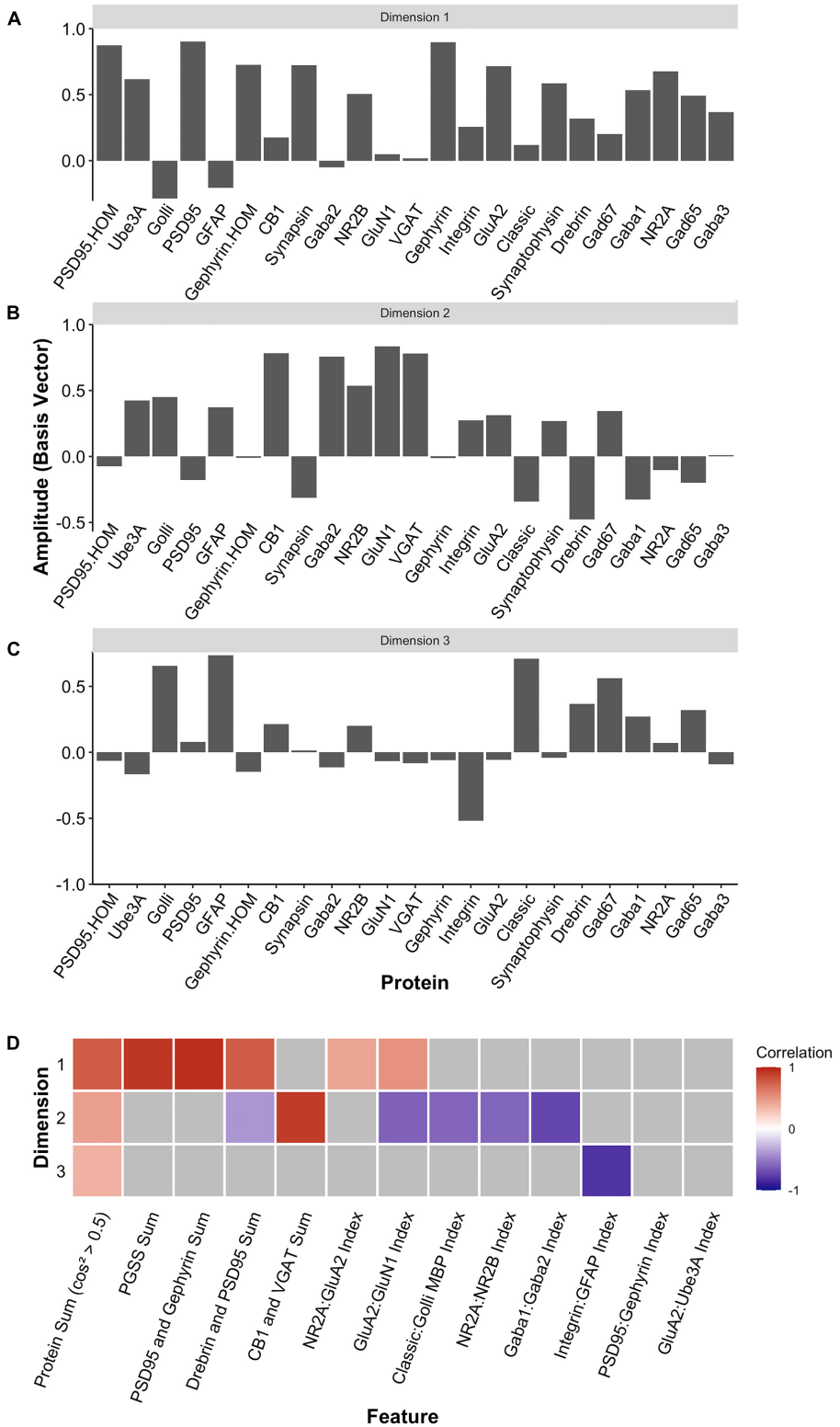

FIGURE 12 | Candidate feature identification using principal component analysis. Histograms showing the amplitude of the basis vector for each protein across **(A)** dimension 1, **(B)** dimension 2, and **(C)** dimension 3. **(D)** The correlations between the protein sums or indices and the first three PCA dimensions. Non-gray cells represent significant correlations after Bonferroni correction, with the color indicating the magnitude and direction of the correlation (negative = blue, zero = white, positive = red).

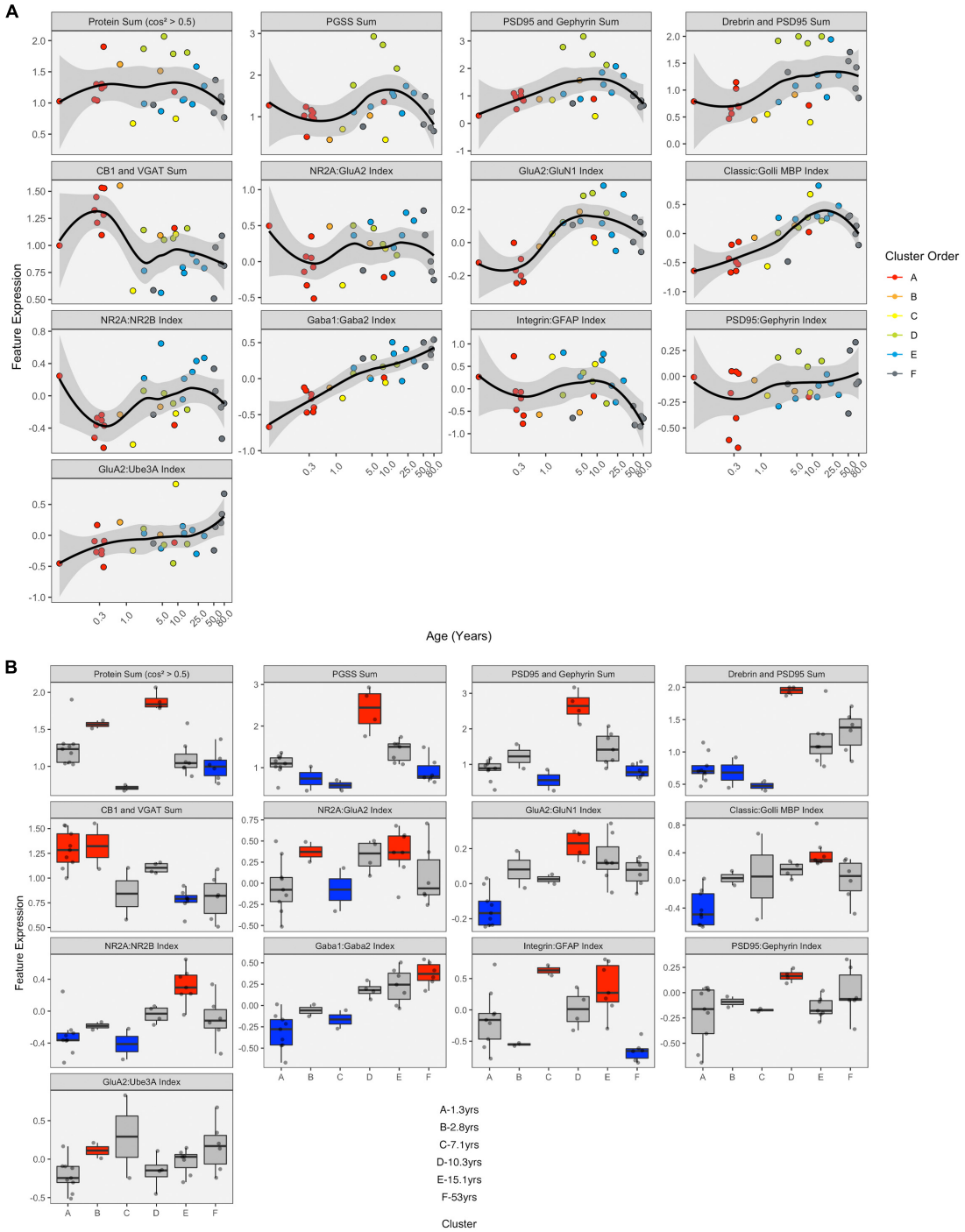

FIGURE 13 | Expression of extracted features with respect to age and by cluster. **(A)** LOESS trajectories illustrating the expression of protein sums and indices. Points are colored corresponding to the clusters in **Figure 8**, and the 95% confidence intervals around each curve are colored in gray. **(B)** Boxplots show the expression of each feature for the six clusters. A simulated normal distribution was sampled to obtain 5th and 95th percentile values. Boxes were colored red (i.e., over-represented) if the 25th percentile of the feature cluster was greater than the 95th percentile of the normal distribution. Boxes were colored blue (i.e., under-represented) if the 75th percentile of the feature cluster was less than the 5th percentile of the simulated distribution. Otherwise, boxes were colored gray.

the $GABA_A$ receptor (Gingrich et al., 1995; Bosman et al., 2002; Heinen et al., 2004; Hashimoto et al., 2009). Finally, on PC3, we noted that GFAP and integrin had the largest amplitudes, and they were in opposite directions. Those two proteins are expressed by astrocytes, and the expression of integrin receptors is increased on reactive astrocytes (Lagos-Cabré et al., 2020), so

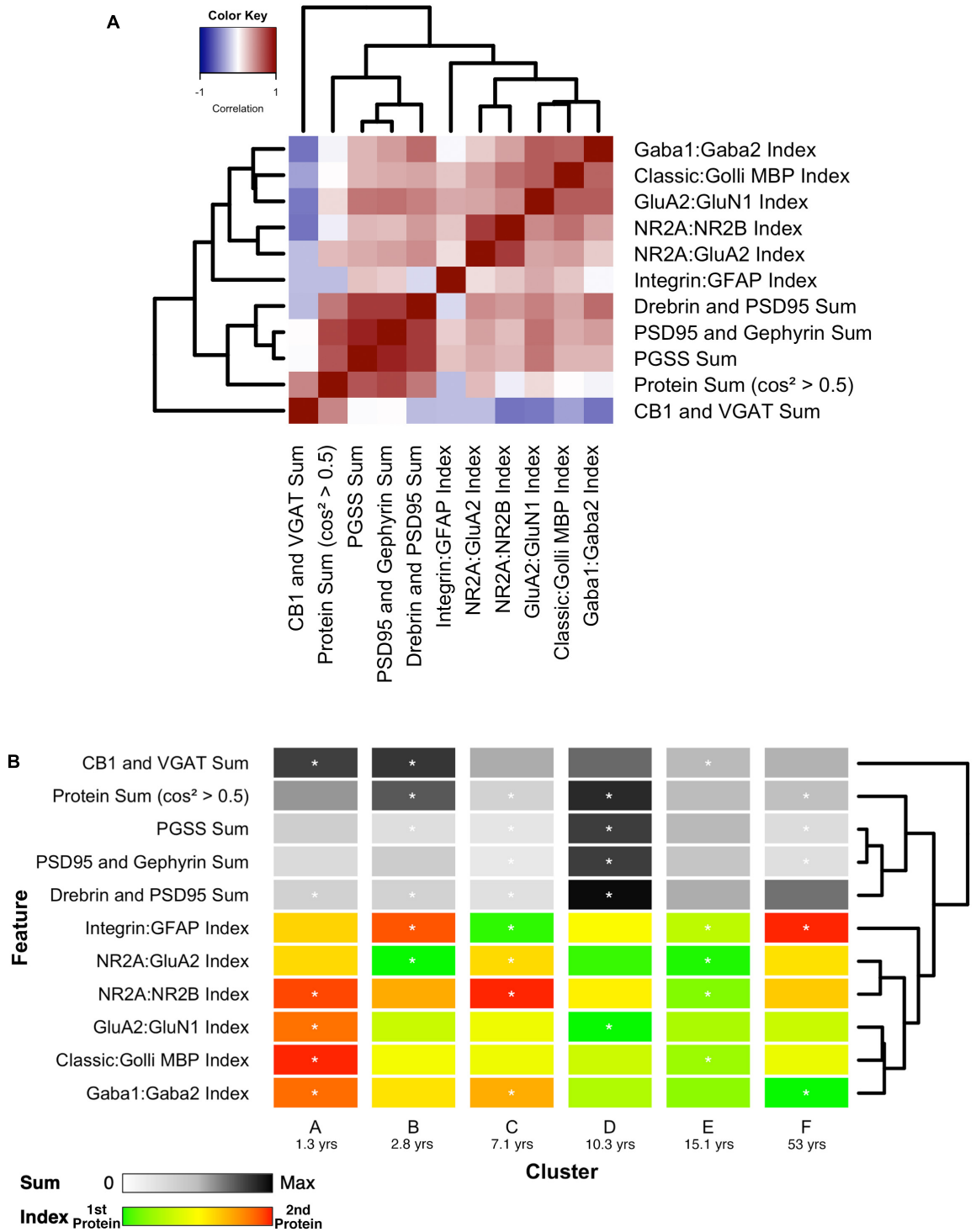

FIGURE 14 | Associations between selected features and feature phenotype by cluster. **(A)** Correlation heat map between protein sums and indices, with strength and direction of Pearson's R correlation represented by the color (negative = blue, zero = white, positive = red), and arranged by similar pairwise correlations using a wrapped dendrogram. Features were selected if they were significantly correlated with any of the first three PCA dimensions. **(B)** The plasticity phenotype was visualized using color-coded horizontal bars representing the median expression of selected features across clusters. For protein sums, the color ranges from white (zero) to gray (midpoint) to black (maximum protein sum across all features). For the protein indices, the color ranges from green (favoring the first protein in the index) to yellow (balance of the two proteins) to red (favoring the second protein in the index). Asterisks indicate features that were found to be either over- or under-represented. The features are arranged according to the same dendrogram generated in **(A)**.

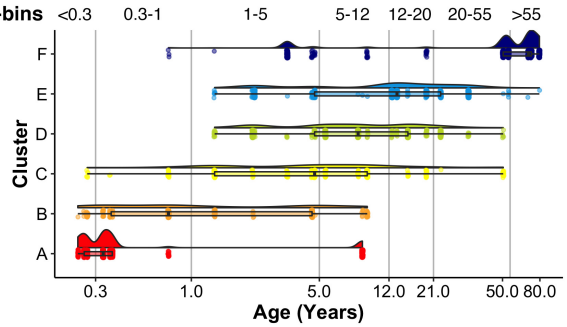

FIGURE 15 | Raincloud plot showing the distribution of sample ages by cluster. Vertical lines correspond to the bounds of pre-defined age bins. Samples corresponding to the youngest age (i.e., 20 days) are not shown.

an index was calculated (GFAP:Integrin). Applying the heuristics resulted in 13 candidate features, including 5 protein sums identified using the basis vector for PC1 and 8 indices from PC2 and PC3 (**Figure 12D**). The features were validated by calculating each feature using the expression values for the 23 proteins (**Supplementary Table 6**) and correlating those with the eigenvalues for the three PC dimensions (**Figure 12D**).

LOESS curves and boxplots were made for all of the candidate features to illustrate how they changed across the lifespan and identify if features were over- or under-represented in a cluster (**Figures 13A,B** and **Supplementary Figure 7**). One aspect of development apparent in the boxplots was the over-representation of the protein sums in cluster D. That cluster has a mean age of 10.3 years (SD 8.4 years), which corresponds with the end of the critical period for developing amblyopia in children (Lewis and Maurer, 2005; Birch, 2013) and a stage of human cortex development often described by synaptic exuberance, growth, and changing state of plasticity. Furthermore, animal research has shown that excess excitation (Fagiolini and Hensch, 2000; Fagiolini et al., 2004) and expression of proteins regulating that activity, especially PSD95 (Huang et al., 2015), can close the critical period.

Analyzing Plasticity Phenotypes for the Robust Sparse *k*-Means Clusters

Finally, the 11 features with significant correlations were used to construct a plasticity phenotype that was combined with the six clusters. A correlation matrix was made using the values for the features calculated from the protein expression for each sample (**Supplementary Table 1**). The matrix and surrounding dendrogram showed that the protein sum and indices were separated into different tree branches. The order of the features in the correlation matrix was used for the bands in the plasticity phenotype visualization. In the phenotype, the median of each feature was represented as a color-coded band for the six clusters (**Figure 14B**). Together, the 66 color-coded feature bands captured the high dimensional pattern of neurobiological changes across the lifespan. The protein sums represented by gray levels convey a pattern with specific groups of proteins that are highly expressed early in development (clusters A and B) and a broad wave of expression in older childhood (cluster D). The indices reflect the multiple timescales of molecular development that are the hallmarks of the human visual cortex (Siu and Murphy, 2018). However, even with undulating features and different timescales, all appear to arrive at a similar level of maturation in cluster E.

Combining the features and clusters into a visualization simplified this complex dataset and facilitated linking the clusters with sets of neurobiologically meaningful features. The asterisks on the feature bands indicate the ones identified as over- or under-represented in **Figure 13B**. Each cluster had a unique group of features that deviate from the average, and those represent the neurobiological mechanisms that differentiate the age-related clusters. For example, the set of 4 red bands for the young visual cortex (cluster A) was unique and showed that the indices were dominated by the NMDA receptor subunits NR2B and GluN1, the Golli family of myelin basic protein (MBP) and the $GABA_A\alpha2$ receptor subunit. In contrast, the older visual cortex (cluster F) was distinguished by a set of 3 light gray protein sum bands, a red band indicating more GFAP and a green band indicating more $GABA_A\alpha1$. Finally, the overall appearance of the protein sums and indices for cluster D gives the impression of a transition stage in the development of the visual cortex when exuberant protein expression (dark gray bands) (Huang et al., 2015) and the shift in protein balances (green bands) (e.g., Quinlan et al., 1999a,b; Fagiolini and Hensch, 2000; Chen et al., 2001; Philpot et al., 2001; Fagiolini et al., 2003, 2004; Hensch, 2005; Hall and Ghosh, 2008) signals the end of the critical period.

A raincloud plot of the samples in the 6 clusters shows the range of ages that correspond with the plasticity phenotypes (**Figure 15**). The distribution of sample ages in the clusters appears like a series of overlapping waves extending beyond the ages of the traditional pre-defined age-bins included in **Figure 15** as vertical lines. For example, for cluster D, the wave's peak falls into the age-bin associated with the end of the period for developing amblyopia (5–12 years). However, cluster D also includes younger and older samples suggesting that the end of the sensitive period may not occur uniformly among individuals. Furthermore, other clusters overlap the 5- to 12-year-old age-bin suggesting that multiple phenotypes can be found during certain age-bins. Thus, the cluster analysis helped reveal aspects of visual cortex development that are obscured by using pre-defined age bins, which is that chronological- and brain-age often diverge (Cole et al., 2019).

DISCUSSION

The current study shows that the application of sparse clustering leverages the high dimensional nature of proteomic and transcriptomic data from human brain development to find age-related clusters that are spread across the lifespan. In particular, RSKC using measurements of proteins or genes from the human visual cortex partitioned samples into clusters that progressed from neonates to older adults. The iterative reweighting of the measurements to focus on the proteins or genes that carry the

most information about lifespan changes led to robust age-related clustering of the data. Furthermore, especially for the datasets focusing on 95 proteins or genes, the clusters represented early development, young childhood, older childhood, adolescence, and adulthood. Thus, sparse clustering provides a robust approach for identifying proteomic or transcriptomic defined brain ages that overlap with behavioral and brain imaging findings of gradual and prolonged human brain maturation.

Many factors come into play when selecting an appropriate clustering algorithm for a study. Here, we considered the goal of the study (to resolve sometimes subtle age-related changes in molecular mechanism), the structure of the dataset ($p \sim n$ to $p >> n$), and the output of the algorithm (is it just the clusters or is feature selection included). Sparse K-means clustering was selected because it fit all of those considerations. We know from previous studies of the molecular development of the human brain that there can be subtle differences between age groups (Murphy et al., 2005; Pinto et al., 2010, 2015; Williams et al., 2010; Siu et al., 2015, 2017), and yet even small changes in protein or gene expression will alter neural function. Therefore, we looked for algorithms designed for omics datasets where subtle changes in a subset of the genes or proteins would identify important characteristics of the data. The development of sparse K-means clustering by Witten and Tibshirani (2010) was partially inspired by the need to better cluster a breast cancer dataset. In that dataset, subtle differences in gene expression significantly impacted patient outcomes, but standard clustering approaches did not pick those up. In addition, sparse clustering was developed to address datasets, like ours and the breast cancer data where the structure is $p \sim n$ to $p >> n$. Sparse K-means clustering is a good fit for those high dimensional structures because it minimizes the within-cluster sum of squares with a dissimilarity measure while maximizing the between-cluster sum of squares by iteratively reweighting the measures. Finally, and most importantly, sparse K-means clustering performs feature selection. The examples in this paper show the reweighted proteins and those distributions identifying how much each protein contributes to partitioning the samples into clusters. That matrix is sparse, with unimportant proteins having near-zero weights and important ones having non-zero weights. Those weights are essential for cluster analysis to help with making neurobiologically relevant interpretations of brain development from the cluster analysis.

Various other algorithms, including linear and non-linear dimension reduction [e.g., tSNE, multidimensional scaling (MDS), and PCA], can separate developmental samples. In this paper, we found that both tSNE and PCA show some age-related progress in the arrangement of the samples. Also, Kang et al. (2011) used MDS to separate the samples across MDS 1 and 2. Then the points were color-coded by pre-defined age bins to show a left to right flow from early prenatal to older adults. However, it was not apparent which genes mapped on to those dimensions. The selection of features in the form of the weights is a key difference between sparse *k*-means clustering and standard clustering approach that was critical for the current study.

The current study is not exhaustive of clustering approaches, as the number of unsupervised clustering algorithms for analyzing high dimensional data is rapidly expanding. For example, new sparse clustering algorithms include innovation at the level of the lasso-type penalty used to adjust observation weights (Brodinová et al., 2019). Accordingly, the "best" algorithm for understanding molecular brain development will continue to change as new approaches are developed. Rather than acting as a prescriptive guide for which algorithm to use, the current study highlights the challenges raised when applying high dimensional clustering to studies using postmortem brain samples. In particular, developmental studies that use postmortem human brain tissue often have more measurements than samples ($p > n$) and require clustering algorithms optimized for high dimensional data structures. The examples showed that the RSKC algorithm worked well for a wide range of observations (p) from 7 to 17,237. However, the age-related progression of the 95 proteins and 88 gene datasets (**Figures 7A,B**) were more distinct than the clustering using 988 SynGO or the full 17,237 gene dataset (**Figures 7C,D**).

The succession of age-related clusters found for the visual cortex aligns with some critical milestones in visual development. Using measurements of molecular mechanisms that regulate experience-dependent plasticity, the clusters illustrated in **Figure 5** show that cluster A overlaps the start of the sensitive period for binocular vision at 4–6 months and cluster B the peak of that sensitive period at 1–3 years (Banks et al., 1975). Furthermore, cluster D aligns with the maturation of contrast sensitivity (Ellemberg et al., 1999), motion perception (Ellemberg et al., 2002), and the end of the period for the susceptibility of developing amblyopia (6–12 years) (Epelbaum et al., 1993; Keech and Kutschke, 1995; Lewis and Maurer, 2005). The oldest cluster, cluster F, highlights ages when cortical changes reduce performance on several visual tasks (Owsley, 2011). The alignment with visual milestones suggests that the clusters might provide insights into the molecular mechanisms that regulate various aspects of visual development and visual function dynamics across the lifespan. Notably, the molecular mechanisms are well studied in animal models. Thus, this information for the human cortex may be seen as a bridge linking results from animal studies with human neurobiology that can help interpret brain imaging and visual perception findings.

By combining the RSKC clustering with PCA, we identified plasticity-related features and constructed a plasticity phenotype that was applied to each cluster (**Figure 14**). The term *plasticity phenotype* has been used before to describe the waxing and waning of gene expression in the developing brain (Smith et al., 2019). Here we used the term to describe an extended phenotype (Dawkins, 1982) because the proteins in the dataset have known functions in regulating experience-dependent plasticity in the visual cortex (e.g., Quinlan et al., 1999a,b; Fagiolini et al., 2003, 2004; Hensch, 2004, 2005; McGee et al., 2005; Philpot et al., 2007; Yashiro and Philpot, 2008; Cho et al., 2009; Gainey et al., 2009; Smith et al., 2009; Kubota and Kitajima, 2010; Larsen et al., 2010; Levelt and Hübener, 2012; Lambo and Turrigiano, 2013; Cooke and Bear, 2014; Guo et al., 2017; Turrigiano, 2017). Thus, the plasticity phenotype can be used to infer the potential for experience-dependent plasticity in the different clusters and provide a new perspective on the maturation of the human visual cortex.

Each cluster had a unique set of features that were over- or under-represented in the plasticity phenotype, and those features were apparent in the phenotype visualization. Notably, the features were selected using a semi-supervised approach with a series of heuristics that included protein combinations and balances known to regulate experience-dependent plasticity. As a result, the unique sets of features can be compared with the literature to infer the likely state of experience-dependent plasticity for a cluster. For example, balances in the youngest cluster (A) were dominated by receptors that are known to facilitate experience-dependent plasticity in the visual cortex (Kleinschmidt et al., 1987; Quinlan et al., 1999a,b; Philpot et al., 2001; Fagiolini et al., 2003, 2004; Iwai et al., 2003; Hensch, 2004; Cho et al., 2009; Jiang et al., 2010). In contrast, cluster D overlaps the end of the period of susceptibility to develop amblyopia, and has peaks in protein expression, especially PSD95 that are known to close the critical period in animal models (Huang et al., 2015). The features in cluster D also appeared to mark the transition from juvenile features found in clusters A, B, and C to the mature and aging patterns in clusters E and F. Moreover, the range of ages in a cluster appeared as a series of overlapping waves in the raincloud plot, thereby illustrating that chronological- and brain-age have a complex relationship.

Clustering the data collected from human postmortem tissue samples to reveal the age-related progression in the brain's molecular complexity is just the start of using high dimensional analyses. The application of modern exploratory data-driven approaches reveals novel aspects of human brain development, such as the risk for mental illness (Li et al., 2018) or divergence from other primates (Zhu et al., 2018). Identifying an appropriate high dimensional clustering technique opens the door to many other downstream analyses to interrogate different clusters' molecular makeup. A critical benefit of clustering with RSKC is that it outputs the feature weights. Those weights reveal the impact of specific proteins or genes on differentiating the brain's molecular environment during the progression of lifespan stages. Those proteins and genes can be used as the input to Gene Ontology (GO) analysis to catalog the molecular processes, cellular components, and biological processes that dominate the stages. Or the opposite can be done as shown in the paper where the 988 genes corresponding to the SynGO database were used to cluster the samples. The clusters can also be used for differential gene expression analysis to highlight which features are enriched during various lifespan stages. For example, the top-weighted molecular features from the RSKC analysis may be useful for creating a phenotype that provides a biologically meaningful characterization of the high dimensional changes that occur in different stages of the lifespan (Balsor et al., 2020).

An interesting finding of the current study is the overlapping ages among the clusters. While this may be viewed as imperfect partitioning of samples by the clustering algorithms, it may also reflect the human brain development's true heterogeneity. In other words, developmental periods may not necessarily be described by a single omic phenotype. Instead, the classically defined developmental stages may be characterized by two or more distinct patterns of gene or protein expression in the brain. This molecular heterogeneity may shed light on findings such as the substantial inter-individual variation in cortical responses measured by fMRI studies in infants (Born et al., 2000). Also, the overlapping ages among clusters may reflect periods of stationary fluctuations in the brain's developmental trajectory, representing transitions from one molecular state to the next, similar to language development models (Sanchez-Alonso and Aslin, 2020).

Addressing how human brain development proceeds is an important question that will require large amounts of new data and algorithms that capture the local and global structure of high dimensional trajectories, including ones with gradual noisy changes and non-linear transitions. One approach could include repeated MRI measurements during the ages that overlap among molecular clusters to assess if those ages have heightened intra- or inter-individual variation in brain responses. Those studies will help identify ages during development with gradual but noisy change from ages with non-linear transitions in the gene and protein expression pattern in the developing human brain. Ultimately, the models will need to include multi-omics data and link with brain imaging to understand how the human brain develops fully.

CONCLUSION

The last decade has seen remarkable growth in the number of studies examining the human brain's molecular features. In parallel, high throughput tools have dramatically increased the amount of data collected for every sample. The complexity and high dimensional nature of those datasets have spurred the need for more guidance in selecting appropriate tools to analyze those big data. Some studies are now collecting data from 100 or 1000 s of human brain postmortem samples (e.g., PsychENCODE), but studies of development still have many fewer tissue samples, and the ages of the cases are spread across the lifespan. The small sample sizes of the developmental datasets make it difficult to apply many commonly used high dimensional clustering methods. Those methods lack the sensitivity needed to reveal robust clusters defined by the subtle differences in genes or proteins that occur across the postnatal lifespan. At the same time, sparsity-based clustering algorithms designed for small sample size have emerged. In this guide, we explored the application of sparsity-based clustering and showed that one algorithm, RSKC, is a good fit for revealing the subtle and gradual changes of human brain development that occurs from birth to aging. In the next decade, the amount of data collected from each postmortem brain sample will only continue to grow as single-cell RNA sequencing methods are applied to studying human brain development. Furthermore, the push to integrate multimodal measurements, from molecules to imaging of human brain development will heighten the demand for robust high dimensional analysis tools. Neuroscientists will continue to face many challenges identifying rigorous methods to analyze those sparse and very high dimensional datasets. Nevertheless, careful selection of high dimensional analytical techniques that are designed for small sample sizes can be expected to have an impact on the discovery of novel aspects of human brain development.

DATA AVAILABILITY STATEMENT

The data used to support the findings in this manuscript and code are available at the following link: https://osf.io/6vgrf/.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Hamilton Integrated Research Ethics Board (HiREB) represents the institutions of Hamilton Health Sciences, St. Joseph's Healthcare Hamilton, Research St. Joseph's-Hamilton and the Faculty of Health Sciences at McMaster University and operates in compliance with and is constituted in accordance with the requirements of: the Tri-Council Policy Statement on Ethical Conduct of Research Involving Humans; the International Conference on Harmonization of Good Clinical Practices; Part C Division 5 of the Food and Drug Regulations of Health Canada, and the provisions of the Ontario Personal Health Information Protection Act 2004 and its applicable Regulations. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

JB designed the research, performed the research, analyzed the data, and wrote and revised the manuscript. KA and KM designed the research, analyzed the data, and wrote and revised the manuscript. DS performed the research, analyzed the data, and wrote and revised the manuscript. RK and JZ performed the research, analyzed the data, and revised the manuscript. EJ analyzed the data and revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

NSERC Grant RGPIN-2015-06215 and RGPIN-2020-06403 were awarded to KM, Woodburn Heron OGS was awarded to JB and KA, and NSERC CGS-M was awarded to EJ. The funder had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

ACKNOWLEDGMENTS

We thank Brendan Kumagai for help with some of the coding. Previous studies from our lab collected the proteomic data, and the human postmortem tissue samples were obtained from the NIH NeuroBioBank (Pinto et al., 2010, 2015; Williams et al., 2010; Siu et al., 2015, 2017). The transcriptomic data were from Kang et al. (2011) (GSE25219).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2021.668293/full#supplementary-material

REFERENCES

Adepeju, M., Langton, S., and Bannister, J. (2020). *akmedoids: Anchored Kmedoids for Longitudinal Data Clustering*. Available online at: https://CRAN.R-project.org/package=akmedoids (accessed April 13, 2021).

Aggarwal, C. C., Wolf, J. L., Yu, P. S., Procopiuc, C., and Park, J. S. (1999). Fast algorithms for projected clustering. *Sigmod. Rec.* 28, 61–72. doi: 10.1145/304181.304188

Agrawal, R., Gehrke, J., Gunopulos, D., and Raghavan, P. (1998). Automatic subspace clustering of high dimensional data for data mining applications. *Sigmod. Rec.* 27, 94–105. doi: 10.1145/276304.276314

Balsor, J. L., Ahuja, D., Jones, D. G., and Murphy, K. M. (2020). A Primer on Constructing Plasticity Phenotypes to Classify Experience-Dependent Development of the Visual Cortex. *Front. Cell Neurosci.* 14:245. doi: 10.3389/fncel.2020.00245

Balsor, J. L., Jones, D. G., and Murphy, K. M. (2019). Classification of Visual Cortex Plasticity Phenotypes following Treatment for Amblyopia. *Neural. Plast.* 2019, 2564018–2564023. doi: 10.1155/2019/2564018

Banks, M. S., Aslin, R. N., and Letson, R. D. (1975). Sensitive period for the development of human binocular vision. *Science* 190, 675–677.

Bellman, R. (1983). Mathematical Methods in Medicine. *Math. Methods Med.* 1983, 249–252. doi: 10.1142/9789814503013_bmatter

Bernard, C. (2019). Changing the Way We Report, Interpret, and Discuss Our Results to Rebuild Trust in Our Research. *Eneuro* 6:2019. doi: 10.1523/eneuro.0259-19.2019

Beston, B. R., Jones, D. G., and Murphy, K. M. (2010). Experience-dependent changes in excitatory and inhibitory receptor subunit expression in visual cortex. *Front. Synaptic Neurosci.* 2:138. doi: 10.3389/fnsyn.2010.00138

Birch, E. E. (2013). Amblyopia and binocular vision. *Prog. Retin. Eye Res.* 33, 67–84. doi: 10.1016/j.preteyeres.2012.11.001

Born, A. P., Miranda, M. J., Rostrup, E., Toft, P. B., Peitersen, B., Larsson, H. B. W., et al. (2000). Functional Magnetic Resonance Imaging of the Normal and Abnormal Visual System in Early Life. *Neuropediatrics* 31, 24–32. doi: 10.1055/s-2000-15402

Bosman, L. W. J., Rosahl, T. W., and Brussaard, A. B. (2002). Neonatal development of the rat visual cortex: synaptic function of GABAa receptor α subunits. *J. Physiol.* 545, 169–181. doi: 10.1113/jphysiol.2002.026534

Bray, S. (2017). Age—associated patterns in gray matter volume, cerebral perfusion and BOLD oscillations in children and adolescents. *Hum. Brain Mapp.* 38, 2398–2407. doi: 10.1002/hbm.23526

Breen, M. S., Ozcan, S., Ramsey, J. M., Wang, Z., Ma'ayan, A., Rustogi, N., et al. (2018). Temporal proteomic profiling of postnatal human cortical development. *Transl. Psychiat.* 8, 267. doi: 10.1038/s41398-018-0306-4

Brodinová, Š., Filzmoser, P., Ortner, T., Breiteneder, C., and Rohm, M. (2019). Robust and sparse *k*-means clustering for high-dimensional data. *Adv. Data Anal. Classi.* 13, 905–932. doi: 10.1007/s11634-019-00356-9

Carlyle, B. C., Kitchen, R. R., Kanyo, J. E., Voss, E. Z., Pletikos, M., Sousa, A. M. M., et al. (2017). A multiregional proteomic survey of the postnatal human brain. *Nat. Neurosci.* 20, 1–15. doi: 10.1038/s41593-017-0011-2

Chang, W. C. (1983). On Using Principal Components before Separating a Mixture of Two Multivariate Normal Distributions. *J. R. Stat. Soc. Series C* 32, 267–275. doi: 10.2307/2347949

Charrad, M., Ghazzali, N., Boiteau, V., and Niknafs, A. (2014). NbClust?: An R Package for Determining the Relevant Number of Clusters in a Data Set. *J. Stat. Softw.* 61:6. doi: 10.18637/jss.v061.i06

Charrad, M., Ghazzali, N., Boiteau, V., and Niknafs, A. (2015). *NbClust: Determining the Best Number of Clusters in a Data Set*. Available online at: https://CRAN.R-project.org/package=NbClust (accessed date 13-April-2015).

Chen, L. L., Yang, C. C., and Mower, G. D. G. (2001). Developmental changes in the expression of GABA(A) receptor subunits (alpha(1), alpha(2), alpha(3)) in the cat visual cortex and the effects of dark rearing. *Mol. Brain Res.* 88, 135–143. doi: 10.1016/s0169-328x(01)00042-0

Cho, K. K. A., Khibnik, L., Philpot, B. D., and Bear, M. F. (2009). The ratio of NR2A/B NMDA receptor subunits determines the qualities of ocular dominance plasticity in visual cortex. *Proc. National. Acad. Sci.* 106, 5377–5382. doi: 10.1073/pnas.0808104106

Colantuoni, C., Lipska, B. K., Ye, T., Hyde, T. M., Tao, R., Leek, J. T., et al. (2011). Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature* 478, 519–523. doi: 10.1038/nature10524

Cole, J. H., Marioni, R. E., Harris, S. E., and Deary, I. J. (2019). Brain age and other bodily 'ages': implications for neuropsychiatry. *Mol. Psychiatr.* 24, 266–281. doi: 10.1038/s41380-018-0098-1

Cooke, S. F., and Bear, M. F. (2014). How the mechanisms of long-term synaptic potentiation and depression serve experience-dependent plasticity in primary visual cortex. *Philosophical. Trans. R. Soc. B Biol. Sci.* 369, 20130284–20130284. doi: 10.1098/rstb.2013.0284

Dawkins, R. (1982). *The Extended Phenotype*. Oxford: Oxford University Press.

Disorder, W. C., Disorder, A.-D., Disorder, A. S., Disorder, B., Disorder, M. D., Groups, O.-C., et al. (2021). Virtual Histology of Cortical Thickness and Shared Neurobiology in 6 Psychiatric Disorders. *JAMA Psychiat.* 78, 47–63. doi: 10.1001/jamapsychiatry.2020.2694

Donaldson, J., and Donaldson, M. J. (2010). *Package "tsne."* Available online at: https://cran.r-project.org/package=tsne.

Duncan, C. E., Webster, M. J., Rothmond, D. A., Bahn, S., Elashoff, M., and Weickert, C. S. (2010). Prefrontal GABA(A) receptor alpha-subunit expression in normal postnatal human development and schizophrenia. *J. Psychiatr. Res.* 44, 673–681. doi: 10.1016/j.jpsychires.2009.12.007

Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci.* 95, 14863–14868. doi: 10.1073/pnas.95.25.14863

Ellemberg, D., Lewis, T. L., Liu, C. H., and Maurer, D. (1999). Development of spatial and temporal vision during childhood. *Vis. Res.* 39, 2325–2333.

Ellemberg, D., Lewis, T. L., Maurer, D., Brar, S., and Brent, H. P. (2002). Better perception of global motion after monocular than after binocular deprivation. *Vis. Res.* 42, 169–179. doi: 10.1016/s0042-6989(01)00278-4

Epelbaum, M., Milleret, C., Buisseret, P., and Dufier, J. L. (1993). The sensitive period for strabismic amblyopia in humans. *Ophthalmology* 100, 323–327.

Fagiolini, M., Fritschy, J.-M., Löw, K., Möhler, H., Rudolph, U., and Hensch, T. K. (2004). Specific GABAA circuits for visual cortical plasticity. *Science* 303, 1681–1683. doi: 10.1126/science.1091032

Fagiolini, M., and Hensch, T. K. (2000). Inhibitory threshold for critical-period activation in primary visual cortex. *Nature* 404, 183–186. doi: 10.1038/35004582

Fagiolini, M., Katagiri, H., Miyamoto, H., Mori, H., Grant, S. G. N., Mishina, M., et al. (2003). Separable features of visual cortical plasticity revealed by N-methyl-D-aspartate receptor 2A signaling. *Proc. Natl. Acad. Sci.* 100, 2854–2859. doi: 10.1073/pnas.0536089100

Gainey, M. A., Hurvitz-Wolff, J. R., Lambo, M. E., and Turrigiano, G. G. (2009). Synaptic scaling requires the GluR2 subunit of the AMPA receptor. *J. Neurosci.* 29, 6479–6489. doi: 10.1523/jneurosci.3753-08.2009

Giavarina, D. (2015). Understanding Bland Altman analysis. *Biochem. Med.* 25, 141–151. doi: 10.11613/bm.2015.015

Gingrich, K. J., Roberts, W. A., and Kass, R. S. (1995). Dependence of the GABAA receptor gating kinetics on the alpha−subunit isoform: implications for structure−function relations and synaptic transmission. *J. Physiol.* 489, 529–543. doi: 10.1113/jphysiol.1995.sp021070

Greer, P. L., Hanayama, R., Bloodgood, B. L., Mardinly, A. R., Lipton, D. M., Flavell, S. W., et al. (2010). The Angelman Syndrome protein Ube3A regulates synapse development by ubiquitinating arc. *Cell* 140, 704–716. doi: 10.1016/j.cell.2010.01.026

Gry, M., Rimini, R., Strömberg, S., Asplund, A., Pontén, F., Uhlén, M., et al. (2009). Correlations between RNA and protein expression profiles in 23 human cell lines. *Bmc Genom.* 10:365. doi: 10.1186/1471-2164-10-365

Guo, Y., Zhang, W., Chen, X., Fu, J., Cheng, W., Song, D., et al. (2017). Timing-dependent LTP and LTD in mouse primary visual cortex following different visual deprivation models. *PLoS One* 12:e0176603. doi: 10.1371/journal.pone.0176603

Hall, B. J., and Ghosh, A. (2008). Regulation of AMPA receptor recruitment at developing synapses. *Trends Neurosci.* 31, 82–89. doi: 10.1016/j.tins.2007.11.010

Hashimoto, T., Nguyen, Q. L., Rotaru, D., Keenan, T., Arion, D., Beneyto, M., et al. (2009). Protracted Developmental Trajectories of GABAA Receptor α1 and α2 Subunit Expression in Primate Prefrontal Cortex. *Biol. Psychiat.* 65, 1015–1023. doi: 10.1016/j.biopsych.2009.01.004

Hassani, M. (2015). *Package "subspace" - Interface to OpenSubspace*. Available online at: https://CRAN.R-project.org/package=subspace (accessed date 2015-October-12).

Hastie, T., Tibshirani, R., Botstein, D., and Brown, P. (2001). Supervised harvesting of expression trees. *Genome Biol.* 2:research0003.1. doi: 10.1186/gb-2001-2-1-research0003

Hastie, T., Tibshirani, R., Eisen, M. B., Alizadeh, A., Levy, R., Staudt, L., et al. (2000). "Gene shaving" as a method for identifying distinct sets of genes with similar expression patterns. *Genome Biol.* 1:research0003.1. doi: 10.1186/gb-2000-1-2-research0003

Heinen, K., Bosman, L. W. J., Spijker, S., Pelt, J., van, Smit, A. B., et al. (2004). Gabaa receptor maturation in relation to eye opening in the rat visual cortex. *Neuroscience* 124, 161–171. doi: 10.1016/j.neuroscience.2003.11.004

Hensch, T. K. (2004). Critical period regulation. *Annu. Rev. Neurosci.* 27, 549–579. doi: 10.1146/annurev.neuro.27.070203.144327

Hensch, T. K. (2005). Critical period plasticity in local cortical circuits. *Nat. Rev. Neurosci.* 6, 877–888. doi: 10.1038/nrn1787

Hensch, T. K., and Fagiolini, M. (2005). Excitatory-inhibitory balance and critical period plasticity in developing visual cortex. *Prog. Brain Res.* 147, 115–124. doi: 10.1016/s0079-6123(04)47009-5

Hensch, T. K., and Quinlan, E. M. (2018). Critical periods in amblyopia. *Vis. Neurosci.* 35:E014. doi: 10.1017/s0952523817000219

Huang, X., Stodieck, S. K., Goetze, B., Cui, L., Wong, M. H., Wenzel, C., et al. (2015). Progressive maturation of silent synapses governs the duration of a critical period. *Proc. Natl. Acad. Sci.* 112, E3131–E3140. doi: 10.1073/pnas.1506488112

Huttenlocher, P. R., and Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *J. Comp. Neurol.* 387, 167–178. doi: 10.1002/(sici)1096-9861(19971020)387:2<167::aid-cne1<3.0.co;2-z

Iwai, Y., Fagiolini, M., Obata, K., and Hensch, T. K. (2003). Rapid critical period induction by tonic inhibition in visual cortex. *J. Neurosci.* 23, 6695–6702. doi: 10.1523/jneurosci.23-17-06695.2003

Jiang, B., Huang, S., Pasquale, R. D., Millman, D., Song, L., Lee, H.-K., et al. (2010). The maturation of GABAergic transmission in visual cortex requires endocannabinoid-mediated LTD of inhibitory inputs during a critical period. *Neuron* 66, 248–259. doi: 10.1016/j.neuron.2010.03.021

Jones, D. G., Beston, B. R., and Murphy, K. M. (2007). Novel application of principal component analysis to understanding visual cortical development. *BMC Neurosci.* 8:188. doi: 10.1186/1471-2202-8-s2-p188

Kang, H. J., Kawasawa, Y. I., Cheng, F., Zhu, Y., Xu, X., Li, M., et al. (2011). Spatio-temporal transcriptome of the human brain. *Nature* 478, 483–489. doi: 10.1038/nature10523

Keech, R. V., and Kutschke, P. J. (1995). Upper age limit for the development of amblyopia. *J. Pediatr. Ophthal. Strab.* 32, 89–93.

Kleinschmidt, A., Bear, M. F., and Singer, W. (1987). Blockade of "NMDA" receptors disrupts experience-dependent plasticity of kitten striate cortex. *Science* 238, 355–358.

Kondo, Y. (2016). *RSKC: Robust Sparse K-Means*. Available online at: https://CRAN.R-project.org/package=RSKC (accessed date 2016-August-28).

Kondo, Y., Salibian-Barrera, M., and Zamar, R. (2016). RSKC: An RPackage for a Robust and Sparse K-Means Clustering Algorithm. *J. Stat. Softw.* 72, 1–26. doi: 10.18637/jss.v072.i05

Koopmans, F., Nierop, P., Andres-Alonso, M., Byrnes, A., Cijsouw, T., Coba, M., et al. (2019). SynGO: An Evidence-Based, Expert-Curated Knowledge Base for the Synapse. *Neuron* 103, 217.e–234.e. doi: 10.1016/j.neuron.2019.05.002

Kubota, S., and Kitajima, T. (2010). Possible role of cooperative action of NMDA receptor and GABA function in developmental plasticity. *J. Comput. Neurosci.* 28, 347–359. doi: 10.1007/s10827-010-0212-0

Lagos-Cabré, R., Burgos-Bravo, F., Avalos, A. M., and Leyton, L. (2020). Connexins in Astrocyte Migration. *Front. Pharmacol.* 10:1546. doi: 10.3389/fphar.2019.01546

Lambo, M. E., and Turrigiano, G. G. (2013). Synaptic and intrinsic homeostatic mechanisms cooperate to increase L2/3 pyramidal neuron excitability during a late phase of critical period plasticity. *J. Neurosci.* 33, 8810–8819. doi: 10.1523/jneurosci.4502-12.2013

Larsen, R. S., Rao, D., Manis, P. B., and Philpot, B. D. (2010). STDP in the Developing Sensory Neocortex. *Front. Synaptic Neurosci.* 2:9. doi: 10.3389/fnsyn.2010.00009

Law, A. J., Weickert, C. S., Webster, M. J., Herman, M. M., Kleinman, J. E., and Harrison, P. J. (2003). Expression of NMDA receptor NR1, NR2A and NR2B subunit mRNAs during development of the human hippocampal formation. *Eur. J. Neurosci.* 18, 1197–1205. doi: 10.1046/j.1460-9568.2003.02850.x

Lebenberg, J., Mangin, J.-F., Thirion, B., Poupon, C., Hertz-Pannier, L., Leroy, F., et al. (2018). Mapping the asynchrony of cortical maturation in the infant brain: A MRI multi-parametric clustering approach. *Neuroimage* 185, 641–653. doi: 10.1016/j.neuroimage.2018.07.022

Levelt, C. N., and Hübener, M. (2012). Critical-period plasticity in the visual cortex. *Annu. Rev. Neurosci.* 35, 309–330. doi: 10.1146/annurev-neuro-061010-113813

Lewis, T. L., and Maurer, D. (2005). Multiple sensitive periods in human visual development: Evidence from visually deprived children. *Dev. Psychobiol.* 46, 163–183. doi: 10.1002/dev.20055

Li, M., Santpere, G., Kawasawa, Y. I., Evgrafov, O. V., Gulden, F. O., Pochareddy, S., et al. (2018). Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* 362:eaat7615. doi: 10.1126/science.aat7615

Liao, Z., Patel, Y., Khairullah, A., Paker, N., and Paus, T. (2021). Pubertal Testosterone and the Structure of the Cerebral Cortex in Young Men. *Cereb Cortex* 1991:389. doi: 10.1093/cercor/bhaa389

Maechler, M. (2019). *Cluster: "Finding Groups in Data": Cluster Analysis Extended Rousseeuw et al.* Available online at: https://cran.r-project.org/web/packages/cluster/index.html.

McGee, A. W., Yang, Y., Fischer, Q. S., Daw, N. W., and Strittmatter, S. M. (2005). Experience-driven plasticity of visual cortex limited by myelin and Nogo receptor. *Science* 309, 2222–2226. doi: 10.1126/science.1114362

Murphy, K. M., Beston, B. R., Boley, P. M., and Jones, D. G. (2005). Development of human visual cortex: a balance between excitatory and inhibitory plasticity mechanisms. *Dev. Psychobiol.* 46, 209–221. doi: 10.1002/dev.20053

Norbom, L. B., Ferschmann, L., Parker, N., Agartz, I., Andreassen, O. A., Paus, T., et al. (2021). New insights into the dynamic development of the cerebral cortex in childhood and adolescence: Integrating macro- and microstructural MRI findings. *Prog. Neurobiol.* 204:102109. doi: 10.1016/j.pneurobio.2021.102109

Owsley, C. (2011). Aging and vision. *Vis. Res.* 51, 1610–1622. doi: 10.1016/j.visres.2010.10.020

Paus, T. (2016). Chapter 2 Population neuroscience. *Handb. Clin. Neurol.* 138, 17–37. doi: 10.1016/b978-0-12-802973-2.00002-1

Philpot, B. D., Cho, K. K. A., and Bear, M. F. (2007). Obligatory role of NR2A for metaplasticity in visual cortex. *Neuron* 53, 495–502. doi: 10.1016/j.neuron.2007.01.027

Philpot, B. D., Sekhar, A. K., Shouval, H. Z., and Bear, M. F. (2001). Visual experience and deprivation bidirectionally modify the composition and function of NMDA receptors in visual cortex. *Neuron* 29, 157–169. doi: 10.1016/s0896-6273(01)00187-8

Pinto, J. G. A., Hornby, K. R., Jones, D. G., and Murphy, K. M. (2010). Developmental changes in GABAergic mechanisms in human visual cortex across the lifespan. *Front. Cell Neurosci.* 4:16. doi: 10.3389/fncel.2010.00016

Pinto, J. G. A., Jones, D. G., Williams, C. K., and Murphy, K. M. (2015). Characterizing synaptic protein development in human visual cortex enables alignment of synaptic age with rat visual cortex. *Front. Neural. Circuit* 9:3. doi: 10.3389/fncir.2015.00003

Quinlan, E. M., Olstein, D. H., and Bear, M. F. (1999a). Bidirectional, experience-dependent regulation of N-methyl-D-aspartate receptor subunit composition in the rat visual cortex during postnatal development. *Proc. Natl. Acad. Sci.* 96, 12876–12880. doi: 10.1073/pnas.96.22.12876

Quinlan, E. M., Philpot, B. D., Huganir, R. L., and Bear, M. F. (1999b). Rapid, experience-dependent expression of synaptic NMDA receptors in visual cortex in vivo. *Nat. Neurosci.* 2, 352–357. doi: 10.1038/7263

Sanchez-Alonso, S., and Aslin, R. N. (2020). Predictive Modeling of Neurobehavioral State and Trait Variation Across Development. *Dev. Cogn. Neuros-neth* 45:100855. doi: 10.1016/j.dcn.2020.100855

Sato, M., and Stryker, M. P. (2010). Genomic imprinting of experience-dependent cortical plasticity by the ubiquitin ligase gene Ube3a. *Proc. Natl. Acad. Sci.* 107, 5611–5616. doi: 10.1073/pnas.1001281107

Sekula, M. (2020). *optCluster: Determine Optimal Clustering Algorithm and Number of Clusters*. Available online at: https://cran.r-project.org/web/packages/optCluster/index.html (accessed date 2020-April-01).

Sigelman, C. K., and Rider, E. A. (2017). *Life-Span Human Development. 9e ed*. Stamford CT: Nelson Education.

Siu, C. R., Balsor, J. L., Jones, D. G., and Murphy, K. M. (2015). Classic and Golli Myelin Basic Protein have distinct developmental trajectories in human visual cortex. *Front. Neurosci-switz* 9:138. doi: 10.3389/fnins.2015.00138

Siu, C. R., Beshara, S. P., Balsor, J. L., Mancini, S. J., and Murphy, K. M. (2018). *Use of Synaptoneurosome Samples to Study Development and Plasticity of Human Cortex in Synaptosomes*. New York, NY: Humana Press, 269–286. doi: 10.1007/978-1-4939-8739-9_15

Siu, C. R., Beshara, S. P., Jones, D. G., and Murphy, K. M. (2017). Development of Glutamatergic Proteins in Human Visual Cortex across the Lifespan. *J. Neurosci.* 37:6042. doi: 10.1523/jneurosci.2304-16.2017

Siu, C. R., and Murphy, K. M. (2018). The development of human visual cortex and clinical implications. *Eye Brain* 10, 25–36. doi: 10.2147/eb.s130893

Smith, G. B., Heynen, A. J., and Bear, M. F. (2009). Bidirectional synaptic mechanisms of ocular dominance plasticity in visual cortex. *Philosoph. Transac. R. Soc. B Biol. Sci.* 364, 357–367. doi: 10.1098/rstb.2008.0198

Smith, M. R., Readhead, B., Dudley, J. T., and Morishita, H. (2019). Critical period plasticity-related transcriptional aberrations in schizophrenia and bipolar disorder. *Schizoph. Res.* 207, 12–21. doi: 10.1016/j.schres.2018.10.021

Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., et al. (1999). Interpreting patterns of gene expression with self-organizing maps: Methods and application to hematopoietic differentiation. *Proc. Natl. Acad. Sci.* 96, 2907–2912. doi: 10.1073/pnas.96.6.2907

Tibshirani, R., Walther, G., and Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 63, 411–423. doi: 10.1111/1467-9868.00293

Turrigiano, G. G. (2017). The dialectic of Hebb and homeostasis. *Philosoph. Transact. R. Soc. B Biol. Sci.* 372, 20160258–20160257. doi: 10.1098/rstb.2016.0258

Vidal-Pineiro, D., Parker, N., Shin, J., French, L., Grydeland, H., Jackowski, A. P., et al. (2020). Cellular correlates of cortical thinning throughout the lifespan. *Sci Rep-uk* 10, 21803. doi: 10.1038/s41598-020-78471-3

Williams, K., Irwin, D. A., Jones, D. G., and Murphy, K. M. (2010). Dramatic Loss of Ube3A Expression during Aging of the Mammalian Cortex. *Front. Aging Neurosci.* 2:18. doi: 10.3389/fnagi.2010.00018

Witten, D. M., and Tibshirani, R. (2010). A framework for feature selection in clustering. *J. Am. Stat. Assoc.* 105, 713–726. doi: 10.1198/jasa.2010.tm09415

Witten, D. M., and Tibshirani, R. (2018). *sparcl: Perform Sparse Hierarchical Clustering and Sparse K-Means Clustering*. Available online at: https://CRAN.R-project.org/package=sparcl (accessed date 2018-October-24).

Yashiro, K., and Philpot, B. D. (2008). Regulation of NMDA receptor subunit expression and its implications for LTD, LTP, and metaplasticity.

Neuropharmacology 55, 1081–1094. doi: 10.1016/j.neuropharm.2008.07.046

Yashiro, K., Riday, T. T., Condon, K. H., Roberts, A. C., Bernardo, D. R., Prakash, R., et al. (2009). Ube3a is required for experience-dependent maturation of the neocortex. *Nat. Neurosci.* 12, 777–783. doi: 10.1038/nn.2327

Zhu, Y., Sousa, A. M. M., Gao, T., Skarica, M., Li, M., Santpere, G., et al. (2018). Spatiotemporal transcriptomic divergence across human and macaque brain development. *Science* 362:eaat8077. doi: 10.1126/science.aat8077

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Systems Neuroscience

ORIGINAL RESEARCH
published: 22 November 2021
doi: 10.3389/fnsys.2021.724805

Developmental Changes in Dynamic Functional Connectivity From Childhood Into Adolescence

Mónica López-Vicente[1,2], Oktay Agcaoglu[3], Laura Pérez-Crespo[4], Fernando Estévez-López[1,2], José María Heredia-Genestar[5], Rosa H. Mulder[1,2], John C. Flournoy[6], Anna C. K. van Duijvenvoorde[7,8], Berna Güroğlu[7,8], Tonya White[1,9], Vince Calhoun[3], Henning Tiemeier[1,10] and Ryan L. Muetzel[1]*

[1] Department of Child and Adolescent Psychiatry and Psychology, Erasmus MC University Medical Center, Rotterdam, Netherlands, [2] The Generation R Study Group, Erasmus MC University Medical Center, Rotterdam, Netherlands, [3] Tri-Institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), Georgia State University, Georgia Institute of Technology, Emory University, Atlanta, GA, United States, [4] Barcelona Institute for Global Health (ISGlobal), Barcelona, Spain, [5] Department of Molecular Genetics, Erasmus MC University Medical Center, Rotterdam, Netherlands, [6] Department of Psychology, Harvard University, Cambridge, MA, United States, [7] Leiden Institute for Brain and Cognition, Leiden University, Leiden, Netherlands, [8] Department of Developmental and Educational Psychology, Leiden University, Leiden, Netherlands, [9] Department of Radiology and Nuclear Medicine, Erasmus MC University Medical Center, Rotterdam, Netherlands, [10] Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, MA, United States

The longitudinal study of typical neurodevelopment is key for understanding deviations due to specific factors, such as psychopathology. However, research utilizing repeated measurements remains scarce. Resting-state functional magnetic resonance imaging (MRI) studies have traditionally examined connectivity as 'static' during the measurement period. In contrast, dynamic approaches offer a more comprehensive representation of functional connectivity by allowing for different connectivity configurations (time varying connectivity) throughout the scanning session. Our objective was to characterize the longitudinal developmental changes in dynamic functional connectivity in a population-based pediatric sample. Resting-state MRI data were acquired at the ages of 10 (range 8-to-12, n = 3,327) and 14 (range 13-to-15, n = 2,404) years old using a single, study-dedicated 3 Tesla scanner. A fully-automated spatially constrained group-independent component analysis (ICA) was applied to decompose multi-subject resting-state data into functionally homogeneous regions. Dynamic functional network connectivity (FNC) between all ICA time courses were computed using a tapered sliding window approach. We used a k-means algorithm to cluster the resulting dynamic FNC windows from each scan session into five dynamic states. We examined age and sex associations using linear mixed-effects models. First, independent from the dynamic states, we found a general increase in the temporal variability of the connections between intrinsic connectivity networks with increasing age. Second, when examining the clusters of dynamic FNC windows, we observed that the time spent in less modularized states, with low intra- and inter-network connectivity, decreased with age. Third, the number of transitions between states also decreased with age. Finally, compared to boys, girls showed a more mature pattern of dynamic brain connectivity, indicated by more time spent in a highly modularized state, less time spent in specific states that are frequently observed at a younger age, and a lower number of transitions between

OPEN ACCESS

Edited by:
Tomáš Paus,
Université de Montréal, Canada

Reviewed by:
Adriana Galván,
University of California, Los Angeles, United States
Ting Xu,
Child Mind Institute, United States

***Correspondence:**
Henning Tiemeier
tiemeier@hsph.harvard.edu

Received: *14 June 2021*
Accepted: *26 October 2021*
Published: *22 November 2021*

Citation:
López-Vicente M, Agcaoglu O, Pérez-Crespo L, Estévez-López F, Heredia-Genestar JM, Mulder RH, Flournoy JC, van Duijvenvoorde ACK, Güroğlu B, White T, Calhoun V, Tiemeier H and Muetzel RL (2021) Developmental Changes in Dynamic Functional Connectivity From Childhood Into Adolescence. Front. Syst. Neurosci. 15:724805. doi: 10.3389/fnsys.2021.724805

states. This longitudinal population-based study demonstrates age-related maturation in dynamic intrinsic neural activity from childhood into adolescence and offers a meaningful baseline for comparison with deviations from typical development. Given that several behavioral and cognitive processes also show marked changes through childhood and adolescence, dynamic functional connectivity should also be explored as a potential neurobiological determinant of such changes.

Keywords: brain development, fMRI, longitudinal, resting state – fMRI, linear mixed effect model

INTRODUCTION

Neurodevelopment from childhood into adolescence represents a pivotal period, marked by several cognitive, social, and behavioral milestones, and is also beset with the emergence of many forms of psychopathology (Nelson et al., 2005; Paus et al., 2008; Luciana, 2013). Typical neurodevelopment provides a baseline framework for understanding how deviations in brain development are associated with mental and neurological illness, and it has been characterized *in vivo* using structural and functional magnetic resonance imaging (MRI) for over two decades (Giedd et al., 1999; Rubia et al., 2000; Luna et al., 2001; Gogtay et al., 2004; Lebel et al., 2008). More recently, resting-state functional MRI (rs-fMRI) has been used to study brain development. This brain imaging modality is used to measure intrinsic functional brain connectivity, or the spontaneous, correlated activations among brain networks (Biswal et al., 1995; Cole et al., 2010). The connectivity patterns of these networks exhibit high reproducibility between individuals, representing a reliable indicator of brain development (Allen et al., 2011). Despite widespread application, the vast majority of neurodevelopmental studies using rs-fMRI have been cross-sectional in design, lacking crucial insights from repeated measures (Kraemer et al., 2000).

Traditional rs-fMRI analysis approaches focus on the average functional connectivity across the scanning session, effectively assuming the connectivity is 'static'. Studies of static brain connectivity have observed intra- and inter-network connectivity associations with age, and a number of networks show abnormal connectivity patterns in the presence of psychiatric disorders (Di Martino et al., 2014; Muetzel et al., 2016; Bos et al., 2017). Certain resting-state networks, such as the precuneus and the lateral frontal, increase their connectivity during brain development while others, such as the right frontoparietal and sensory networks, decrease with age (Muetzel et al., 2016). While static brain connectivity studies provide information about topological organization of functional brain networks during development, changes in connectivity throughout the scanning session are not captured by using this approach (Delamillieure et al., 2010; Calhoun et al., 2014).

Dynamic brain connectivity is a novel functional MRI analysis technique that allows connectivity between brain areas to vary over time, relaxing the stationarity assumption (Allen et al., 2014; Calhoun et al., 2014). Though several novel methods exist to estimate dynamic connectivity, one popular framework identifies different connectivity configurations, or states, across the scanning session and offers summary metrics, such as the time spent in each of these states. Although the general structure and topology of functional connectivity states are stable across age, there are age-related changes in the frequency of certain states and the time spent in each of them (Hutchison and Morton, 2015; Marusak et al., 2017). In the Generation R Study (n = 774, 6–10 years old), Rashid et al. (2018) found that older children spent more time in a state that showed a modular organization of functional connectivity in distinguished networks (Rubinov and Sporns, 2010), named 'globally modularized dynamic state'. In this state, the nodes comprising a network were positively interconnected among them and those of different networks were negatively correlated. Contrarily, younger children spent more time in a globally disconnected state (Rashid et al., 2018). In addition, girls spent more time in a default mode modularized state compared to boys, which could indicate an earlier maturation of functional connectivity (Rashid et al., 2018). In the PING Study (n = 421, 3–21 years old), Faghiri et al. (2018) showed that age was negatively correlated with the time spent in states with strong connectivity between cognitive control and default mode domains, while older participants stayed longer in states showing positive intra-network connectivity within the default mode domain. Although the number of transitions between different states has not been associated with age in cross-sectional studies, some of these studies observed that the overall connectivity between intrinsic networks becomes more variable (higher standard deviation, SD) across the scanning session from childhood to adulthood (Hutchison and Morton, 2015; Qin et al., 2015; Marusak et al., 2017). For instance, Marusak et al. (2017) reported positive age associations with the variability of functional connectivity between core neurocognitive networks, which may afford greater cognitive and behavioral flexibility.

Currently, the literature examining associations between age and dynamic brain connectivity indicators in children has been comprised exclusively of cross-sectional studies. While these studies have established the fundamental basis for our understanding of age- and sex-related differences in functional brain connectivity, cross-sectional neurodevelopmental research provides limited information and it does not take into account inter-individual variability (Kraemer et al., 2000). Therefore, longitudinal studies are needed to explore individual growth changes, which is of key importance to understand the deviations in neurodevelopment after various exposures (e.g., early life adversities) or in psychopathology (Kraemer et al., 2000; Crone and Elzinga, 2015). Therefore, our objective was to characterize the longitudinal developmental changes in dynamic functional connectivity from childhood into adolescence in a

large, population-based sample, as a follow-up of the cross-sectional findings observed in the Generation R Study by Rashid et al. (2018) at a younger age. We also aimed to understand whether maturation of dynamic functional connectivity was distinct in boys and girls, as sex differences have been observed using rs-fMRI (Satterthwaite et al., 2015) as well as other MRI modalities (Lenroot and Giedd, 2006; Perrin et al., 2009; López-Vicente et al., 2020). Our main focus was on global summary metrics of dynamic connectivity, specifically those related to time spent in a given connectivity configuration and the number of transitions between different connectivity configurations. Since previous cross-sectional studies suggested that the overall connectivity between intrinsic networks becomes more variable during development, we additionally tested the longitudinal age associations with the temporal variability (SD) of functional connections across the scanning session. We hypothesized that, over time, children would show more variable connections and they would spend more time in modularized states. This is in line with research indicating that brain development is characterized by the increase of "integration" of functional networks (Fair et al., 2007) and also with the existing dynamic connectivity literature (Faghiri et al., 2018; Rashid et al., 2018). Given differential developmental patterns previously reported with various neuroimaging modalities, we hypothesized that girls would show slightly faster development than boys, showing more mature dynamic connectivity patterns.

MATERIALS AND METHODS

Participants

The current study is part of the Generation R Study, a population-based birth cohort in Rotterdam, the Netherlands (Kooijman et al., 2016). Data in this study includes rs-fMRI data acquired at the age-10 visit (mean age 10 years, range 8-to-12, n = 3,327) and the age-14 visit (mean age 14 years, range 13-to-15, n = 2,404). Data collection was carried out between March 2013 and November 2015 for the age-10 visit (White et al., 2018) and between October 2016 and January 2020 for the age-14 visit. A flow-chart outlining data inclusion/exclusion for the study can be found in **Figure 1**. Few data were excluded due to the presence of prominent incidental findings (n_{age10} = 19 and n_{age14} = 24). Due to excessive motion, 698 datasets (21%) were excluded from the age-10 visit and 168 datasets (7%) were excluded from the age-14 visit (see Section "Image Quality Assurance" for details). Data were also excluded due to poor registration (spatial normalization) (0.7% from the age-10 visit and 0.3% from the age-14 visit). After this filtering, the total number of datasets available for analysis in this study was 2,586 at the age-10 visit and 2,204 at the age-14 visit. 1,031 participants had data in both visits (**Figure 1**). **Supplementary Figure 1** shows the age at scan for all the individuals with repeated measures. All parents provided written informed consent and children provided assent (younger than 12 years) or consent (12 years or older). All study procedures were approved by the local medical ethics committee of the Erasmus MC University Medical Center.

Magnetic Resonance Imaging Data Acquisition

Magnetic resonance images were acquired on a study-dedicated 3 Tesla GE Discovery MR750w MRI System (General Electric, Milwaukee, WI, United States) scanner using an 8-channel head coil. No hardware upgrades or major software upgrades have taken place since the study began in 2012 in order to keep the system stable for longitudinal research.

After a brief mock scanning session to acclimate the participants to the MRI environment, structural T_1-weighted images were obtained using a 3D coronal inversion recovery fast spoiled gradient recalled (IR-FSPGR, BRAVO) sequence using ARC acceleration [T_R = 8.77 ms, T_E = 3.4 ms, T_I = 600 ms, flip angle = 10°, matrix = 220 × 220, field of view (FOV) = 220 mm × 220 mm, slice thickness = 1 mm]. 200 volumes of rs-fMRI data were acquired using an interleaved axial echo planar imaging sequence with the following parameters: T_R = 1,760 ms, T_E = 30 ms, flip angle = 85°, matrix = 64 × 64, FOV = 230 mm × 230 mm, slice thickness = 4 mm (White et al., 2018). The total duration of the resting-state scan was 5 min 52 s. Children were instructed to stay awake and keep their eyes closed.

Image Preprocessing

Data were first converted from DICOM to Nifti format using dcm2niix (Li et al., 2016). Data were subsequently preprocessed through the FMRIPrep package (version 20.1.1 singularity container) (Esteban et al., 2019). Briefly, this included volume realignment for translation and rotation motion, slice-timing correction, and inter-subject registration.

Spatial normalization to the ICBM 152 Non-linear Asymmetrical template version 2009c (Fonov et al., 2009) was performed through non-linear registration with the antsRegistration tool of ANTs v2.1.0 (Avants et al., 2008), using brain-extracted versions of both T1w volume and template. The resulting functional data were resampled to 3 mm × 3 mm × 3 mm isotropic voxels.

From the volume realignment, we obtained the time series corresponding to the first temporal derivatives of the six base motion parameters (3 translations and 3 rotations), together with their quadratic terms, resulting in the total 24 head motion parameters (6 base motion parameters + 6 temporal derivatives + 12 quadratic terms) to be used as confound regressors (see below) (Satterthwaite et al., 2013).

Group-Independent Component Analysis and Dynamic Functional Network Connectivity Analysis

Group-independent component analysis and dynamic functional network connectivity analyses were performed using the Group ICA Of fMRI Toolbox (GIFT) software[1] (GroupICAT v4.0b) (Calhoun et al., 2001; Calhoun and Adalí, 2012) in MATLAB R2020a.

[1]https://trendscenter.org/software/gift

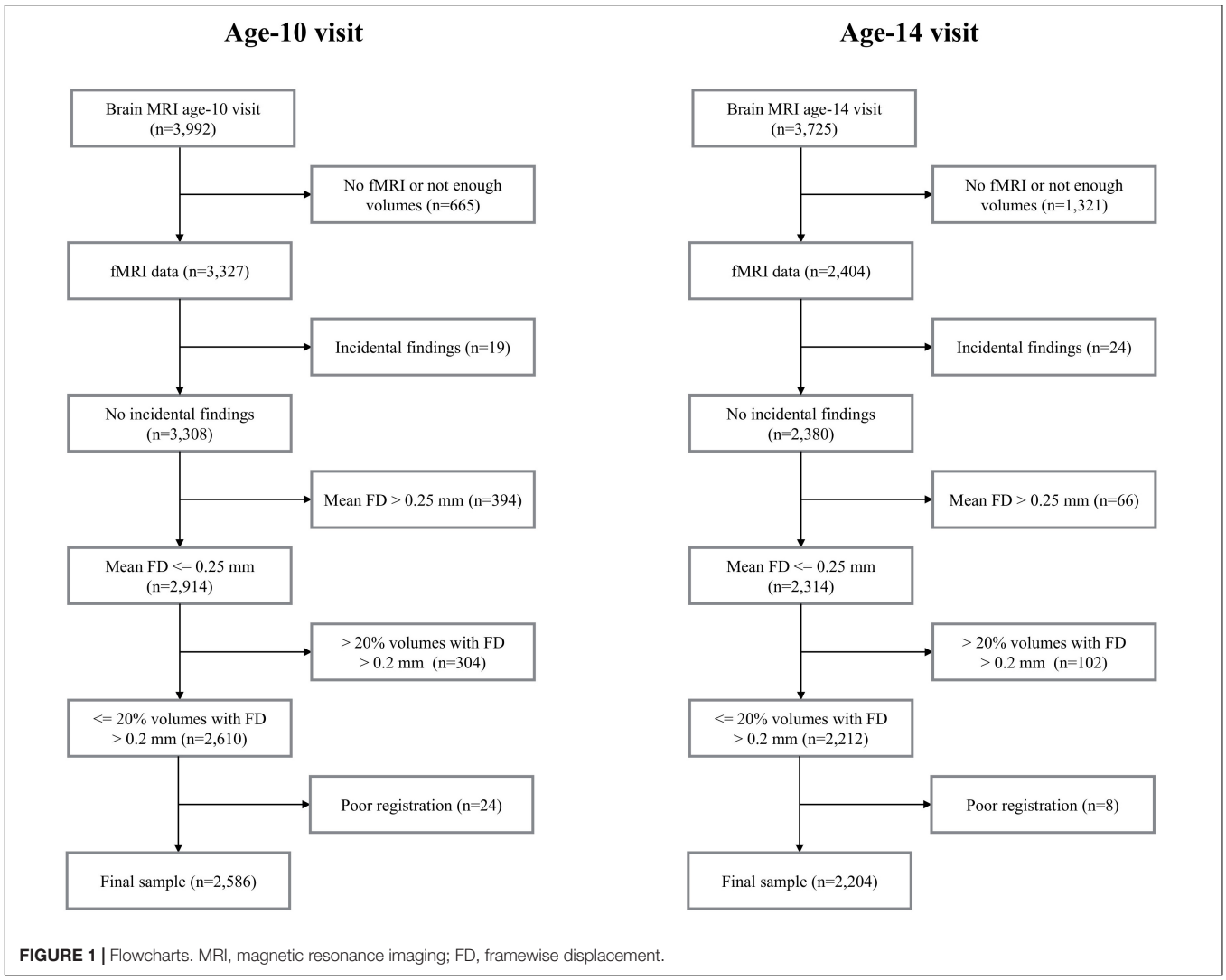

FIGURE 1 | Flowcharts. MRI, magnetic resonance imaging; FD, framewise displacement.

Group-Independent Component Analysis

Prior to analysis, the first 4 volumes of each subject were excluded to ensure magnetic stabilization. Resting-state data was decomposed into functionally homogeneous regions applying a spatially constrained group-independent component analysis (scICA) via the multi-objective optimization ICA with reference algorithm (Du and Fan, 2013). The scICA method is a fully automated approach which uses aggregate component maps from previous group ICA analysis as reference to estimate subject specific independent components. This technique has been previously used on adult studies (Salman et al., 2019; Du et al., 2020). Here we used 51 reference components derived from the Dev-CoG developmental imaging study, and grouped them into seven networks: subcortical (SC), auditory (AUD), sensorimotor (SM), visual (VIS), default-mode (DMN), cognitive control (CC), and cerebellar (CB) (**Supplementary Figure 2** and **Supplementary Table 1**) (Agcaoglu et al., 2019).

The subject specific time courses corresponding to the components were detrended, despiked, and the 24 motion parameters were regressed out. As correlation among brain networks is primarily driven by low frequency fluctuations (Cordes et al., 2001), time courses were also filtered using a fifth-order Butterworth low-pass filter with a high frequency cut-off of 0.15 Hz.

Dynamic Functional Network Connectivity Analysis

Dynamic functional network connectivity (FNC) between all independent components time courses was computed using a tapered sliding window approach. This method provides multiple correlation matrices (one per window per dataset) that correspond to different temporal portions of data. We used a window size of 25 TR (44 s) in steps of 1 TR and the alpha parameter of the Gaussian sliding window was 3 TRs (Allen et al., 2014; Qin et al., 2015), which yielded 171 FNC windows per dataset. We estimated covariance from regularized inverse covariance matrices (Smith et al., 2011) using a graphical LASSO framework (Friedman et al., 2008) as estimation of covariance matrices of short time series can be noisy. Also, we imposed an additional L1 norm constraint on the inverse covariance matrix to enforce sparsity. The regularization parameter was optimized for

TABLE 1 | Sample characteristics.

	Single measurement		Repeated measurements
	Age-10 visit	**Age-14 visit**	**Age-10 and -14 visits**
N	1,555	1,173	1,031
Age (years, mean ± SD)	10.13 ± 0.58	14.18 ± 0.67	10.15 ± 0.62; 13.85 ± 0.51
Sex (n, %)			
Boys	787 (50.61)	546 (46.55)	466 (45.2)
Girls	768 (49.39)	627 (53.45)	565 (54.8)
Ethnicity (n, %)			
Dutch	948 (60.96)	671 (57.2)	642 (62.27)
Other western	138 (8.87)	107 (9.12)	94 (9.12)
Non-western	440 (28.3)	363 (30.95)	274 (26.58)
Maternal education (n, %)			
None or primary	38 (2.44)	30 (2.56)	24 (2.33)
Secondary	451 (29)	387 (32.99)	317 (30.75)
Higher	835 (53.7)	565 (48.17)	589 (57.13)
Household income (n, %)			
Very low (<1,200€/month)	83 (5.34)	52 (4.43)	40 (3.88)
Low (1,200€–2,400€/month)	235 (15.11)	215 (18.33)	168 (16.29)
Modal and higher (>2,400€/month)	950 (61.09)	666 (56.78)	675 (65.47)

SD, standard deviation. The percentages of missing values in subjects with single measurement at age-10 visit were 2% for ethnicity, 15% for maternal education and 18% for household income. The percentages of missing values in subjects with single measurement at age-14 visit were 3% for ethnicity, 16% for maternal education and 20% for household income. The percentages of missing values in subjects with repeated measurements were 2% for ethnicity, 10% for maternal education and 14% for household income.

each subject/visit by evaluating the log-likelihood of unseen data in a cross-validation framework, that is, splitting time courses into training and testing sets. Finally, to stabilize variance, the dynamic FNC values were Fisher-*Z* transformed.

Clustering

Using data from both visits, we computed k-means clustering on the resulting 171 dynamic FNC windows of 44 s from each scan session in order to identify patterns of connectivity that reoccur over time (within the scan session) and across subjects and visits (between the scan sessions). The number of clusters, or states, was set to five to match previous studies (Faghiri et al., 2018; Agcaoglu et al., 2020). We used the correlation distance function and the clustering algorithm was repeated 500 times to increase the chances of escaping local minima, with random initialization of centroid positions. We determined the modularity of the dynamic states qualitatively. First, a state was described as fully modularized when a clear modular organization, thus positive intra-network connectivity and negative inter-network connectivity, was observed. Next, if a state was not fully modularized, but presented sub-modules within networks with different connectivity configurations, we defined it as being partially modularized. Lastly, if a state did not possess any or very little characteristics of being modularized, we labeled it as a non-modularized state.

Outcome Measures

For each individual and visit, we calculated three different outcomes. First, the SD of the functional connections between the 51 components as a measure of temporal variability. Second, the mean dwell time (MDT) in each dynamic state. This variable was obtained by first identifying every change between states, calculating the number of windows in each state and computing the average time a participant remained in the specific states [for a more detailed explanation, see Rashid et al. (2018)]. Third, the number of transitions between states.

Image Quality Assurance

Scans with excessive motion defined as having a mean framewise displacement (FD) higher than 0.25 mm or having more than 20% of the volumes with a FD higher than 0.2 mm, were excluded (Parkes et al., 2018). Image co-registration was visually inspected for accuracy by merging all co-registered images into a single 4D Nifti image and scrolling through the images. Scans were also screened for major artifacts (e.g., from dental retainers, or other scanner-related artifacts) as well as whole-brain coverage (e.g., missing from field of view). All the scans flagged as being of poor quality for the above-mentioned reasons were excluded.

Sample Characteristics

Descriptive characteristics of the participants are presented with means and standard deviations or proportions. Child sex and date of birth were determined from medical records obtained at birth. Child ethnicity was defined based on the country of birth of the parents and was coded into three categories (Dutch, non-western, and other western). Maternal education level and household income, proxies of socioeconomic status, were assessed by questionnaire during pregnancy.

Statistical Analyses

Statistical analyses were conducted using the R Statistical Software (version 3.6.0).

We compared the summary metrics (MDT in each state and number of transitions between states) between age-10 and age-14 visits and between boys and girls in each visit using Wilcoxon tests. For the visit comparison we used Wilcoxon signed rank test, while for the sex comparisons, we used Wilcoxon rank sum test.

Age was centered to the mean age of the sample at the age-10 visit. The distributions of all the dependent variables were visually inspected using histograms. Since MDT outcomes were right-skewed, we applied a Box–Cox transformation (Box and Cox, 1964) using the 'bestNormalize' package (version 1.6.1) (Peterson and Cavanaugh, 2020) to obtain a more homogeneous range of values.

The age associations with the temporal variability of connectivity (SD) between components, and the summary metrics (transformed MDT in each state and number of transitions between states) were estimated using linear mixed-effects models, implemented in the 'nlme' package (version 3.1-139) (Pinheiro et al., 2019), including age and sex as fixed effects and subject as random effect, allowing the intercept to vary randomly across subjects. For the summary metrics, we also tested the quadratic term of age in order to capture the possible non-linearity in the growth changes, and we added an interaction term of age-by-sex into the regression models to detect potential differential age associations in boys and girls, following a step-wise approach. We performed the likelihood ratio (LR) test for model comparisons using ML estimation. Stratified analyses by sex were performed when we observed statistically significant interactions. Given the important role of socioeconomic status in the brain development (Brito and Noble, 2014), we additionally included maternal education in the models as a precision covariate. The models were performed separately for each outcome. A false discovery rate (FDR) was applied to control for Type-I error. Associations with $p_{\text{corrected}} < 0.05$ were considered significant.

The associations between age and the summary metrics were graphically represented in the original scale using the 'ggplot2' package (version 3.3.2) (Wickham, 2016). To estimate the variation of the values in the population, we applied the bootstrap technique (Efron, 1979), using 2,000 resamples with replacement.

RESULTS

Sample Characteristics

The sample characteristics are shown separately for participants with a single measurement (age-10 or age-14 visit) and those with repeated measurements in **Table 1**. The mean age and variation at the age-10 visit were very similar between those with and without repeated measurements, as well as at the age-14 visit. The mean duration between visits was 4 years (range 1–6 years). Although the proportion of boys and girls was balanced in the participants with data only at the age-10 visit, there were more girls than boys with single measurement at age-14 visit (53% girls vs. 47% boys) and with data at both visits (55% girls vs. 45% boys). Around 60% of the participants were of Dutch origin and between 26 and 31% were of non-western national origin, with small differences between single/repeated measurement groups. The groups also differed slightly in terms of maternal education and household income, although the relative proportions were constant between them. The participants with data only at the age-14 visit had the highest proportion of low (3%) or secondary (33%) maternal education and lower household income (<2,400€ per month, 23%).

Temporal Variability in Functional Connections Within Scan Session

Figure 2 shows the average variability (SD) in the correlations between the time courses of the 51 components over the measurement period across participants and visits. In general, the SDs were between 0.20 and 0.25. The smallest variability was observed within the VIS network, indicating more stable connections over the scan session. The most variable connections were observed between SM and VIS networks, CC and VIS networks, and within the CC network.

Dynamic States

The *k*-means clustering method allowed us to identify five dynamic states, or patterns of connectivity that reoccurred over time (within the scan session) and across subjects and visits (between the scan sessions). We obtained three modularized states with components showing intra- and inter-network connectivity (states 1, 2, and 3) and two non- or only partially modularized states (states 4 and 5) (**Figure 3**). In state-1 (15% of occurrences), the SC and CB networks showed positive intra-network connectivity and negative inter-network connectivity, mainly with the sensory networks (AUD, SM, and VIS). Thus, the components that comprise the SC and CB networks were positively correlated within themselves, and they were negatively correlated with the components of the sensory networks. The SM and the VIS networks also showed strong positive intra-network connectivity. The frequency of occurrence of state-1 increased with scan progression (i.e., occurred more as the scan went on, particularly, around 7% more windows of time were part of this state at the end of the session as compared to the beginning) (**Figure 4**). In state-2 (22% of occurrences), SM and DMN networks showed positive intra-network connectivity and negative inter-network connectivity between them. Other components, such as CC components, showed opposite connectivity patterns with SM and DMN. For instance, frontal CC components were positively correlated with DMN and negatively correlated with SM components, and posterior CC components showed the opposite pattern. The frequency of state-2 decreased by 5% with scan progression (**Figure 4**). State-3 (20% of occurrences) was characterized by positive correlations within the DMN and negative correlations between this network and the other networks, except some CC components from the frontal lobe. The SM and the VIS networks also showed strong positive intra-network connectivity in this state. State-3 shared some traits with state-2. However, the negative inter-network connectivity between DMN and the other

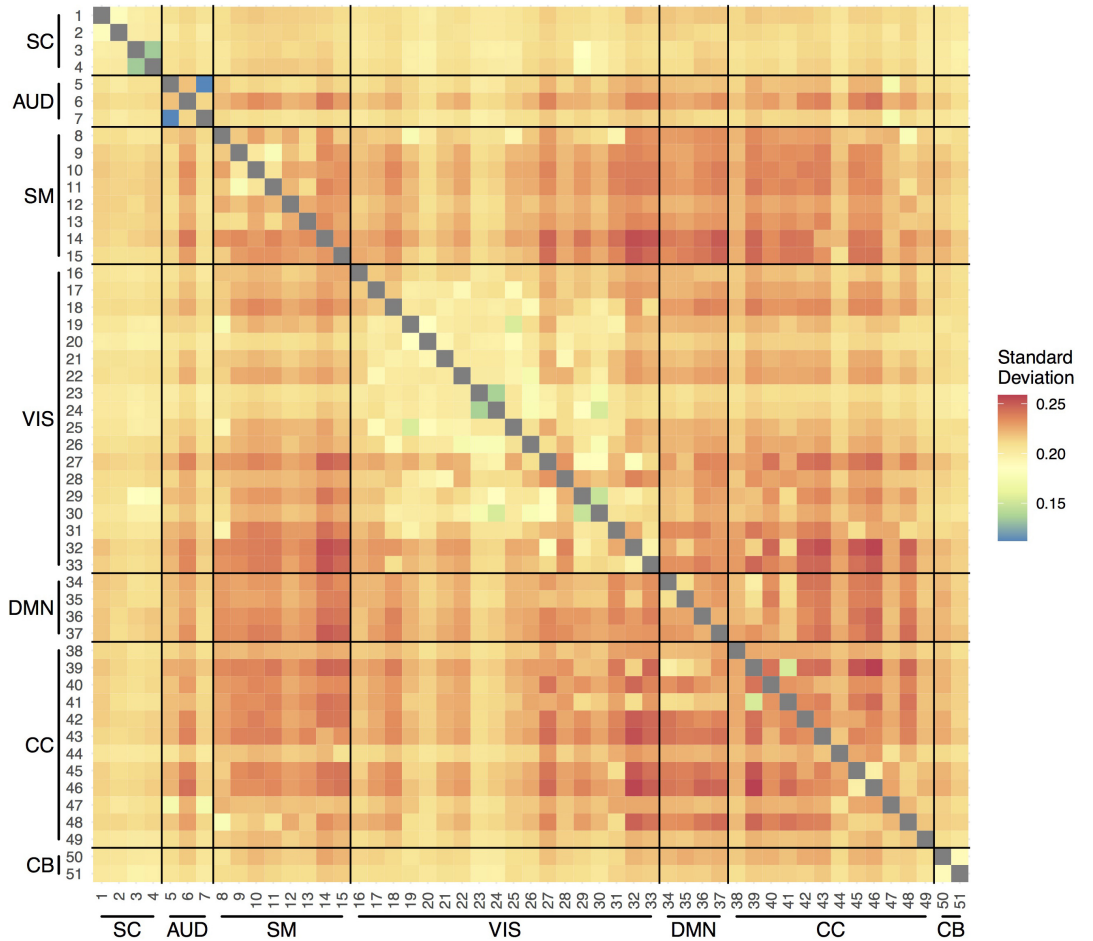

FIGURE 2 | Temporal variability in pairwise functional connections between the 51 components within scan session. SC, subcortical network; AUD, auditory network; SM, sensorimotor network; VIS, visual network; DMN, default-mode network; CC, cognitive control network; CB, cerebellar network.

networks was more "global," although weaker, in state-3 than in state-2, in which the inter-network connectivity was focused on specific networks. The frequency of state-3 also decreased by 3% of windows of time over the scan session (**Figure 4**).

State-4 (18% of occurrences) was non-modularized, thus a clear modular organization of functional connectivity in distinguished networks was absent in this state. The putamen (SC, 1 and 2), the middle temporal gyrus (CC, 38), and the cerebellum were negatively correlated with all the other components. The opposite was observed in the postcentral gyrus component (SM, 8), which was positively correlated to several components. The frequency of state-4 increased by 4% with scan progression (**Figure 4**).

Finally, state-5 (24% of occurrences) was partially modularized, presenting sub-modules within networks with different connectivity configurations. For instance, regarding the SC network, the putamen (SC, 1 and 2) was negatively correlated with visual components, while the thalamus (SC, 3 and 4) showed positive correlations. The postcentral gyrus component (SM, 8) was positively correlated with visual components, and the rest of the SM components were negatively correlated with those components. As in state-3, the DMN was positively connected within network and with frontal CC components. The frequency of state-5 showed a general mild decreasing trend over the scan session (**Figure 4**).

Longitudinal Changes in the Temporal Variability

We observed increases in overall temporal variability in functional connections between components across age (**Figure 5**). The SD, which ranged between 0.20 and 0.25, increased on average a 1% (coefficient = 0.0025) per year. Some connections showed less variability at older ages, such as the Insular Cortex component (CC, 44) connections with components from other networks.

Longitudinal Changes in Dynamic States

The distributions of the summary metrics (MDT and number of transitions) by visit and sex are depicted in the original scale in **Figure 6**, and the individual observations are shown

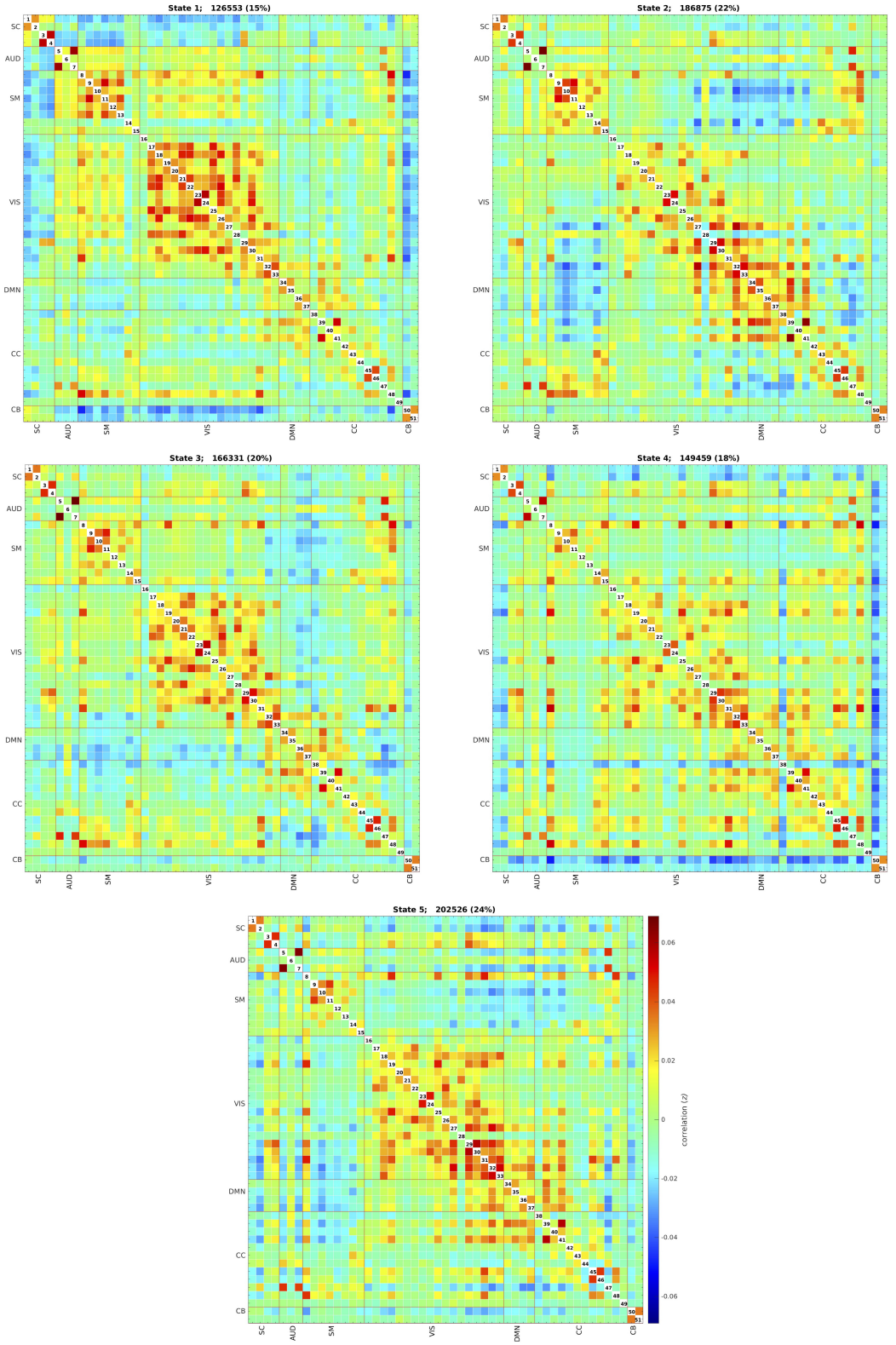

FIGURE 3 | Five dynamic states after clustering across all individuals and visits. Each state captures a particular connectivity 'configuration' a participant may display over the course of the MRI scan. The total number and percentage of occurrences (windows of time) is listed above each state. SC, subcortical network; AUD, auditory network; SM, sensorimotor network; VIS, visual network; DMN, default-mode network; CC, cognitive control network; CB, cerebellar network.

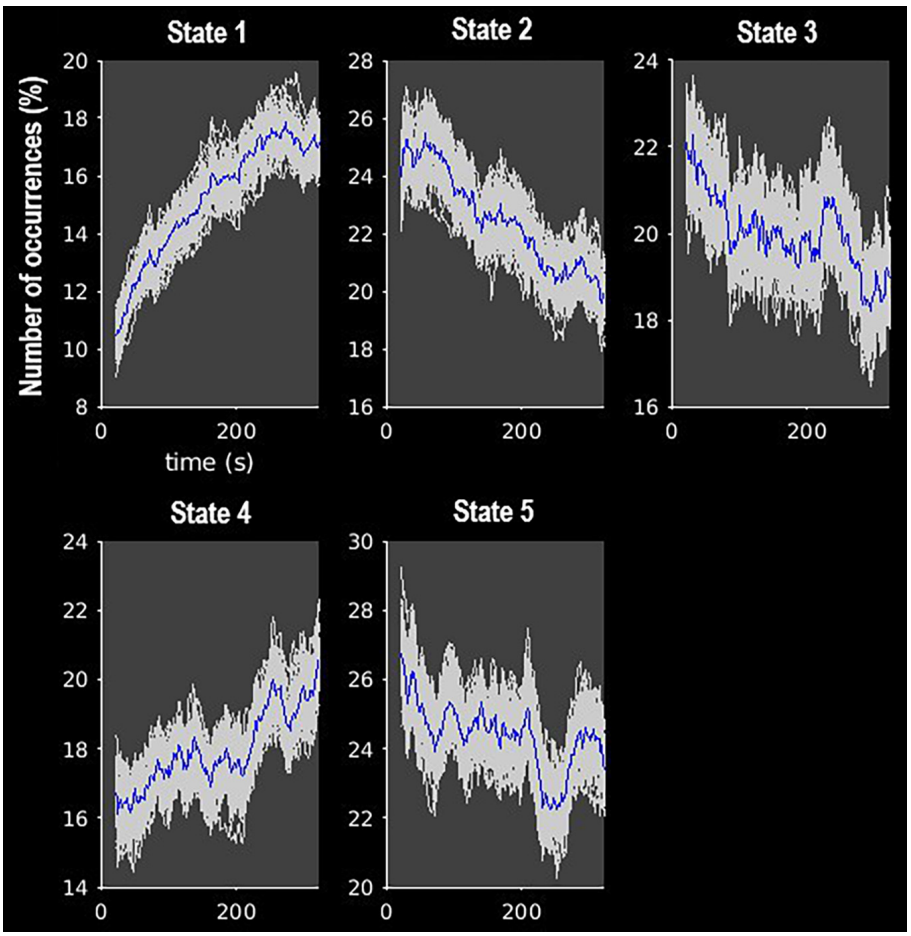

FIGURE 4 | Frequency of occurrence of each state over the course of the MRI scan. Gray lines indicate the estimated occurrence profiles of each state for 100 bootstrap resamples with replacement.

in **Supplementary Figure 3**. We observed differences between age-10 and age-14 visits in all the outcomes except state-2 MDT. Boys and girls also showed differences in the MDT of states 1, 3, and 5 (age-10 visit), states 2 and 4 (both visits), and the number of transitions between states (age-14 visit). As outlined in the Section "Statistical Analyses", the linear mixed-effects models were performed using transformed MDT outcomes (Box–Cox transform). **Table 2** shows the age and sex associations with the MDT of each state and the number of transitions between states. Only the linear term of age was included in the models because the LR test indicated that the model fit was not significantly improved with the quadratic term addition. Overall, the MDT in state-1 increased with age, while the MDT in states 3, 4, and 5 decreased with age. The number of transitions between states decreased over time. Girls spent more time in state-2, less time in states 3 and 4, and showed fewer transitions between states compared to boys.

From baseline to follow-up, increases were observed for the time spent (MDT) in state-1, which is characterized by negative inter-network connectivity between subcortical and sensorimotor networks. Thus, children spent more time in state-1 as they grew older. **Figure 7**, depicts the predicted number of windows spent in state-1 increased slightly more in absolute terms at older ages than at younger ages. No sex differences were observed in MDT for state-1, and adding the age-by-sex interaction term did not improve the model (LR test $p = 0.491$). Regarding state-2 (the default-mode/sensorimotor modularized), we found differences by sex, with girls spending around 1% more time (2 windows of time) than boys in this state across the whole age range (**Figure 7**). However, no significant interaction was observed (LR test $p = 0.051$).

We observed decreases in MDT for state-3 (the default-mode network modularized state), state-4 (the non-modularized state), and state-5 (the partially modularized state) associated with age. Children spent around 0.6% less time (1 window) per year in state-3. In addition, girls spent around 0.6% less time (1 window) in this state than boys across the whole age period, with no age-by-sex interaction (LR test $p = 0.120$) (**Figure 7**). Regarding state-4, we found an interaction between age and sex (LR test $p = 0.006$) (**Table 3** and

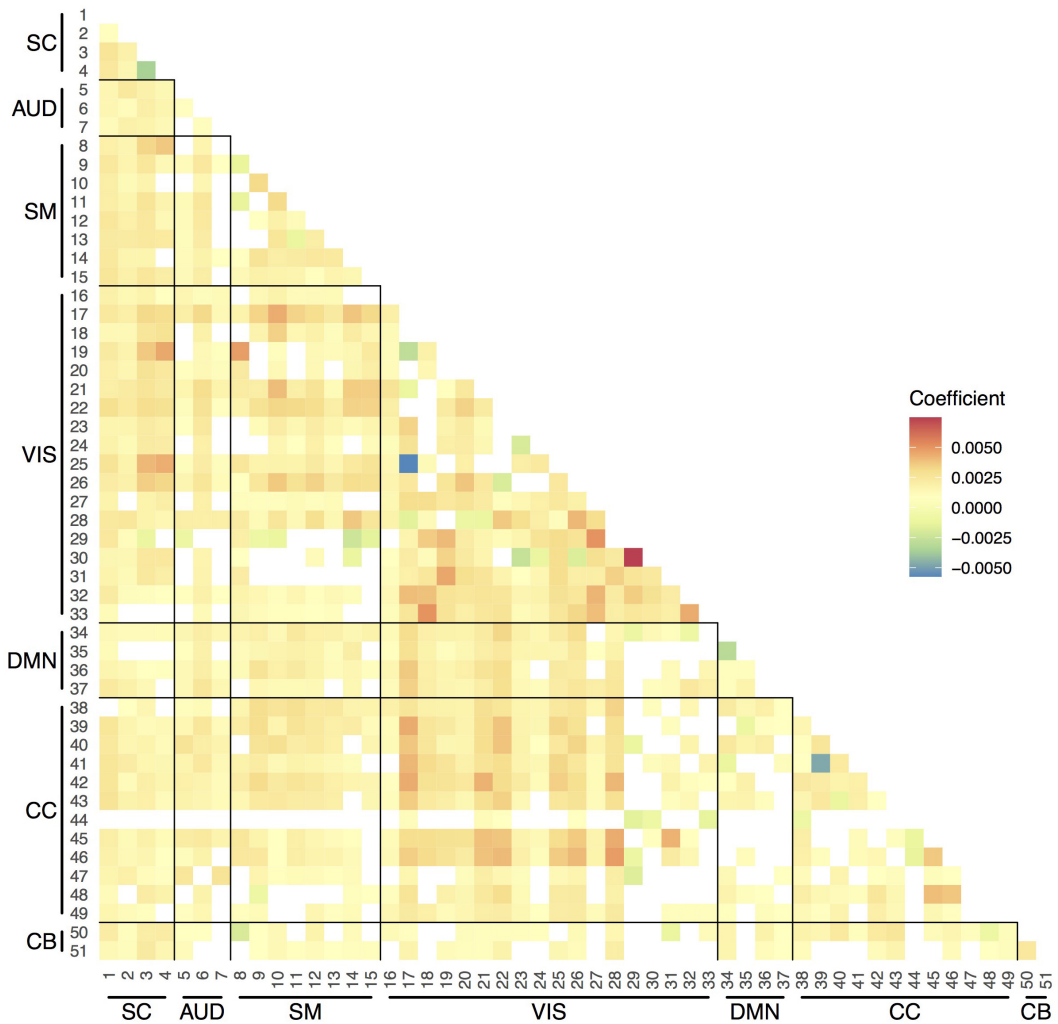

FIGURE 5 | Strength of the associations between age and variability in pairwise functional connections between components. Linear mixed-effects models adjusted for sex (random effect: subject). Only the results that survived the false discovery rate (FDR) multiple comparison correction threshold of pFDR = 0.05 are shown here. SC, subcortical network; AUD, auditory network; SM, sensorimotor network; VIS, visual network; DMN, default-mode network; CC, cognitive control network; CB, cerebellar network.

Figure 7). The stratified analyses showed that the negative association between age and MDT in state-4 was stronger in girls (**Table 4**). Girls spent 0.3% less time (half window) per year in this state. In boys, the slope decreased more slowly with age (**Figure 7**). The change in state-5 was steeper, MDT decreased by 1% (2 windows) per year both in boys and in girls, with no age-by-sex interaction (LR test $p = 0.361$) (**Figure 7**).

The number of transitions between states decreased over time, and this association was stronger in girls than in boys, with age-by-sex interaction (LR test $p = 0.016$) (**Tables 3**, **4**). The predicted number of transitions changed from NT = 8.5 around age-9 to NT = 7 at age-14 in girls (**Figure 7**).

Similar results were observed in the models that were additionally adjusted for maternal education (**Supplementary Table 2**).

DISCUSSION

This is the largest longitudinal population-based study describing individual changes in dynamic brain connectivity from childhood into adolescence. We highlight three findings that show developmental patterns. First, we found a general increase in the variability of the connections between intrinsic connectivity networks with increasing age. Second, the time spent in a modularized state increased with age, while the time spent in less modularized states decreased with age. Third, the number of transitions between states decreased with age. Girls showed a more mature pattern of dynamic brain connectivity, spending more time in a highly modularized state, less time in specific states that were more frequently observed at a younger age, transitioning less between states and showing a faster decrease of time spent in a non-modularized state across age than boys.

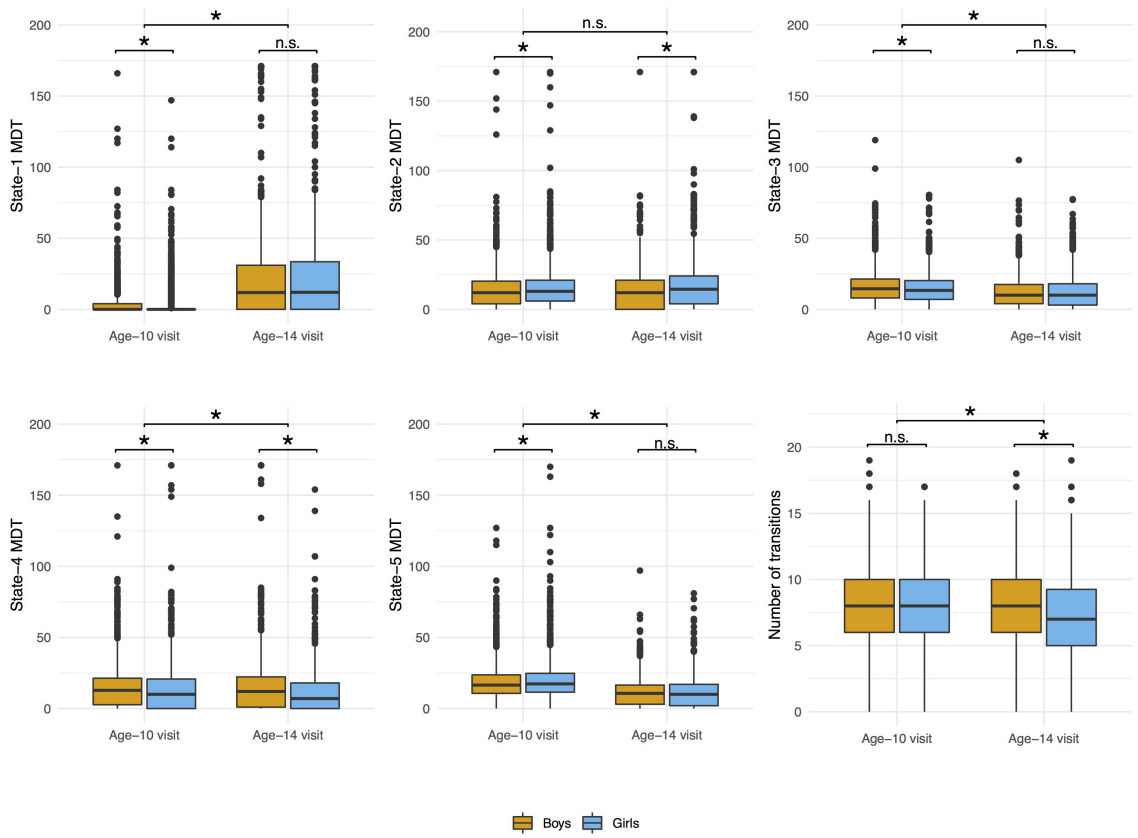

FIGURE 6 | Distributions of mean dwell time (MDT, number of time windows) in each state and number of transitions between states by visit and sex. Wilcoxon signed rank test was used to compare the values between age-10 and age-14 visits and Wilcoxon rank sum test was used to compare the values between boys and girls in each visit. *p-value < 0.05; n.s., non-significant.

TABLE 2 | Age- and sex-associations with transformed mean dwell time (MDT, number of time windows) in each state and number of transitions (NT) between states.

	Age				Sex (ref. boys)				AIC	BIC
		95% CI				95% CI				
	Estimate	Lower	Upper	*P*-value	Estimate	Lower	Upper	*P*-value		
State-1 MDT	0.208	0.196	0.221	< 0.001*	−0.052	−0.104	−0.001	0.046	12678.65	12711.02
State-2 MDT	0.001	−0.013	0.015	0.872	0.152	0.096	0.209	< 0.001*	13594.54	13626.9
State-3 MDT	−0.082	−0.096	−0.068	< 0.001*	−0.074	−0.130	−0.018	0.009*	13480.18	13512.55
State-4 MDT	−0.036	−0.049	−0.022	< 0.001*	−0.244	−0.302	−0.186	< 0.001*	13476.86	13509.23
State-5 MDT	−0.171	−0.185	−0.158	< 0.001*	0.040	−0.014	0.093	0.145	13003.42	13035.78
NT	−0.178	−0.221	−0.134	< 0.001*	−0.292	−0.468	−0.115	0.001*	24485.35	24517.71

*Linear mixed-effects models (random effect: subject). The MDT outcomes were transformed using Box–Cox. Age was centered to the mean age of the sample at age-10 visit. *P-value corrected for multiple comparisons (FDR) < 0.05. AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion.*

The higher variability in the connections between networks observed with increasing age is consistent with previous cross-sectional studies (Hutchison and Morton, 2015; Marusak et al., 2017). This broader repertoire of functional connections between brain regions could be a neural substrate of a higher cognitive complexity. Some of our findings regarding the associations between age and the time spent in specific dynamic states are consistent with previous research. Using cross-sectional data of the Generation R Study, but a younger age visit than this study (6–10 years old), Rashid et al. (2018) also found that older children showed longer MDT in a globally modularized state, characterized by intra- and inter-network connectivity. We found negative age associations with MDT in state-3, in which the DMN was negatively correlated with the other networks. We expected this type of connectivity pattern to be positively associated with age, given the modularity of the state and the fact that the

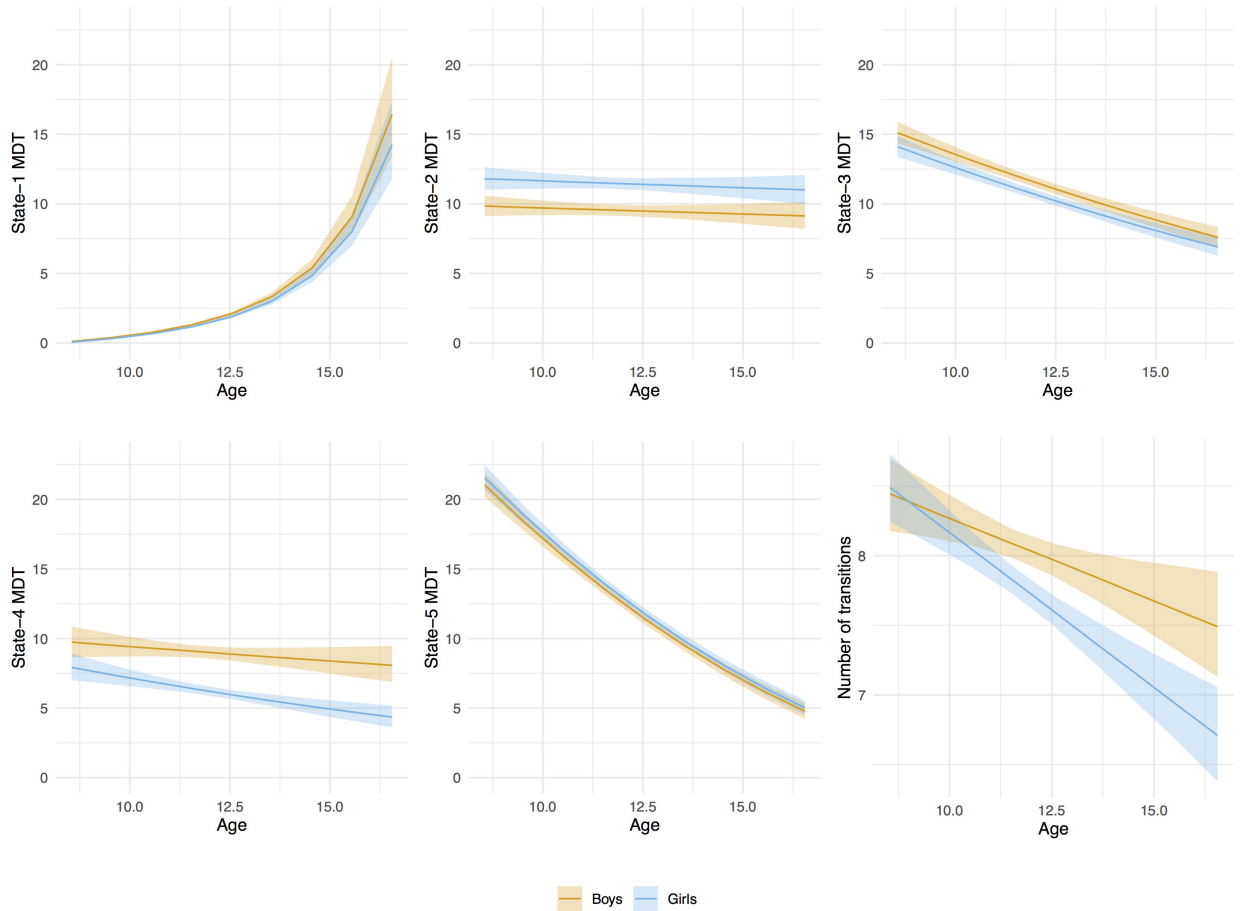

FIGURE 7 | Age-associations with mean dwell time (MDT, number of time windows) in each state and number of transitions between states by sex. Linear mixed-effects models (random effect: subject). The MDT values were transformed back to the original scale for the graphical representation of the associations. A bootstrap technique was applied using 2,000 resamples with replacement to estimate the variation of the values in the population.

efficiency of the DMN increases as children grow older. It is possible that state-3 is a precursor of state-2, and older children tend to spend less time in state-3 because they transition to the other modularized states (1 and 2).

We reported negative associations between age and time spent in less modularized states, such as states 4 and 5. These results are supported by previous research showing that the developing brain is characterized by an increase of "integration" of functional networks (Fair et al., 2007). The components of a network in those states do not show consistent intra-network connectivity, nor inter-network connectivity. This suggests that the integration, or the increased connectivity within the brain regions that comprise a network, is low. This is expected given adolescence is a period of transition to more efficient brain connectivity, in which widely distributed areas are integrated into complex brain systems. This type of developmental process, in which connections change during adolescence, has recently been identified as "disruptive mode," in contrast to "conservative mode," in which connections already established become more strong (Váša et al., 2020). Myelination and synaptic pruning processes that take place during brain development likely contribute to these changes in functional connections by supporting more efficient neuronal communication. The establishment of these complex functional systems has an impact on higher-order cognition (Giedd et al., 1999; Luna and Sweeney, 2004; Bunge and Wright, 2007; Fair et al., 2007).

We found negative associations between age and the number of transitions between dynamic states. Previous studies did not find such association during rest (Hutchison and Morton, 2015; Marusak et al., 2017). Our findings were statistically significant, however, the change we observed was relatively small. Given the large size of the current sample, the discrepancy in findings could be explained by the higher power of our study. Overall, our findings suggest that older participants show more complex connectivity patterns and they remain longer in specific connectivity configurations.

In terms of the composition of the different dynamic states, or configurations, some of the connections between components were stable between states, such as intra-network connectivity within the SC and the VIS networks. Consistent with previous

TABLE 3 | Age-, sex-, and age-by-sex associations with transformed mean dwell time (MDT, number of time windows) in each state and number of transitions (NT) between states.

	Age				Sex (ref. boys)				Age*sex interaction				AIC	BIC
		95% CI				95% CI				95% CI				
	Estimate	Lower	Upper	*P*-value	Estimate	Lower	Upper	*P*-value	Estimate	Lower	Upper	*P*-value		
State-1 MDT	0.203	0.185	0.222	< 0.001*	−0.068	−0.136	0.000	0.050	0.009	−0.016	0.034	0.492	12687.03	12725.87
State-2 MDT	−0.014	−0.034	0.007	0.187	0.103	0.028	0.178	0.007*	0.028	0.000	0.056	0.051	13599.4	13638.24
State-3 MDT	−0.094	−0.114	−0.073	< 0.001*	−0.113	−0.187	−0.039	0.003*	0.022	−0.006	0.049	0.121	13486.46	13525.3
State-4 MDT	−0.016	−0.035	0.004	0.116	−0.179	−0.254	−0.104	< 0.001*	−0.037	−0.063	−0.010	0.006*	13478.19	13517.03
State-5 MDT	−0.165	−0.184	−0.146	< 0.001*	0.061	−0.009	0.132	0.090	−0.012	−0.039	0.014	0.361	13011.37	13050.21
NT	−0.120	−0.184	−0.056	< 0.001*	−0.102	−0.337	0.132	0.391	−0.107	−0.194	−0.020	0.016*	24485.94	24524.78

*Linear mixed-effects models (random effect: subject). The MDT outcomes were transformed using Box–Cox. Age was centered to the mean age of the sample at age-10 visit. *P-value corrected for multiple comparisons (FDR) < 0.05. AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion.*

TABLE 4 | Sex stratified age-associations with transformed mean dwell time (MDT, number of time windows) in state-4 and number of transitions (NT) between states.

		95% CI		
	Estimate	Lower	Upper	*P*-value
State-4 MDT				
Boys	−0.016	−0.035	0.004	0.111
Girls	−0.053	−0.071	−0.034	< 0.001
NT				
Boys	−0.120	−0.184	−0.056	< 0.001
Girls	−0.227	−0.286	−0.168	< 0.001

Linear mixed-effects models (random effect: subject). The MDT outcome was transformed using Box–Cox. Age was centered to the mean age of the sample at age-10 visit.

work, this finding suggests that network organization in humans is a combination of both static and dynamic connections (Calhoun et al., 2014; Faghiri et al., 2018). The age-related changes in the dynamic connectivity metrics reported in this study indicate that the organization of human connectivity patterns develop progressively across the age spectrum (Faghiri et al., 2018). The largest change in MDT associated with age was observed in state-1. This connectivity configuration resembles one that has previously been identified as a "drowsiness pattern" (Allen et al., 2014; Damaraju et al., 2014, 2020). Interestingly, the frequency of this state increased with the scan progression, which could indicate an increase in the fatigue or a decrease in the anxiety of the participants along the session. At the same time, other states that are likely related to a higher awareness, such as state-2, showed the opposite pattern, its frequency decreased with scan progression. The detection of this drowsiness-related state could be beneficial for other rs-fMRI studies, since it allows to remove the effect of potential drowsiness from the data. Additionally, it could also prove interesting from a clinical perspective, where a particular disorder shows differential associations within this connectivity configuration.

Overall, girls showed a faster development of dynamic connectivity than boys. This is consistent with previous literature (Satterthwaite et al., 2015; Rashid et al., 2018) and may be due to an earlier onset of puberty in girls. We observed that girls spent more time than boys in the default-mode/sensorimotor modularized state (state-2) state across the whole age range. In this state, the DMN and the SM showed opposite activation patterns and they were negatively correlated between them; when one of those networks is activated, the other is deactivated. The DMN has been linked to internally focused thought, episodic memory, and planning the future (Buckner et al., 2008), while the SM is related to the processing of external stimuli and motor information. Hence, girls were more prone than boys to show a connectivity configuration in which the DMN and SM networks were negatively correlated, which could indicate a higher efficiency in the synchronized activation and deactivation of those networks. Similar patterns of modular organization between sensory systems and DMN have been observed in young adults (Allen et al., 2014). Faghiri et al. (2018) using

cross-sectional data from a broader age range (3–21 years old), observed that older participants spent more time in states showing intra-network connectivity within the DMN, and inter-network connectivity between the SM and CC networks. These connectivity traits were similar to the ones of the default-mode/sensorimotor modularized state (state-2) that we obtained. However, in our study we did not find associations between age and the time spent in this state. Girls showed 1 year of advantage in state-3 MDT decrease in relation to boys and they showed a faster MDT decrease across age in the non-modularized state-4, more commonly observed in younger children. In addition, girls transitioned less between states than boys, which also suggests more mature connectivity.

One of the most important limitations of developmental fMRI studies is motion (Power et al., 2012). In order to reduce the impact of motion, we implemented various strategies at several levels of the analysis. First, we excluded datasets with excessive motion applying a strict threshold (Parkes et al., 2018). Second, we used a spatially constrained group-independent component analysis approach, only including components that were identified as not being noise. Third, we added a standard set of motion regressors to the dynamic FNC analysis (Satterthwaite et al., 2013). Indeed, there was no relationship between age and motion in our sample after excluding the datasets with excessive motion. The drawback of these actions is that participants who move more are underrepresented in the analyses, potentially leading to selection bias. In fact, we observed some differences between the participants with data only at the age-10 visit, those with data only at age-14 visit, and those with repeated measurements. The proportion of girls was higher in the second visit, the socioeconomic status was lower in the participants with data only at the second visit, while a higher socioeconomic level was more common among those with data at the two time points. Despite these small differences in the socioeconomic status between the groups, the inclusion of maternal education in the models as a precision variable did not change the results. Future work should explore the role of socioeconomic status on the development of dynamic connectivity. For example, socioeconomic status has proven an important factor in structural neurodevelopment (Brito and Noble, 2014). In this study, we used a group ICA template generated from a model order of 150, however, analyzing dynamic FNC with higher and lower group ICA model-based templates would also be interesting. Indeed, a full multi-spatial scale FNC analysis appears to provide additional information (Iraji et al., 2021). In terms of the tapered sliding-window approach used in this study, one limitation is related to the selection of the window size. However, it has been demonstrated that 44 s provides reliable connectivity estimations and it is also sensitive to abrupt brain activity (Allen et al., 2014; Qin et al., 2015). In addition, the observations were weighted according the their position within the window to avoid the effect of influential points (Allen et al., 2014). Another relevant limitation of this study was the distribution of the MDT outcomes with a relatively skewed distributions. Despite the Box–Cox transformation, the residuals from the linear mixed model for state-1 were not fully normally distributed. However, different transformations as well as not transforming the data at all, yielded highly similar results, and the large sample size of this study ensures the robustness of the estimates obtained even in non-ideal conditions (Knief and Forstmeier, 2021).

The longitudinal design, the large and multiethnic sample, which was based on the general population, and the use of a single MRI scanner are the main strengths of this study. Longitudinal studies are key to study the development of the brain, since they allow to control for interindividual variability (Kraemer et al., 2000). The advantages of studying the brain at the population level as opposed to using small samples include the higher statistical power, the lower bias and the higher generalizability of the results (Paus, 2010; White et al., 2013; LeWinn et al., 2017). The use of a single scanner is important as it reduces vendor- and hardware-dependent differences, and it avoids the possible influence of the system updates on the longitudinal estimates.

To summarize, we observed longitudinal changes in dynamic connectivity from ages 8–15 years. Particularly, as children mature, they show: (1) a higher variability in the connections between networks; (2) less time in less modularized states; and (3) less transitions between states. Girls showed a more mature pattern of dynamic connectivity. Resting-state functional connectivity is a reliable tool for studying functional neurodevelopment as it does not require an explicit task-based framework and the connectivity of intrinsic networks exhibits high reproducibility between individuals. Dynamic brain connectivity approaches offer a more comprehensive view of functional connectivity than static connectivity alone and they provide summary metrics, which are likely more reproducible than many thousands of individual edge comparisons. In conclusion, the changes of dynamic connectivity over the course of development presented in this work provide a meaningful baseline for comparison in deviations from typical development.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because of legal and ethical regulations. Requests to access the datasets should be directed to Vincent Jaddoe, v.jaddoe@erasmusmc.nl.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the local medical ethics committee of the Erasmus MC University Medical Center, Rotterdam, Netherlands. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

ML-V and RLM contributed to conception and design of the study and wrote the first draft of the manuscript. RLM preprocessed the data. ML-V, LP-C, FE-L, RHM, JF, and RLM contributed to data curation. ML-V, OA, JH-G, VC, and

RLM contributed to methodology and software. ML-V performed the analyses. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This project was funded by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement no 707404 (ML-V and FE-L). The opinions expressed in this document reflect only the author's view. The European Commission is not responsible for any use that may be made of the information it contains. Neuroimaging acquisition, image analysis and/or infrastructure were supported by the Sophia Foundation (S18-20, RLM), the Erasmus University Fellowship (RLM), the Netherlands Organization for Health Research and Development (ZonMw) TOP project number 91211021 (TW), and by NIH grants R01MH118695 (VC) and R01EB020407 (VC). Supercomputing resources were provided by the Dutch Scientific Organization (NWO) and SurfSara (Cartesius compute system). RHM was supported by The Netherlands Organization for Scientific Research (NWA Startimpuls 400.17.602). LP-C was supported by the Spanish Ministry of Science, Innovation and Universities through the "Centro de Excelencia Severo Ochoa 2019–2023" Program (CEX2018-000806-S), and the Generalitat de Catalunya through the CERCA Program.

ACKNOWLEDGMENTS

The Generation R Study is conducted by the Erasmus Medical Center in close collaboration with Faculty of Social Sciences of the Erasmus University Rotterdam, the Municipal Health Service Rotterdam area, Rotterdam, and the Stichting Trombosedienst & Artsenlaboratorium Rijnmond (STAR-MDC), Rotterdam. We gratefully acknowledge the contribution of children and parents, general practitioners, hospitals, midwives, and pharmacies in Rotterdam. We also thank Nicole Erler for her statistical support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnsys.2021.724805/full#supplementary-material

REFERENCES

Agcaoglu, O., Wilson, T. W., Wang, Y. P., Stephen, J. M., and Calhoun, V. D. (2020). Dynamic resting-state connectivity differences in eyes open versus eyes closed conditions. *Brain Connect.* 10, 504–519. doi: 10.1089/brain.2020.0768

Agcaoglu, O., Wilson, T. W., Wang, Y. P., Stephen, J., and Calhoun, V. D. (2019). Resting state connectivity differences in eyes open versus eyes closed conditions. *Hum. Brain Mapp.* 40, 2488–2498.

Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., and Calhoun, V. D. (2014). Tracking whole-brain connectivity dynamics in the resting state. *Cereb. Cortex* 24, 663–676.

Allen, E. A., Erhardt, E. B., Damaraju, E., Gruner, W., Segall, J. M., Silva, R. F., et al. (2011). A baseline for the multivariate comparison of resting-state networks. *Front. Syst. Neurosci.* 5:2. doi: 10.3389/fnsys.2011.00002

Avants, B. B., Epstein, C. L., Grossman, M., and Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med. Image Anal.* 12, 26–41. doi: 10.1016/j.media.2007.06.004

Biswal, B., Yetkin, F. Z., Haughton, V. M., and Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* 34, 537–541.

Bos, D. J., Oranje, B., Achterberg, M., Vlaskamp, C., Ambrosino, S., de Reus, M. A., et al. (2017). Structural and functional connectivity in children and adolescents with and without attention deficit/hyperactivity disorder. *J. Child Psychol. Psychiatry* 58, 810–818.

Box, G. E. P., and Cox, D. R. (1964). An analysis of transformations. *J. R. Stat. Soc.* 26, 211–243.

Brito, N. H., and Noble, K. G. (2014). Socioeconomic status and structural brain development. *Front. Neurosci.* 8:276. doi: 10.3389/fnins.2014.00276

Buckner, R. L., Andrews-Hanna, J. R., and Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Ann. N. Y. Acad. Sci.* 1124, 1–38.

Bunge, S. A., and Wright, S. B. (2007). Neurodevelopmental changes in working memory and cognitive control. *Curr. Opin. Neurobiol.* 17, 243–250. doi: 10.1016/j.conb.2007.02.005

Calhoun, V. D., Adali, T., Pearlson, G. D., and Pekar, J. J. (2001). A method for making group inferences from functional MRI data using independent component analysis. *Hum. Brain Mapp.* 14, 140–151.

Calhoun, V. D., and Adalí, T. (2012). Multisubject independent component analysis of fMRI: a decade of intrinsic networks, default mode, and neurodiagnostic discovery. *IEEE Rev. Biomed. Eng.* 5, 60–73. doi: 10.1109/RBME.2012.2211076

Calhoun, V. D., Miller, R., Pearlson, G., and Adalı, T. (2014). The chronnectome: time-varying connectivity networks as the next frontier in fMRI data discovery. *Neuron* 84, 262–274. doi: 10.1016/j.neuron.2014.10.015

Cole, D. M., Smith, S. M., and Beckmann, C. F. (2010). Advances and pitfalls in the analysis and interpretation of resting-state FMRI data. *Front. Syst. Neurosci.* 4:8. doi: 10.3389/fnsys.2010.00008

Cordes, D., Haughton, V. M., Arfanakis, K., Carew, J. D., Turski, P. A., Moritz, C. H., et al. (2001). Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data. *AJNR Am. J. Neuroradiol.* 22, 1326–1333.

Crone, E. A., and Elzinga, B. M. (2015). Changing brains: how longitudinal functional magnetic resonance imaging studies can inform us about cognitive and social-affective growth trajectories. *Wiley Interdiscip. Rev. Cogn. Sci.* 6, 53–63. doi: 10.1002/wcs.1327

Damaraju, E., Allen, E. A., Belger, A., Ford, J. M., McEwen, S., Mathalon, D. H., et al. (2014). Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia. *Neuroimage Clin.* 5, 298–308.

Damaraju, E., Tagliazucchi, E., Laufs, H., and Calhoun, V. D. (2020). Connectivity dynamics from wakefulness to sleep. *NeuroImage* 220:117047. doi: 10.1016/j.neuroimage.2020.117047

Delamillieure, P., Doucet, G., Mazoyer, B., Turbelin, M. R., Delcroix, N., Mellet, E., et al. (2010). The resting state questionnaire: an introspective questionnaire for evaluation of inner experience during the conscious resting state. *Brain Res. Bull.* 81, 565–573. doi: 10.1016/j.brainresbull.2009.11.014

Di Martino, A., Fair, D. A., Kelly, C., Satterthwaite, T. D., Castellanos, F. X., Thomason, M. E., et al. (2014). Unraveling the miswired connectome: a developmental perspective. *Neuron* 83, 1335–1353. doi: 10.1016/j.neuron.2014.08.050

Du, Y., and Fan, Y. (2013). Group information guided ICA for fMRI data analysis. *Neuroimage* 69, 157–197. doi: 10.1016/j.neuroimage.2012.11.008

Du, Y., Fu, Z., Sui, J., Gao, S., Xing, Y., Lin, D., et al. (2020). NeuroMark: an automated and adaptive ICA based pipeline to identify reproducible fMRI markers of brain disorders. *NeuroImage* 28:102375. doi: 10.1016/j.nicl.2020.102375

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Stat.* 7, 1–26.

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., et al. (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nat. Methods* 16, 111–116. doi: 10.1038/s41592-018-0235-4

Faghiri, A., Stephen, J. M., Wang, Y. P., Wilson, T. W., and Calhoun, V. D. (2018). Changing brain connectivity dynamics: from early childhood to adulthood. *Hum. Brain Mapp.* 39, 1108–1117.

Fair, D. A., Dosenbach, N. U., Church, J. A., Cohen, A. L., Brahmbhatt, S., Miezin, F. M., et al. (2007). Development of distinct control networks through segregation and integration. *Proc. Natl. Acad. Sci. U.S.A.* 104, 13507–13512. doi: 10.1073/pnas.0705843104

Fonov, V. S., Evans, A. C., McKinstry, R. C., Almli, C. R., and Collins, D. L. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage* 47:S102.

Friedman, J., Hastie, T., and Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* 9, 432–441.

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., et al. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Nat. Neurosci.* 2, 861–863.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl. Acad. Sci. U.S.A.* 101, 8174–8179. doi: 10.1073/pnas.0402680101

Hutchison, R. M., and Morton, J. B. (2015). Tracking the brain's functional coupling dynamics over development. *J. Neurosci.* 35, 6849–6859. doi: 10.1523/jneurosci.4638-14.2015

Iraji, A., Faghiri, A., Fu, Z., Rachakonda, S., Kochunov, P., Belger, A., et al. (2021). Multi-spatial scale dynamic interactions between functional sources reveal sex-specific changes in schizophrenia. *bioRxiv* [Preprint] doi: 10.1101/2021.01.04.425222

Knief, U., and Forstmeier, W. (2021). Violating the normality assumption may be the lesser of two evils. *Behav. Res. Methods* doi: 10.3758/s13428-021-01587-5 [Epub ahead of print].

Kooijman, M. N., Kruithof, C. J., van Duijn, C. M., Duijts, L., Franco, O. H., van, I. M. H., et al. (2016). The generation R study: design and cohort update 2017. *Eur. J. Epidemiol.* 31, 1243–1264.

Kraemer, H. C., Yesavage, J. A., Taylor, J. L., and Kupfer, D. (2000). How can we learn about developmental processes from cross-sectional studies, or can we? *Am. J. Psychiatry* 157, 163–171. doi: 10.1176/appi.ajp.157.2.163

Lebel, C., Walker, L., Leemans, A., Phillips, L., and Beaulieu, C. (2008). Microstructural maturation of the human brain from childhood to adulthood. *Neuroimage* 40, 1044–1055. doi: 10.1016/j.neuroimage.2007.12.053

Lenroot, R. K., and Giedd, J. N. (2006). Brain development in children and adolescents: insights from anatomical magnetic resonance imaging. *Neurosci. Biobehav. Rev.* 30, 718–729. doi: 10.1016/j.neubiorev.2006.06.001

LeWinn, K. Z., Sheridan, M. A., Keyes, K. M., Hamilton, A., and McLaughlin, K. A. (2017). Sample composition alters associations between age and brain structure. *Nat. Commun.* 8:874.

Li, X., Morgan, P. S., Ashburner, J., Smith, J., and Rorden, C. (2016). The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *J. Neurosci. Methods* 264, 47–56. doi: 10.1016/j.jneumeth.2016.03.001

López-Vicente, M., Lamballais, S., Louwen, S., Hillegers, M., Tiemeier, H., Muetzel, R. L., et al. (2020). White matter microstructure correlates of age, sex, handedness and motor ability in a population-based sample of 3031 school-age children. *Neuroimage* 227:117643. doi: 10.1016/j.neuroimage.2020.117643

Luciana, M. (2013). Adolescent brain development in normality and psychopathology. *Dev. Psychopathol.* 25(4 Pt 2), 1325–1345. doi: 10.1017/s0954579413000643

Luna, B., and Sweeney, J. A. (2004). The emergence of collaborative brain function: FMRI studies of the development of response inhibition. *Ann. N. Y. Acad. Sci.* 1021, 296–309. doi: 10.1196/annals.1308.035

Luna, B., Thulborn, K. R., Munoz, D. P., Merriam, E. P., Garver, K. E., Minshew, N. J., et al. (2001). Maturation of widely distributed brain function subserves cognitive development. *Neuroimage* 13, 786–793. doi: 10.1006/nimg.2000.0743

Marusak, H. A., Calhoun, V. D., Brown, S., Crespo, L. M., Sala-Hamrick, K., Gotlib, I. H., et al. (2017). Dynamic functional connectivity of neurocognitive networks in children. *Hum. Brain Mapp.* 38, 97–108. doi: 10.1002/hbm.23346

Muetzel, R. L., Blanken, L. M., Thijssen, S., van der Lugt, A., Jaddoe, V. W., Verhulst, F. C., et al. (2016). Resting-state networks in 6-to-10 year old children. *Hum. Brain Mapp.* 37, 4286–4300. doi: 10.1002/hbm.23309

Nelson, E. E., Leibenluft, E., McClure, E. B., and Pine, D. S. (2005). The social re-orientation of adolescence: a neuroscience perspective on the process and its relation to psychopathology. *Psychol. Med.* 35, 163–174. doi: 10.1017/s0033291704003915

Parkes, L., Fulcher, B., Yücel, M., and Fornito, A. (2018). An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. *Neuroimage* 171, 415–436. doi: 10.1016/j.neuroimage.2017.12.073

Paus, T. (2010). Population neuroscience: why and how. *Hum. Brain Mapp.* 31, 891–903. doi: 10.1002/hbm.21069

Paus, T., Keshavan, M., and Giedd, J. N. (2008). Why do many psychiatric disorders emerge during adolescence? *Nat. Rev. Neurosci.* 9, 947–957. doi: 10.1038/nrn2513

Perrin, J. S., Leonard, G., Perron, M., Pike, G. B., Pitiot, A., Richer, L., et al. (2009). Sex differences in the growth of white matter during adolescence. *Neuroimage* 45, 1055–1066. doi: 10.1016/j.neuroimage.2009.01.023

Peterson, R. A., and Cavanaugh, J. E. (2020). Ordered quantile normalization: a semiparametric transformation built for the cross-validation era. *J. Appl. Stat.* 47, 2312–2327. doi: 10.1080/02664763.2019.1630372

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team (2019). *nlme: Linear and Nonlinear Mixed Effects Models. R package Version 3.1-139.*

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59, 2142–2154. doi: 10.1016/j.neuroimage.2011.10.018

Qin, J., Chen, S. G., Hu, D., Zeng, L. L., Fan, Y. M., Chen, X. P., et al. (2015). Predicting individual brain maturity using dynamic functional connectivity. *Front. Hum. Neurosci.* 9:418. doi: 10.3389/fnhum.2015.00418

Rashid, B., Blanken, L. M. E., Muetzel, R. L., Miller, R., Damaraju, E., Arbabshirani, M. R., et al. (2018). Connectivity dynamics in typical development and its relationship to autistic traits and autism spectrum disorder. *Hum. Brain Mapp.* 39, 3127–3142. doi: 10.1002/hbm.24064

Rubia, K., Overmeyer, S., Taylor, E., Brammer, M., Williams, S. C., Simmons, A., et al. (2000). Functional frontalisation with age: mapping neurodevelopmental trajectories with fMRI. *Neurosci. Biobehav. Rev.* 24, 13–19. doi: 10.1016/s0149-7634(99)00055-x

Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059–1069. doi: 10.1016/j.neuroimage.2009.10.003

Salman, M. S., Du, Y., Lin, D., Fu, Z., Fedorov, A., Damaraju, E., et al. (2019). Group ICA for identifying biomarkers in schizophrenia: 'adaptive' networks via spatially constrained ICA show more sensitivity to group differences than spatio-temporal regression. *Neuroimage Clin.* 22:101747. doi: 10.1016/j.nicl.2019.101747

Satterthwaite, T. D., Elliott, M. A., Gerraty, R. T., Ruparel, K., Loughead, J., Calkins, M. E., et al. (2013). An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *Neuroimage* 64, 240–256. doi: 10.1016/j.neuroimage.2012.08.052

Satterthwaite, T. D., Wolf, D. H., Roalf, D. R., Ruparel, K., Erus, G., Vandekar, S., et al. (2015). Linked sex differences in cognition and functional connectivity in youth. *Cereb. Cortex* 25, 2383–2394. doi: 10.1093/cercor/bhu036

Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., et al. (2011). Network modelling methods for FMRI. *Neuroimage* 54, 875–891. doi: 10.1016/j.neuroimage.2010.08.063

Váša, F., Romero-Garcia, R., Kitzbichler, M. G., Seidlitz, J., Whitaker, K. J., Vaghi, M. M., et al. (2020). Conservative and disruptive modes of adolescent change in human brain functional connectivity. *Proc. Natl. Acad. Sci. U.S.A.* 117, 3248–3253. doi: 10.1073/pnas.1906144117

White, T., El Marroun, H., Nijs, I., Schmidt, M., van der Lugt, A., Wielopolki, P. A., et al. (2013). Pediatric population-based neuroimaging and the generation R study: the intersection of developmental neuroscience and epidemiology. *Eur. J. Epidemiol.* 28, 99–111. doi: 10.1007/s10654-013-9768-0

White, T., Muetzel, R. L., El Marroun, H., Blanken, L. M. E., Jansen, P., Bolhuis, K., et al. (2018). Paediatric population neuroimaging and the generation R study: the second wave. *Eur. J. Epidemiol.* 33, 99–125. doi: 10.1007/s10654-017-0319-y

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Human Neuroscience

ORIGINAL RESEARCH
published: 13 January 2022
doi: 10.3389/fnhum.2021.781554

General Psychopathology, Cognition, and the Cerebral Cortex in 10-Year-Old Children: Insights From the Adolescent Brain Cognitive Development Study

*Yash Patel[1], Nadine Parker[1], Giovanni A. Salum[2], Zdenka Pausova[3] and Tomáš Paus[1,4]**

[1] Institute of Medical Sciences, University of Toronto, Toronto, ON, Canada, [2] Department of Psychiatry, Federal University of Rio Grande do Sul, Porto Alegre, Brazil, [3] The Hospital for Sick Children, University of Toronto, Toronto, ON, Canada, [4] Departments of Psychiatry and Neuroscience, Faculty of Medicine and Centre Hopitalier Universitaire Sainte-Justine, University of Montreal, Montreal, QC, Canada

OPEN ACCESS

Edited by:
Urvakhsh Meherwan Mehta, National Institutes of Health (NIH), United States

Reviewed by:
Bharath Holla, National Institute of Mental Health and Neurosciences (NIMHANS), India
Rakshathi Basavaraju, Columbia University Irving Medical Center, United States

***Correspondence:**
Tomáš Paus
tomas.paus@umontreal.ca

Specialty section:
This article was submitted to Brain Imaging and Stimulation, a section of the journal Frontiers in Human Neuroscience

Received: *22 September 2021*
Accepted: *29 November 2021*
Published: *13 January 2022*

Citation:
Patel Y, Parker N, Salum GA, Pausova Z and Paus T (2022) General Psychopathology, Cognition, and the Cerebral Cortex in 10-Year-Old Children: Insights From the Adolescent Brain Cognitive Development Study. Front. Hum. Neurosci. 15:781554. doi: 10.3389/fnhum.2021.781554

General psychopathology and cognition are likely to have a bidirectional influence on each other. Yet, the relationship between brain structure, psychopathology, and cognition remains unclear. This brief report investigates the association between structural properties of the cerebral cortex [surface area, cortical thickness, intracortical myelination indexed by the T1w/T2w ratio, and neurite density assessed by restriction spectrum imaging (RSI)] with general psychopathology and cognition in a sample of children from the Adolescent Brain Cognitive Development (ABCD) study. Higher levels of psychopathology and lower levels of cognitive ability were associated with a smaller cortical surface area. Inter-regionally—across the cerebral cortex—the strength of association between an area and psychopathology is strongly correlated with the strength of association between an area and cognition. Taken together, structural deviations particularly observed in the cortical surface area influence both psychopathology and cognition.

Keywords: MRI, brain development, cerebral cortex, growth, cohort

INTRODUCTION

There is overwhelming evidence demonstrating the shared heritability of psychiatric disorders (Anttila et al., 2018). Given the high rates of comorbidity (Plana-Ripoll et al., 2019), commonality in genetic and environmental risk factors (Uher and Zwicker, 2017), a transdiagnostic perspective is warranted. The "p" factor, or general psychopathology factor, is one such approach capturing latent structures of psychopathology across many disorders (Lahey et al., 2012; Caspi et al., 2014). Impairments in cognitive functioning are observed across many psychiatric disorders (Gale et al., 2010; Urfer-Parnas et al., 2010). Conceptual frameworks have suggested a bidirectional relationship between cognitive function and psychopathology (Batty et al., 2005; Calvete et al., 2013; Schweizer and Hankin, 2018). With the advent of large-scale magnetic resonance imaging (MRI) studies, group differences in the structural properties of the cerebral cortex (predominantly cortical thickness) have been reported in common psychiatric disorders (Ching et al., 2020; Hoogman et al., 2020;

Thompson et al., 2020; Van den Heuvel et al., 2020; Patel et al., 2021), as well as in relation to general psychopathology (Mewton et al., 2020; Romer et al., 2021a,b) and cognitive ability (Shaw et al., 2006; Karama et al., 2014).

In this study, we investigate the association between several properties of the cerebral cortex, namely the surface area, cortical thickness, the T1w/T2w ratio (potentially an index of myelination) (Glasser and Van Essen, 2011), and neurite density [as indexed by restriction spectrum imaging (RSI)] (White et al., 2013), with general psychopathology and cognitive ability in a large sample of children from the Adolescent Brain Cognitive Development Study (ABCD) (Casey et al., 2018).

MATERIALS AND METHODS

Magnetic resonance imaging data (T1-weighted, T2-weighted, and diffusion tensor imaging) from the ABCD study of 11,753 children (mean age, 9.9 years; 48% female) were acquired and processed as described previously (Casey et al., 2018). For twin pairs, only one twin was selected at random to assess unrelated individuals only. Following quality control of the FreeSurfer pipeline (Fischl, 2012) (as described in the ABCD white papers) and removing outliers based on three times the standard deviation, there were 8,869, 8,885, 8,474, and 8,301 participants for the cortical area, thickness, T1w/T2w, and neurite density, respectively. The ABCD study conducted manual quality control of the FreeSurfer cortical surface reconstruction by scoring the extent/severity of artifacts, namely motion, intensity in homogeneity, white-matter underestimation, pial overestimation, and magnetic susceptibility artifacts. Cortical measures were averaged between the two hemispheres for each of the 34 regions of the Desikan—Killiany atlas derived by FreeSurfer (Desikan et al., 2006). Cortical thickness and the surface area were estimated through the FreeSurfer cortical reconstruction pipeline (Fischl, 2012). Neurite density was estimated by RSI using the restricted normalized directional maps, indexing intracellular and directional movement of water through neurites (White et al., 2013). The T1w/T2w ratio was quantified as the ratio of T1-weighted and T2-weighted maps sampled within the cortical ribbon (detailed in ABCD white papers, and Casey et al., 2018).

A bi-factor confirmatory factor analysis on the Child Behavior Check List (parent completed) was used to extract a general psychopathology factor, and internalizing and externalizing factors using the R package "lavaan" (Rosseel, 2012). A total of 12 questions from the CBCL questionnaire were not included in the model as they occurred with very low frequency in the sample population (<1%). The model was fit using the diagonally weighted least squares estimators implemented in "lavaan." P-factor model item loadings, model fit, and the 12 excluded questions are presented in **Supplementary Tables 1–3**. The comparative fit index for the bi-factor model is 0.964, which agrees with a generally accepted threshold of good model fit of >0.95. To quantify cognitive ability, a total cognitive composite score was extracted from the youth NIH Toolbox cognitive battery. NIH Total composite measure included the following cognitive tests: flanker, dimensional change card, picture sequence memory, list sorting, pattern, oral reading, and picture vocabulary (Weintraub et al., 2013).

The relationships between psychopathology and cortical measures were modeled using linear mixed effects where psychopathology and cognition were modeled as a function of fixed effects (cortical measure, age, sex, and ethnicity), and random effects for MRI serial number (due to multiple scanners used in the ABCD study). *P*-values were corrected for multiple comparisons (34 regions tested and for each of the 4 MRI modalities for a total of 136 tests) using False Discovery Rate (FDR) (Benjamini and Hochberg, 1995).

To test the presence of mediation by cognition (or psychopathology) on the relationship between the surface area and psychopathology (or cognition), we used a simple mediation framework implemented by the "mediation" R package (Imai et al., 2011; Tingley et al., 2014). Specifically, we used a similar linear mixed effects model as above, adjusting for age, sex, ethnicity, and MRI serial number to estimate the direct effect (i.e., average direct effect, ADE), indirect effect (average causal mediation effect, ACME), and the proportion of total effect mediated. Confidence intervals were estimated using quasi-Bayesian Monte Carlo approximation (Tingley et al., 2014).

RESULTS

We reveal subtle yet robust associations between cortical structure and general psychopathology and cognitive ability (**Figure 1** and **Supplementary Figure 1**). The cortical surface area is negatively associated with psychopathology and positively associated with cognition across all cortical regions (FDR $p < 0.05$). The cortical T1w/T2w ratio and neurite density are positively associated with psychopathology, predominately in the frontal lobe. Sex-stratified analyses are reported in **Supplementary Figures 2**, **3**. Cortical thickness is associated with cognition in eight cortical regions but shows very little association with psychopathology (**Figure 1**). Little to no associations are present between cognition and either T1w/T2w ratio or neurite density.

Across individuals, cognitive function is weakly correlated with psychopathology ($R^2 = 0.03$, $p < 0.0001$). But across cortical regions, we observe a robust association between the two interregional profiles; associations between the cortical area and psychopathology correlate—across the 34 regions—with associations between the cortical area and cognitive function ($R^2 = 0.74$, $p < 0.0001$; **Figure 2**). To some extent, this relationship between association-based profiles is found also with thickness, T1w/T2w ratio, and neurite density. Regression model statistics can be found in **Supplementary Tables 4**, **5**.

Finally, we examined if cognition (or psychopathology) mediates the relationship between the surface area and psychopathology (or cognition, **Supplementary Figure 4**). Cognition mediates between 20 and 40% of the total effect between the surface area and psychopathology, varying across the 34 regions (**Supplementary Figures 4A,B**). On the other

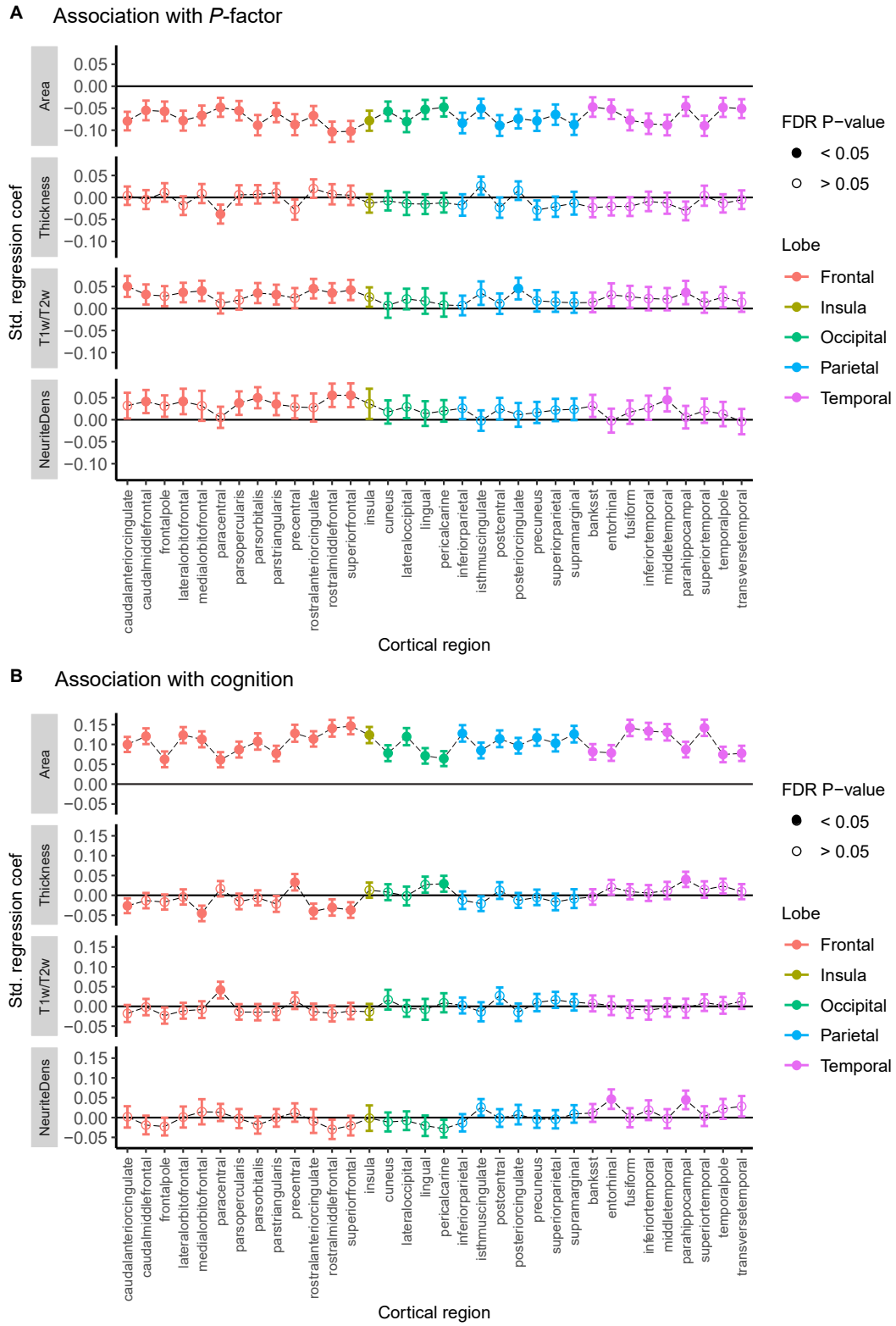

FIGURE 1 | Association between general psychopathology **(A)** and cognitive score **(B)** with cortical thickness, the surface area, the T1w/T2w ratio, and neurite density (labeled "NeuriteDens") across the 34 cortical regions of the Desikan—Killiany atlas. Standardized effect sizes (betas) plotted on the *y*-axis from linear mixed models adjusting for the effect of age, sex, and scanner effects. Error bars represent 95% confidence intervals for the estimates. Filled-in circles represent FDR-corrected $p < 0.05$.

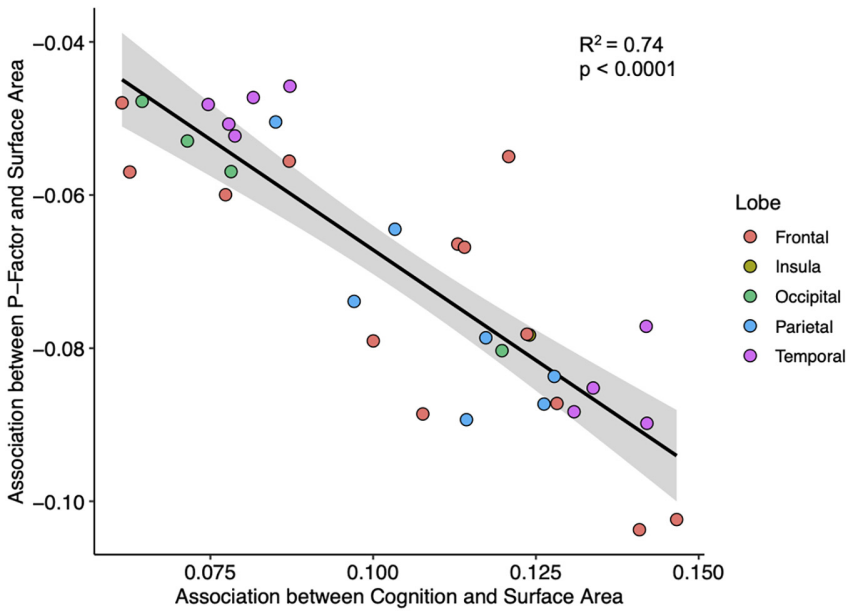

FIGURE 2 | A plot of interregional associations between the surface area and cognition (*x*-axis) and interregional associations between the surface area and general psychopathology (*y*-axis) across the 34 regions of the cerebral cortex. Each point represents 1 of the 34 cortical regions. Line and shaded portions represent, respectively, linear fit and 95% confidence interval for the linear fit. Pearson correlation = −0.86, R^2 = 0.74, $p < 0.0001$.

hand, psychopathology mediates between 6 and 13% of the total effect between the surface area and cognition (**Supplementary Figures 4B,C**). This is a tentative analysis and should be interpreted with caution as mediation analysis of cross-sectional data cannot assess the directionality of these results.

DISCUSSION

This report examines the relationship between multimodal measures of the cerebral cortex with cognition and general psychopathology in a large set of children from the ABCD study. There are robust associations between the radial growth of the cerebral cortex (as reflected in the cortical surface area), and both general psychopathology and cognition, as well as more subtle variations with cytoarchitectonic (neurite density) and myeloarchitectonic (T1W/TW2 ratio) features. These associations may reflect variations in developmental trajectories likely starting prenatally (the surface area) (Rakic, 1988), and continuing postnatally (intracortical myelination, dendritic branching) (Hill et al., 2010; Whitaker et al., 2016; Patel et al., 2019). It is important to note that the neurobiological underpinnings for many of these MRI-derived indices are not fully clear, and are unlikely to be specific to a single microstructural feature, such as myelin or neurite density.

In addition, there is a strong, inverse relationship between associations of the cortical area with cognition and psychopathology, respectively. Hypothetically, this may indicate an overlap between genetic (Shin et al., 2020) and environmental factors imparting—in parallel—the two behavioral phenotypes *via* the radial expansion of the cerebral cortex during prenatal development and the first few years of life. A majority of the expansion of the cerebral cortex, reflected in the surface area, occurs during prenatal and perinatal time periods (Li et al., 2013). Cross-disorder psychiatric genome-wise association studies (GWAS) point toward a role in prenatal neurodevelopment across multiple conditions, and also reveal a negative genetic correlation with GWAS of cognitive ability (Lee et al., 2019). Similarly, genetic studies of intelligence reveal the importance of neurodevelopmental processes (Savage et al., 2018). Finally, neuron progenitor specific regulatory elements are enriched with GWAS loci associated with the cortical surface area, psychiatric disorders (e.g., schizophrenia, autism, and major depression), and with intelligence and education attainment (Liang et al., 2021). Taken together, it is possible that neurodevelopment during gestation connects the processes underlying cortical growth with psychopathology and cognition. It is also possible that such early developmental events cascade into later cognitive development and psychopathology in a sequential manner (e.g., from lower cognitive abilities to higher psychopathology or *vice versa*) (Schweizer and Hankin, 2018). We have shown that a number of adverse perinatal events (e.g., hypoxia, maternal hypertension) share their molecular architecture with that underlying neurodevelopmental processes involved in cortical growth during the same period (Patel and Paus, under review).

Finally, the mediation analysis revealed differences in the amount of mediation by cognition as compared with psychopathology. This aligns with the observed lower levels of premorbid IQ in those who later develop various mental illnesses, including schizophrenia, mood disorders, substance use disorders, and any disorder, in general (Gale et al., 2010).

However, it is important to note that this analysis is highly exploratory and limited by the cross-sectional nature of the analysis in this report. Temporal precedence is required *via* longitudinal data to assess directionality and ensure correct model specification (MacKinnon et al., 2007). Similarly, mediation analysis relies on strong assumptions of sequential ignorability, such that there are no unobserved covariates that influence between the independent variable (area) and the mediator, or between the mediator and the dependent variable (MacKinnon et al., 2007; Imai et al., 2010). Modeling of forthcoming longitudinal data will provide much-needed insights into the directionality of brain-psychopathology-cognition relationships, and possible strategies for modifying (unfavorable) developmental trajectories.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation. The data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Development SM (ABCD) Study (https://abcdstudy.org), held in the NIMH Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children aged 9–10 and follow them over 10 years into early adulthood.

ETHICS STATEMENT

This study was reviewed and approved by the Hospital for Sick Children Institutional Review Board (IRB # 1000073323), and by the NIMH Data Archive (data access request ID 6959). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

YP and TP conceived the research. YP wrote the first draft. GS and NP supervised the calculation of P factor. YP, NP, GS, ZP, and TP reviewed the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

YP was supported by the Alexander Graham Bell Canada Graduate Scholarship from the Natural Sciences and Engineering Research Council of Canada. The ABCD Study is supported by the National Institutes of Health and additional federal partners under award Nos. U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, and U24DA041147. A full list of supporters is available at https://abcdstudy.org/federal-partners.html. A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/consortium_members/. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in the analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnhum.2021.781554/full#supplementary-material

REFERENCES

Anttila, V., Bulik-Sullivan, B., Finucane, H. K., Walters, R. K., Bras, J., Duncan, L., et al. (2018). Analysis of shared heritability in common disorders of the brain. *Science* 360:eaa8757. doi: 10.1126/science.aap8757

Batty, G. D., Mortensen, E. L., and Osler, M. (2005). Childhood IQ in relation to later psychiatric disorder: evidence from a Danish birth cohort study. *Br. J. Psychiatry* 187, 180–181. doi: 10.1192/bjp.187.2.180

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc. Ser. B* 57, 289–300.

Calvete, E., Orue, I., and Hankin, B. L. (2013). Transactional relationships among cognitive vulnerabilities, stressors, and depressive symptoms in adolescence. *J. Abnorm. Child Psychol.* 41, 399–410. doi: 10.1007/s10802-012-9691-y

Casey, B., Cannonier, T., Conley, M. I., Cohen, A. O., Barch, D. M., Heitzeg, M. M., et al. (2018). The adolescent brain cognitive development (ABCD) study: imaging acquisition across 21 sites. *Dev. Cogn. Neurosci.* 32, 43–54. doi: 10.1016/j.dcn.2018.03.001

Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., et al. (2014). The p factor: one general psychopathology factor in the structure of psychiatric disorders? *Clin. Psychol. Sci.* 2, 119–137. doi: 10.1177/2167702613497473

Ching, C. R., Hibar, D. P., Gurholt, T. P., Nunes, A., Thomopoulos, S. I., Abé, C., et al. (2020). What we learn about bipolar disorder from large-scale neuroimaging: findings and future directions from the ENIGMA Bipolar Disorder Working Group. *Hum. Brain Mapp.* Epub online ahead of print. doi: 10.1002/hbm.25098

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31, 968–980. doi: 10.1016/j.neuroimage.2006.01.021

Fischl, B. (2012). FreeSurfer. *Neuroimage* 62, 774–781. doi: 10.1016/j.neuroimage.2012.01.021

Gale, C. R., Batty, G. D., Tynelius, P., Deary, I. J., and Rasmussen, F. (2010). Intelligence in early adulthood and subsequent hospitalisation and admission rates for the whole range of mental disorders: longitudinal study of 1,049,663 men. *Epidemiology* 21:70.

Glasser, M. F., and Van Essen, D. C. (2011). Mapping human cortical areas in vivo based on myelin content as revealed by T1-and T2-weighted MRI. *J. Neurosci.* 31, 11597–11616. doi: 10.1523/JNEUROSCI.2180-11.2011

Hill, J., Inder, T., Neil, J., Dierker, D., Harwell, J., and Van Essen, D. (2010). Similar patterns of cortical expansion during human development and evolution. *Proc. Natl. Acad. Sci.* 107, 13135–13140. doi: 10.1073/pnas.1001229107

Hoogman, M., Van Rooij, D., Klein, M., Boedhoe, P., Ilioska, I., Li, T., et al. (2020). Consortium neuroscience of attention deficit/hyperactivity disorder and autism

spectrum disorder: the ENIGMA adventure. *Hum. Brain Mapp.* Epub Online ahead of print. doi: 10.1002/hbm.25029

Imai, K., Keele, L., and Tingley, D. (2010). A general approach to causal mediation analysis. *Psychol. Methods* 15:309. doi: 10.1037/a0020761

Imai, K., Keele, L., Tingley, D., and Yamamoto, T. (2011). Unpacking the black box of causality: learning about causal mechanisms from experimental and observational studies. *Am. Polit. Sci. Rev.* 105, 765–789.

Karama, S., Bastin, M. E., Murray, C., Royle, N. A., Penke, L., Maniega, S. M., et al. (2014). Childhood cognitive ability accounts for associations between cognitive ability and brain cortical thickness in old age. *Mol. Psychiatry* 19, 555–559. doi: 10.1038/mp.2013.64

Lahey, B. B., Applegate, B., Hakes, J. K., Zald, D. H., Hariri, A. R., and Rathouz, P. J. (2012). Is there a general factor of prevalent psychopathology during adulthood? *J. Abnorm. Psychol.* 121:971. doi: 10.1037/a0028355

Lee, P. H., Anttila, V., Won, H., Feng, Y.-C. A., Rosenthal, J., Zhu, Z., et al. (2019). Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. *Cell* 179, 1469–1482. doi: 10.1016/j.cell.2019.11.020

Li, G., Nie, J., Wang, L., Shi, F., Lin, W., Gilmore, J. H., et al. (2013). Mapping region-specific longitudinal cortical surface expansion from birth to 2 years of age. *Cerebral Cortex* 23, 2724–2733. doi: 10.1093/cercor/bhs265

Liang, D., Elwell, A. L., Aygün, N., Krupa, O., Wolter, J. M., Kyere, F. A., et al. (2021). Cell-type-specific effects of genetic variation on chromatin accessibility during human neuronal differentiation. *Nat. Neurosci.* 24, 941–953. doi: 10.1038/s41593-021-00858-w

MacKinnon, D. P., Fairchild, A. J., and Fritz, M. S. (2007). Mediation analysis. *Annu. Rev. Psychol.* 58, 593–614.

Mewton, L., Lees, B., Squeglia, L., Forbes, M. K., Sunderland, M., Krueger, R., et al. (2020). The relationship between brain structure and general psychopathology in preadolescents. PsyArXiv doi: 10.31234/osf.io/r4pxy

Patel, Y., Parker, N., Shin, J., Howard, D., French, L., Thomopoulos, S. I., et al. (2021). Virtual histology of cortical thickness and shared neurobiology in 6 psychiatric disorders. *JAMA Psychiatry* 78, 47–63. doi: 10.1001/jamapsychiatry.2020.2694

Patel, Y., Shin, J., Gowland, P., Pausova, Z., and Paus, T. (2019). Maturation of the human cerebral cortex during adolescence: myelin or dendritic arbor? *Cerebral Cortex* 29, 3351–3362. doi: 10.1093/cercor/bhy204

Plana-Ripoll, O., Pedersen, C. B., Holtz, Y., Benros, M. E., Dalsgaard, S., De Jonge, P., et al. (2019). Exploring comorbidity within mental disorders among a Danish national population. *JAMA Psychiatry* 76, 259–270. doi: 10.1001/jamapsychiatry.2018.3658

Rakic, P. (1988). Specification of cerebral cortical areas. *Science* 241, 170–176. doi: 10.1126/science.3291116

Romer, A. L., Elliott, M. L., Knodt, A. R., Sison, M. L., Ireland, D., Houts, R., et al. (2021a). Pervasively thinner neocortex as a transdiagnostic feature of general psychopathology. *Am. J. Psychiatry* 178, 174–182. doi: 10.1176/appi.ajp.2020.19090934

Romer, A. L., Knodt, A. R., Sison, M. L., Ireland, D., Houts, R., Ramrakha, S., et al. (2021b). Replicability of structural brain alterations associated with general psychopathology: evidence from a population-representative birth cohort. *Mol. Psychiatry* 26, 3839–3846. doi: 10.1038/s41380-019-0621-z

Rosseel, Y. (2012). Lavaan: an R package for structural equation modeling and more. Version 0.5–12 (BETA). *J. Statist. Softw.* 48, 1–36.

Savage, J. E., Jansen, P. R., Stringer, S., Watanabe, K., Bryois, J., De Leeuw, C. A., et al. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat. Genet.* 50, 912–919. doi: 10.1038/s41588-018-0152-6

Schweizer, T. H., and Hankin, B. L. (2018). "Cognitive Risks: translating stress into psychopathology" in *The Oxford Handbook of Stress and Mental Health*. (eds) K. Harkness and E.P. Hayden (New York: Oxford University Press, Inc).

Shaw, P., Greenstein, D., Lerch, J., Clasen, L., Lenroot, R., Gogtay, N., et al. (2006). Intellectual ability and cortical development in children and adolescents. *Nature* 440, 676–679. doi: 10.1038/nature04513

Shin, J., Ma, S., Hofer, E., Patel, Y., Vosberg, D., Tilley, S., et al. (2020). Global and regional development of the human cerebral cortex: molecular Architecture and Occupational Aptitudes. *Cerebral Cortex* 30, 4121–4139. doi: 10.1093/cercor/bhaa035

Thompson, P., Jahanshad, N., Ching, C. R., Salminen, L. E., Thomopoulos, S. I., Bright, J., et al. (2020). ENIGMA and global neuroscience: a decade of large-scale studies of the brain in health and disease across more than 40 countries. *Transl. Psychiatry* 10:100. doi: 10.1038/s41398-020-0705-1

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., and Imai, K. (2014). Mediation: R package for causal mediation analysis. *J. Stat. Softw.* 59, 1–38.

Uher, R., and Zwicker, A. (2017). Etiology in psychiatry: embracing the reality of poly-gene-environmental causation of mental illness. *World Psychiatry* 16, 121–129. doi: 10.1002/wps.20436

Urfer-Parnas, A., Mortensen, E. L., Saebye, D., and Parnas, J. (2010). Pre-morbid IQ in mental disorders: a Danish draft-board study of 7486 psychiatric patients. *Psychol. Med.* 40, 547–556. doi: 10.1017/S0033291709990754

Van den Heuvel, O. A., Boedhoe, P. S., Bertolin, S., Bruin, W. B., Francks, C., Ivanov, I., et al. (2020). An overview of the first 5 years of the ENIGMA obsessive–compulsive disorder working group: the power of worldwide collaboration. *Hum. Brain Mapp.* Epub online ahead of print. doi: 10.1002/hbm.24972

Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J., et al. (2013). Cognition assessment using the NIH Toolbox. *Neurology* 80, S54–S64.

Whitaker, K. J., Vértes, P. E., Romero-Garcia, R., Váša, F., Moutoussis, M., Prabhu, G., et al. (2016). Adolescence is associated with genomically patterned consolidation of the hubs of the human brain connectome. *Proc. Natl. Acad. Sci.* 113, 9105–9110. doi: 10.1073/pnas.1601745113

White, N. S., Leergaard, T. B., D'Arceuil, H., Bjaalie, J. G., and Dale, A. M. (2013). Probing tissue microstructure with restriction spectrum imaging: histological and theoretical validation. *Hum. Brain Mapp.* 34, 327–346. doi: 10.1002/hbm.21454

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 23 February 2022
doi: 10.3389/fnagi.2022.795764

No Association Between Loneliness, Episodic Memory and Hippocampal Volume Change in Young and Healthy Older Adults: A Longitudinal European Multicenter Study

Cristina Solé-Padullés[1]†, Dídac Macià[1,2]†, Micael Andersson[3,4], Mikael Stiernstedt[3,4], Sara Pudas[3,4], Sandra Düzel[5,6], Enikő Zsoldos[7], Klaus P. Ebmeier[7], Julia Binnewies[8], Christian A. Drevon[9,10], Andreas M. Brandmaier[5,6], Athanasia M. Mowinckel[11,12], Anders M. Fjell[11,12], Kathrine Skak Madsen[13,14], William F. C. Baaré[13], Ulman Lindenberger[5,6], Lars Nyberg[3,4,15], Kristine B. Walhovd[11,12] and David Bartrés-Faz[1,16]*

[1] Department of Medicine, Faculty of Medicine and Health Sciences, Institute of Neurosciences, University of Barcelona, Barcelona, Spain, [2] ISGlobal, Hospital Clínic – University of Barcelona, Barcelona, Spain, [3] Department of Integrative Medical Biology, Umeå University, Umeå, Sweden, [4] Umeå Center for Functional Brain Imaging, Umeå University, Umeå, Sweden, [5] Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany, [6] Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Berlin, Germany, [7] Department of Psychiatry, Wellcome Centre for Integrative Neuroimaging, University of Oxford, Oxford, United Kingdom, [8] Department of Psychiatry, Amsterdam Neuroscience, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, Netherlands, [9] Vitas Ltd., Oslo, Norway, [10] Department of Nutrition, Institute of Basic Medical Sciences, Faculty Medicine, University of Oslo, Oslo, Norway, [11] Center for Lifespan Changes in Brain and Cognition, University of Oslo, Oslo, Norway, [12] Department of Radiology and Nuclear Medicine, Oslo University Hospital, Oslo, Norway, [13] Danish Research Centre for Magnetic Resonance, Centre for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital Amager and Hvidovre, Copenhagen, Denmark, [14] Radiography, Department of Technology, University College Copenhagen, Copenhagen, Denmark, [15] Department of Radiation Sciences, Umeå University, Umeå, Sweden, [16] August Pi i Sunyer Biomedical Research Institute (IDIBAPS), Barcelona, Spain

Background: Loneliness is most prevalent during adolescence and late life and has been associated with mental health disorders as well as with cognitive decline during aging. Associations between longitudinal measures of loneliness and verbal episodic memory and brain structure should thus be investigated.

Methods: We sought to determine associations between loneliness and verbal episodic memory as well as loneliness and hippocampal volume trajectories across three longitudinal cohorts within the Lifebrain Consortium, including children, adolescents (N = 69, age range 10–15 at baseline examination) and older adults (N = 1468 over 60). We also explored putative loneliness correlates of cortical thinning across the entire cortical mantle.

Results: Loneliness was associated with worsening of verbal episodic memory in one cohort of older adults. Specifically, reporting medium to high levels of loneliness over time was related to significantly increased memory loss at follow-up examinations. The significance of the loneliness-memory change association was lost when eight participants were excluded after having developed dementia in any of the subsequent follow-up assessments. No significant structural brain correlates of loneliness were found, neither hippocampal volume change nor cortical thinning.

OPEN ACCESS

Edited by:
Stephanie Debette, Université de Bordeaux, France

Reviewed by:
Katya Numbers, University of New South Wales, Australia
Judith Okely, The University of Edinburgh, United Kingdom
Dorina Cadar, Brighton and Sussex Medical School, United Kingdom
Leslie Grasset, Université de Bordeaux, France

***Correspondence:**
Cristina Solé-Padullés
csolepadulles@ub.edu

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Neurocognitive Aging and Behavior, a section of the journal Frontiers in Aging Neuroscience

Received: *15 October 2021*
Accepted: *17 January 2022*
Published: *23 February 2022*

Citation:
Solé-Padullés C, Macià D, Andersson M, Stiernstedt M, Pudas S, Düzel S, Zsoldos E, Ebmeier KP, Binnewies J, Drevon CA, Brandmaier AM, Mowinckel AM, Fjell AM, Madsen KS, Baaré WFC, Lindenberger U, Nyberg L, Walhovd KB and Bartrés-Faz D (2022) No Association Between Loneliness, Episodic Memory and Hippocampal Volume Change in Young and Healthy Older Adults: A Longitudinal European Multicenter Study. Front. Aging Neurosci. 14:795764. doi: 10.3389/fnagi.2022.795764

Conclusion: In the present longitudinal European multicenter study, the association between loneliness and episodic memory was mainly driven by individuals exhibiting progressive cognitive decline, which reinforces previous findings associating loneliness with cognitive impairment and dementia.

Keywords: loneliness, episodic memory, hippocampus, cortical thickness, adolescence, cognitive decline

INTRODUCTION

Loneliness is a subjective and negative emotion related to dissatisfaction with the quantity or quality of social connections (Hawkley and Cacioppo, 2010). Previous literature states that loneliness confers increased risk of all-cause mortality as well as cardiovascular disorders (Holt-Lunstad and Smith, 2016), which might be mediated by unhealthy lifestyle or depression (Holwerda et al., 2016). Indeed, people who report feeling lonely are at higher risk of depression, and, similarly, depression reinforces feelings of loneliness (Cacioppo et al., 2006). Although loneliness and social isolation are related, with the latter describing an objective state of minimal social contact or even lack of social support (Ong et al., 2016; Yanguas et al., 2018), both entities represent independent risk factors for cognitive decline and dementia with advanced age (Holwerda et al., 2014; Kuiper et al., 2015; Sundström et al., 2020; Sutin et al., 2020). While some researchers have used both terms indistinctly, more recently, they have been differentiated, although they both appear to be common predictors of social frailty and mortality (Yanguas et al., 2018).

Associations between greater loneliness and lower cognitive functioning have been found in cross-sectional studies with advanced age (reviewed in Boss et al., 2015), as well as in a previous longitudinal study with repeated measures of loneliness and cognition (Wilson et al., 2007). But loneliness does not only affect older people, and there appears to be a U-shaped relationship, where younger and older adults present the highest prevalence (Lasgaard et al., 2016). Hence, interventions to alleviate loneliness among young people have been carefully designed (reviewed in Eccles and Qualter, 2021). Studies aiming to explore associations between loneliness and cognitive functions among young populations are lacking, and loneliness and stress are entangled in a way that the former may contribute to strengthen the acknowledged implication of stress on hippocampal neurogenesis and memory formation (see review by Kim and Diamond, 2002).

Research attempting to identify structural brain correlates of loneliness have emerged with previous studies using a Voxel-Based Morphometry (VBM) approach, emphasizing that regions linked to processing of social information, empathy and emotional regulation would be particularly compromised; namely, fronto-temporal and limbic areas in both young (Kong et al., 2015; Nakagawa et al., 2015) and older adults (Cacioppo et al., 2014; Düzel et al., 2019). These results were derived from cross-sectional studies associating loneliness and brain structure without investigating how progression of loneliness may relate to brain structure in successive evaluations. Only one previous study explored prefrontal cortical thickness and loneliness associations before and after an exercise intervention in older adults aged 60–79 (Ehlers et al., 2017), failing to find any direct associations with loneliness.

Due to its involvement in learning and memory (Squire, 1991; Burgess et al., 2002) as well as emotion regulation (Santangelo et al., 2018), the hippocampus is a key brain structure to be considered, especially in older adults. A previous cross-sectional study found an association between this structure and loneliness (Düzel et al., 2019), highlighting its role in both cognitive and social processes linked to self-perception of social isolation. A very recent study also showed that individuals reporting loneliness and social isolation presented higher brain age (de Lange et al., 2021), relative to chronological age, which is acknowledged as a marker of brain integrity and health. There have been, however, no studies examining the association between repeated assessment of loneliness and the brain, and more specifically the hippocampus. In a recent review by Campagne (2019), it was argued that loneliness is related to circulating stress hormones, immune system as well as glutamate system functioning. As stated above, loneliness is very much associated with depression, anxiety and stress (Beutel et al., 2017; Campagne, 2019). Chronic stress is acknowledged to cause activation of the hypothalamic-pituitary-adrenal (HPA) axis, which leads to elevated circulating glucocorticoids (Sapolsky, 1996). The hippocampus presents a high concentration of glucocorticoid receptors and it has been demonstrated that one of the causes associated with an accelerated damage of hippocampal neurons is a prolonged high concentration of corticoids (Sapolsky, 1996). The fact that loneliness is experienced as a feeling possibly leading to mental distress further contributes to hypothesize that loneliness might also promote a stress-related chain, with overactivation of HPA axis, and possible impact on hippocampal volume trajectories, particularly among older adults. Furthermore, this putative association might also be seen among preadolescents. Wiedenmayer et al. (2006) found associations between cortisol levels and specific portions of the hippocampus morphology among children, which were positive for anterior parts and negative for lateral portions, and loneliness has been previously linked to stress (Campagne, 2019). Due to the complexity of the developing brain at these early stages, with the acknowledged initial decrease of gray matter volume and cortical thickness (Gennatas et al., 2017), it is of relevance to explore putative associations between loneliness and hippocampus among children, as compared to old adults.

In a recent review, authors stressed that large and diverse longitudinal cohorts are needed to elucidate the neurobiology of loneliness (Lam et al., 2021). In this line, it might be relevant to explore loneliness as a long-term feeling. Likewise, longitudinal investigation of brain structural and cognitive

correlates of loneliness in younger and older adults, i.e., groups that exhibit the highest loneliness rates, are required to gain understanding of how loneliness may be associated with poor brain health. Furthermore, a longitudinal approach offers a unique opportunity to model how changes in loneliness status may relate to a particular cognitive and structural trajectory, and this is especially important when taking into account typically-developing children as well as older adults.

With longitudinal data from the European Lifebrain Consortium[1] (Walhovd et al., 2018), we aimed to explore the association between loneliness and verbal episodic memory, as well as loneliness and hippocampal (HPC) volume change across young and older participants. Furthermore, we explored possible associations between loneliness and cortical thickness across the cortical mantle with no previous hypothesis, as no preceding studies had used this technique in the study of structural correlates of loneliness. Thus, we consider this analysis to be exploratory.

MATERIALS AND METHODS

Subjects

A total of 1,537 participants were drawn from three cohorts within Lifebrain: BETULA (aged 60–85), BASE-II (aged 60–86), and HUBU (aged 10–15 at baseline). **Table 1** shows further details on cohort size, waves of assessment for loneliness, verbal episodic memory and Magnetic Resonance Imaging (MRI) of the brain for each participating cohort. **Figure 1** depicts timelines of assessments for each cohort, indicating the year of assessment, mean age at each particular time point and measures undertaken at each evaluation.

A subset of the BETULA cohort with data on loneliness, cognition and MRI was included in the study. Because this was a study with rolling recruitment, sample size increased at each wave of assessment (see **Table 1**). Exclusion criteria included severe visual or auditory impairment, intellectual or developmental disabilities, suspected dementia, having a mother tongue other than Swedish, any contraindication to MRI, neurological disorders, Mini-Mental State Examination (MMSE) < 24, brain surgery or substantial anatomical deviations (Nilsson et al., 1997; Gorbach et al., 2020). BASE-II cohort included healthy community-dwelling older adults living in the greater Berlin metropolitan area with normal or corrected to normal vision. Exclusion criteria included MMSE scores < 25, any history of psychiatric or neurological conditions or history of head injuries (Bertram et al., 2014; Gerstorf et al., 2016). Two waves of assessment with cognition and MRI were included; this latter only for a subset of 215 (see **Table 1**). A third follow-up was implemented by the GendAge study, with measures of loneliness and cognition, but not MRI, for the majority of the initial sample (Demuth et al., 2021).

The longitudinal HUBU cohort includes typically-developing children older than seven who were recruited from three elementary schools in the Copenhagen suburban area in 2007. Exclusion criteria included any known history of neurological or psychiatric disorders or significant brain injury, as reported by parents (Madsen et al., 2018). For these participants, it is important to note that they underwent four longitudinal assessments of loneliness and MRI, encompassing an age range from 13 to 18 in the last time-point of loneliness evaluation and thus covering a period from late childhood to late adolescence.

All volunteers had been drawn from studies where appropriate informed consent was obtained from themselves or their parents/legal guardians (Nilsson et al., 1997; Bertram et al., 2014; Madsen et al., 2018). In addition, local ethical approvals for data sharing were acquired for each participating site (Walhovd et al., 2018).

Loneliness Measures

Loneliness scores had been obtained for each participating cohort based on the following scales. For BETULA, the following item from the Center for Epidemiological Studies – Depression scale (CES-D; Radloff, 1977) was included: "I felt lonely in the past week," with scores ranging from 0 to 3 ('0 – rarely or less than 1 day' to '3 – most or all of the time-5–7 days'). For BASE-II, the UCLA Loneliness Scale 7 item-version was available (Russell et al., 1984). Statements such as: "I feel isolated from others," were presented and scored on a 5-point Likert scale ranging from 1 to 5 ('1 – strongly disagree' to '5 – strongly agree'). The mean of the seven items was computed as published elsewhere (Düzel et al., 2019), with larger values indicating greater loneliness. For the HUBU cohort, an item included in the Junior Eysenck Personality Questionnaire (J-EPQ, Eysenck, 1965) was the measure taken to compute loneliness among the youngest: "Do you often feel lonely?", which was scored from 0 to 3 ('0 – strongly disagree' to '3 – strongly agree').

Harmonization of Loneliness

Since loneliness for the BASE-II cohort were derived from seven Likert-based questions with a final quantitative score, as opposed to the comparable one single Likert-based question for BETULA and HUBU, harmonization was required. The technique employed to harmonize loneliness into a qualitative variable for the three cohorts was Statistical Matching (D'Orazio et al., 2006). This method entails the assumption that distribution of loneliness scores is comparable among cohorts. In this line, it was assumed that the proportion of participants feeling high, middle and low loneliness from BETULA (also old adults) would be comparable with the ones from BASE-II. BETULA and BASE-II are both cohorts from Northern Europe and the European Commission analyzed the incidence of loneliness among European from surveys administered between 2010 and 2014. Ensuing conclusions were that "the lowest share of people who feel lonely is found in the Netherlands and Denmark (3%), Finland (4%) as well as Germany, Ireland and Sweden (5%)"[2]. As a result, the assumption of equal distribution between BETULA and BASE-II could be accepted. The abovementioned percentage of 5% matched perfectly the one we obtained in the Swedish cohort of BETULA (5% of participants scored 2 or 3 in the

[1] https://www.lifebrain.uio.no/

[2] https://ec.europa.eu/jrc/en/news/how-lonely-are-europeans

TABLE 1 | Lifebrain eligible study cohorts: sample sizes and waves of assessment for each measure of loneliness, episodic memory, and magnetic resonance imaging (MRI).

Study cohort (city/country) and sample characteristics	Loneliness (*N*: sample size for each time point)	Memory (*N*: sample size for each time point)	MRI image (T1) (*N*: sample size for each time point)
BETULA (Umeå/Sweden): aged 60–85 (Nilsson et al., 1997).	Four time-points (*N* = 143, 185, 260, and 250). Interval period: 5 years.	Four time-points (*N* = 143, 185, 260, and 250). Interval period: 5 years.	Two time-points of MRI (*N* = 230 and 168). Interval period: 4 years.
BASE-II (Berlin/Germany): aged 60–86 (Bertram et al., 2014).	Three time-points (*N* = 1325, 219, and 844). Mean interval periods ranging from 2.3 to 3.23, with a mean of 5.54 years (*SD* 0.45) between first and last assessment.	Three time-points: (*N* = 1323, 218, and 749). Mean interval periods ranging from 2.3 to 3.24, with a mean of 5.57 (*SD* 0.45) between first and last assessment.	Two time-points (*N* = 215 and 215). Mean interval period: 2.29 years (*SD* 0.45)
HUBU (Copenhagen/Denmark): aged 10–15 (Madsen et al., 2018).	Four time-points (*N* = 69, 68, 59, and 39). Mean interval periods from 1.11 to 1.34, with a mean interval period of 3.43 years (*SD* 0.43) between first and last assessment.	Two time-points (*N* = 59 and 31). Mean memory interval period: 4.04 (*SD* 0.23).	Four time-points (*N* = 66, 64, 42, and 38). Mean interval periods from 1.08 to 1.35, with a mean interval period of 3.42 years (*SD* 0.45) between first and last MRI.

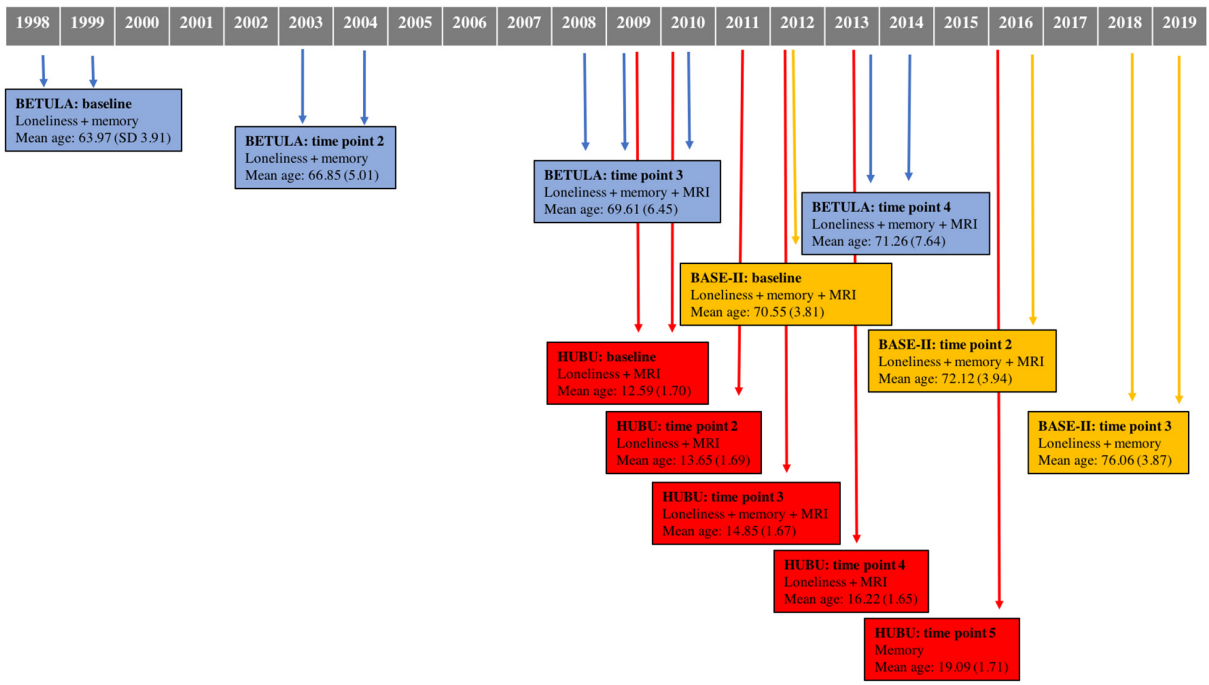

FIGURE 1 | Data collection timeline for each cohort, indicating time of assessment, measures conducted for each time point, as well as mean age and standard deviation. MRI, magnetic resonance imaging.

question "I felt lonely in the past week") and these two scores were joined into the category of 'high loneliness.' Following this, the same percentage was applied for BASE-II. Therefore, 5% of German participants with the highest scores of loneliness were classified as 'high loneliness.' The same was done for subjects that scored 1 ('medium loneliness') for BETULA, which represented 15% of the cohort. Again, the same percentage was applied for those participants from BASE-II with the highest loneliness scores, after having classified and ruled out the high loneliness group. Finally, the remaining participants were categorized as 'low loneliness,' which represented 80% of older participants. The loneliness scale used for the youth (HUBU) also included values ranging from 0 to 3, making it comparable to the categories described above. In this sample of older children and adolescents 13% of participants reported usually feeling lonely (score of 2 and 3: high loneliness) at baseline examination, while 40.7% felt moderately lonely (score of 1, classified as medium loneliness) and 46.3% were classified in the low group (score of 0).

The above classification of participants according to low, medium or high loneliness was applied for all time points. BETULA was again taken as a reference and since percentages of high, medium and low loneliness did not significantly change

across time points, we took the same proportions (5% for high, 15% for medium, and 80% for low) for matching the remaining time points of BASE-II.

Verbal Episodic Memory

All three cohorts included measures of verbal episodic memory, with words or sentences as items to be encoded and later retrieved. Specifically, a composite score of immediate free and cued recall of sentences was used for BETULA. More details on memory assessment are described elsewhere (Nilsson et al., 1997). For the BASE-II cohort, recognition accuracy from the Verbal Learning and Retention Test (Helmstaedter and Durwen, 1990), computed as hits minus false alarms, was used as a measure of episodic memory. For the HUBU cohort, the Verbal Affective Memory Test-26 (VAMT-26; Hjordt et al., 2020) was available, which allowed to obtain a composite score from total free and delayed recall of 10 positive, 10 negative, and 6 neutral words.

Depressive Symptoms and Life Events

For BETULA and BASE-II cohorts, a measure of mental health status was available for each wave, matching both loneliness and memory assessments. For the former cohort, and as mentioned above, the CES-D scale was used. This scale contains 20 items scored from 0 to 3 with (maximum score of 60) and a cut-off value of 16 has been considered to identify individuals at risk of clinical depression (Lewinsohn et al., 1997). To compute depressive symptoms, we disregarded the item of loneliness, as previously reported (CES-D Minus Loneliness: CES-DML; Cacioppo et al., 2017). Baseline mean score for this cohort was 6.78 (SD: 5.74).

For the BASE-II cohort, a baseline score was available from the 15-item version of the Geriatric Depression Scale (GDS; Yesavage, 1988). Normal scores are considered from 0 to 5. The baseline mean score for the cohort was 2.17 (SD: 1.67).

For the younger sample (HUBU), a measure of exposure to negative life events was used as covariate in the statistical models. The Child and Adolescent Survey of Experiences (CASE, Allen et al., 2012) was available for all time points except for the last one, matching all loneliness measures. Thirty-four negative life events such as "family member really sick," "been teased or bullied," "parent split up," or "parent lost job" amongst others were included. Maximum score was 34.

Magnetic Resonance Imaging and Pre-processing

As detailed in **Table 1**, a subsample of 511 participants underwent an MRI session, including a 3D structural T1-weighted scan of the whole brain acquisition. More than one thousand observations over time were considered for the longitudinal analyses.

At each site, structural images were acquired with a 3 Tesla MRI scanner, with the following parameters: (1) BETULA: Discovery GE scanner; TR: 8.2 ms, TE: 3.2 ms, TI: 450 ms, flip angle: 12°, slice thickness: 1 mm, FoV 250 mm × 250 mm, 176 slices; (2) BASE-II: Tim Trio Siemens scanner; TR: 2,500 ms, TE: 4.77 ms, TI: 1,100 ms, flip angle: 7°, slice thickness: 1 mm, FoV 256 mm × 256 mm, 176 slices, and (3) HUBU: Magnetom Trio Siemens scanner; TR: 1,550 ms, TE: 3.04 ms, TI: 800 ms, flip angle: 9°, slice thickness: 1 mm, FOV 256 mm × 256 mm, 192 slices.

Structural brain images were processed with the longitudinal processing stream available in FreeSurfer 6.0[3]. Hippocampal (HPC) and estimated total intracranial volumes (TIV) were extracted for each time-point. All images were visually inspected for quality control.

Statistical Analyses

To explore the associations between loneliness and age on the one hand, and loneliness and sex on the other hand, for each cohort separately, we first conducted general linear models including baseline loneliness as a categorical dependent variable, sex as a fixed factor and baseline age as a covariate of interest, with SPSS Statistics for Windows (IBM Corp. Released 2020. Version 27.0. Armonk, NY, United States: IBM Corp.). Subsequently, partial correlations between baseline loneliness and memory were performed for each cohort, adjusted for age and additionally for years of education in the older cohorts. Baseline loneliness and HPC volume associations were also examined, with age, sex, and total intracranial volume as covariates.

To examine associations between changes in verbal episodic memory or HPC volume and loneliness across time-points, statistical models were computed separately for each cohort as described below. Exploratory vertex-wise analyses associating loneliness and cortical thickness were additionally conducted (see below).

Loneliness and Verbal Episodic Memory/Hippocampal Volume Associations

Linear mixed-effects models run in RStudio 1.4. (RStudio Team, 2020) with the package lme4 (Bates et al., 2015), were conducted to explore associations between loneliness and verbal episodic memory and loneliness and HPC volume. We also tested a potential differential memory decline or rate of volume change for different levels of loneliness using a loneliness-by-age interaction term. The outcome variable was episodic memory or HPC volume; fixed factors included age (linear and quadratic), sex, years of education, baseline depression (for BETULA and BASE-II), categorical loneliness (as a time-varying covariate) and loneliness-by-age interaction. Change in HPC volume or change in memory was not computed prior to modeling. Interaction terms of fix effect regression with age were used to capture age-related changes. Since the age regression coefficient can be interpreted as the annual rate of change of the outcome and any interaction with it will represent an effect modifier of such expected annual rate of change of the outcome.

Within-subject random effects included intercept. We also considered adding random linear and quadratic terms to model the slope as in the fixed-effects if sufficient degrees of freedom were available (sufficient repeated measures per subject) and also a better goodness of fit was obtained (tested by comparing models with increasing number of random effects with likelihood ratio tests).

[3]https://surfer.nmr.mgh.harvard.edu/

To test hypotheses involving more than one fixed effect regression coefficient, we framed them as a comparison of nested models, using a Chi squared likelihood test for inference. Fixed effect under hypothesis usually involved loneliness and its interaction with the age linear term. Significance of single fix effect regression coefficients (betas) was estimated with two-tailed t-tests using the Satterwhite approximation for the effective degrees of freedom with R package lmerTest (Kuznetsova et al., 2017).

The HUBU cohort included two time-points with verbal episodic memory performance data only overlapped with one of the four time-points in which loneliness was assessed; hence a multiple regression model was implemented with the difference of verbal episodic memory performance as a dependent variable and baseline loneliness, age, sex and negative life events as independent variables. The model with hippocampal volume as outcome, for this younger cohort included as fixed effects: age (linear and quadratic), negative life events, sex and categorical loneliness (all variables measured at each time point).

Loneliness and Cortical Thickness Associations

The same approach using linear mixed models as described above was used to examine associations between loneliness and cortical thickness at each vertex of extracted cortical surfaces. To account for spatial correlation of the vertex-wise tests, we used Freesurfer's LME toolbox (Reuter et al., 2012; Bernal-Rusiel et al., 2013), for each cohort separately, with an FDR-corrected significance threshold of $p < 0.05$ for multiple comparisons.

RESULTS

Baseline Associations Between Loneliness and Age, and Loneliness and Sex

Table 2 lists main demographic variables from the three included Lifebrain cohorts, as well as associations between loneliness and age and loneliness and sex at baseline examination. As shown in **Table 2**, there was a positive association between age and loneliness in the BASE-II cohort, with older volunteers feeling lonelier. No associations were observed for the BETULA or HUBU cohorts.

For BETULA, a *post hoc* analysis considering all observations was carried out to capture the component of trajectories for both age and loneliness variables. By doing this, the association improved, with a trend toward increased loneliness with increasing age ($\chi^2 = 4.9$, $p = 0.08$).

Sex differences were seen within the BASE-II cohort, with males reporting significantly more feelings of loneliness than females (males: $N = 671$, mean = 1.60, $SD = 0.62$; females: $N = 654$, mean = 1.52, $SD = 0.64$). We also observed significant sex differences in loneliness in the younger cohort (HUBU), with female participants having higher scores of loneliness at baseline than males (females: $N = 41$, mean = 0.85, $SD = 0.93$; males: $N = 24$, mean = 0.33, $SD = 0.56$). No sex differences in baseline loneliness were observed for BETULA.

TABLE 2 | Baseline demographics for each participating cohort and associations with loneliness derived from an ANOVA with loneliness as the outcome variable and including sex and age in the same model.

Lifebrain Cohorts	Baseline age: mean (*SD*)	Sex (% female/ % male)	Years of education: mean (*SD*)	Baseline MMSE: mean (*SD*)	Loneliness ~ age *F-/p-value*	Loneliness ~ sex *F-/p-value*	Loneliness ~ memory *r-/p-value*	Loneliness ~ HPC volume *r-/p-value*
BETULA	63.97 (3.91)	57.34/42.66	10.06 (3.90)	28.07 (1.43)	2.47/0.12	1.09/0.29	−0.12/0.19	−0.08/0.21
BASE-II	70.55 (3.81)	50.6/49.4	14.14 (2.88)	28.64 (1.18)	5.99/0.01*	4.93/0.026** (male > female)	−0.03/0.28	−0.01/0.83
HUBU	12.59 (1.70)	62.32/37.68	–	–	0.97/0.33	5.36/0.024*** (female > male)	−0.16/0.23	−0.05/0.7

*Beta = 0.011, std. error = 0.005, partial eta squared = 0.005; **beta = −0.76, std. error = 0.034, partial eta squared = 0.004; ***beta = −0.49, std. error = 0.21, partial eta squared = 0.08. Further partial correlations with baseline episodic memory and hippocampal volume (HPC) are also shown. MMSE, Mini-Mental State Examination.*

Baseline Associations of Loneliness With Verbal Episodic Memory, Hippocampal Volume and Cortical Thickness

While at baseline, all correlations between loneliness and episodic memory, as well as loneliness and HPC volume were negative for all participating cohorts, they were not statistically significant (see **Table 2**). Likewise, in the baseline, i.e., cross-sectional, analysis of associations between loneliness and cortical thickness across the cortical mantle, none of the vertices survived FDR correction. Uncorrected results for baseline loneliness and thickness associations are depicted in the **Supplementary Figure 1**.

Longitudinal Loneliness and Verbal Episodic Memory Change

We observed a significant effect of loneliness on verbal episodic memory for BETULA (χ^2 = 11.25, df = 4, p = 0.02). Such effect was clearly driven by the loneliness-by-age interaction effect with a higher degree of loneliness being associated with poorer memory performance, only at advanced ages (**Figure 2A** and **Supplementary Table 1**). This result entailed that with advancing age, having medium to high levels of loneliness was associated with decreased memory slope. Therefore, if one subject felt lonely at baseline and these feelings of loneliness were stable across time, their memory would be expected to decrease more than one subject never feeling lonely at all. Likewise, the model also predicted that if loneliness levels decreased at some point (being high at baseline and low at posterior evaluations), then memory levels would 'normalize' or the subject would 'jump' from the steeper decreased memory slope (**Figure 2A**, blue or green slopes) to a 'normal' slope of memory decrease, allegedly expected by age (**Figure 2A**, red slope). Therefore, the association between loneliness and memory progress would mainly be seen for those subjects with persistent feelings of loneliness. It is, however, important to note that only medium loneliness reached significant results (see **Supplementary Table 1**), even though the decline in episodic memory can also be seen among subjects categorized with high loneliness. It is likely that these observations may reflect fewer numbers of participants classified as high loneliness as the point estimate of the loneliness-by-age interaction term was of the same order and in the same direction for both loneliness levels.

No evidence for such effects were seen in BASE-II (χ^2 = 1.3, df = 4, p = 0.86, see **Figure 2B** and **Supplementary Table 2**). Likewise, using the original measure of loneliness for this cohort and its interaction term with age did not improve the model significantly (χ^2 = 1.20, df = 2, p = 0.57).

Analyses were repeated for both BETULA and BASE-II after having ruled out eight and one subject, respectively, who had developed dementia in any of the subsequent waves of assessments. For BETULA, the previous association found between memory and loneliness-by-age interaction was not significant anymore: χ^2 = 6.62, df = 4, p = 0.16. For BASE-II results remained minimally changed: χ^2 = 1.18, df = 4, p = 0.56.

Regression model for HUBU did not yield any significant result (F = 0.87, df = 4, p = 0.50). **Supplementary Tables 1–3** provide estimates for the fixed effects of the association between verbal episodic memory and loneliness for all cohorts.

Longitudinal Loneliness and Hippocampal Volume Change

Rates of annual HPC volume loss from age 60 onwards are depicted in **Figure 3** (at age 60 estimated 0.2% for BETULA and 0.4% for BASE-II; at age 80 these were increased to 0.47 and 0.75% for each cohort, respectively). No association between loneliness and HPC volume, as well as loneliness and HPC volume change (loneliness-by-age interaction), where intercept value (0) corresponds to age 60, was found for BETULA (χ^2 = 4.27, df = 4, p = 0.36; **Figure 3A**) and BASE-II (χ^2 = 2.31, df = 4, p = 0.68; **Figure 3B**). Likewise, we did not find any associations between loneliness and HPC volume changes in the younger cohort (HUBU: χ^2 = 2.33, df = 4, p = 0.68). As seen in **Figure 3C**, for this latter cohort, HPC volume increased from age 10 to 14/15 and was followed by stabilization, emulating the non-linear developmental pattern of increased volume during late childhood and early adolescence accompanied by a slight subsequent deceleration, as described elsewhere (Tamnes et al., 2018). **Supplementary Tables 4–6** provide estimates for the fixed effects of the association between hippocampal volume and loneliness, as well as loneliness-by-age interaction for BETULA, BASE-II and HUBU cohorts.

Again, re-analysis disregarding the eight participants who developed dementia along the study (8 for BETULA) yielded no significant changes on the abovementioned statistical outputs: BETULA: χ^2 = 4.30, df = 4, p = 0.37. For BASE-II, the participant with onset of dementia at time point 2, was missing at that particular time point, and did not have MRI data on time point 3, so analyses were not repeated.

Longitudinal Loneliness and Cortical Thickness Change

Longitudinal analyses of associations between loneliness and cortical thickness across the cortical mantle did not yield any statistically significant vertex-wise results after correction for multiple comparisons (FDR < 5%), in any of the cohorts. Uncorrected results are further described in the **Supplementary Figure 1**.

DISCUSSION

To our knowledge, this is the first multicenter European study to incorporate measures of loneliness over time to explore its cognitive and structural brain correlates in distinct samples across the lifespan, covering periods of late childhood, adolescence, and older adulthood.

Significant associations between age and loneliness were observed within the German (BASE-II) cohort, with older participants reporting increased feelings of loneliness; a result that is in line with the acknowledged increase of loneliness in late life, a period where one may experience a number of losses, from the death of a spouse or friends to social disengagement (Singh and Misra, 2009). Despite no baseline age and loneliness

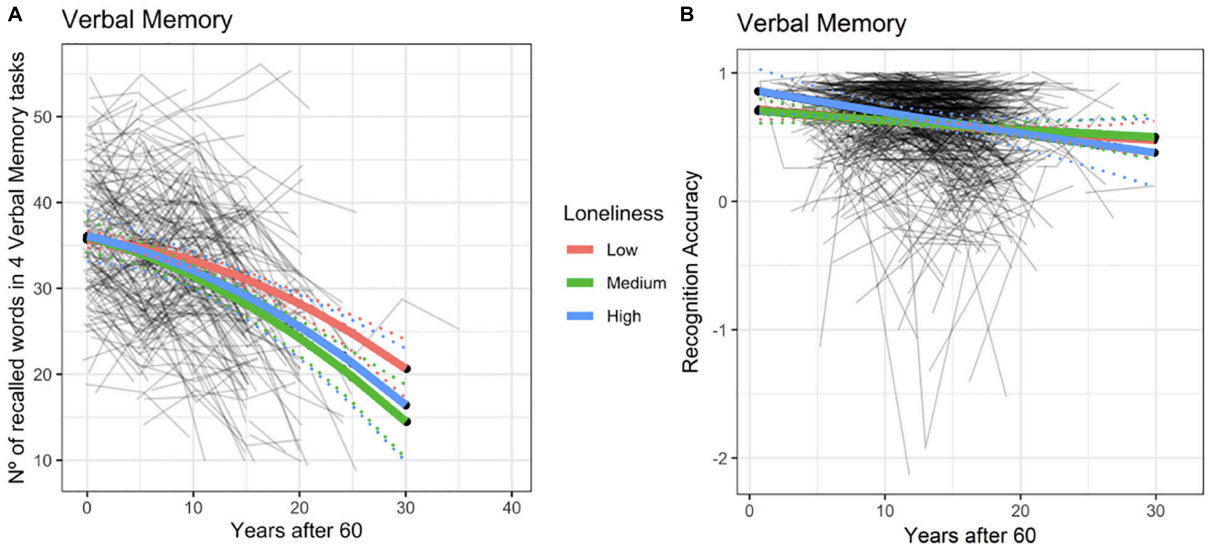

FIGURE 2 | Episodic memory predicted trajectories versus age after 60 for different loneliness levels for Betula **(A)** and BASE-II **(B)**. Association between loneliness and episodic memory change within the BETULA cohort: no effect was seen at age 60 (intercept), but from then onwards, medium and high loneliness scores were related to poorer memory performance over time. Dotted lines show 95% confidence interval for predictions with random effects set to 0.

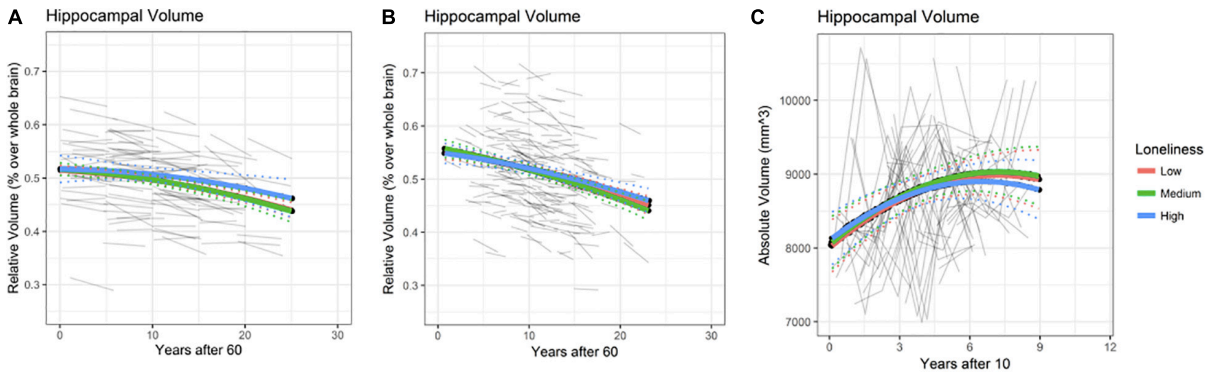

FIGURE 3 | Hippocampal (HPC) volume predicted trajectories versus age after 60 for different loneliness levels for BETULA **(A)** and BASE-II **(B)**. **(C)** Depicts hippocampal trajectories for the younger participants (HUBU cohort) from age 10 onwards, as regards the loneliness variable. Raw hippocampal volumes are shown for HUBU, as absolute values might be more convenient in longitudinal studies of brain development (Goddings et al., 2014). Dotted lines show 95% confidence interval for predictions with random effects set to 0.

associations were seen for the BETULA cohort, a trend toward more loneliness with increasing age was seen after taking into account all observations.

We also found that German male participants presented more feelings of loneliness than their female counterparts. Notably, these participants were born between 1927 and 1953, a generation of working men and 'stay-at-home' women. Therefore, men were more likely to experience the life-changing event of retiring from work, and this may have intensified their feelings of loneliness, as compared to women. Yet, the above associations were not seen for the Swedish cohort of older adults (BETULA). This might be partially explained by the fact that the German cohort included more than one thousand participants at baseline examination, compared to the 143 volunteers for the BETULA cohort. This larger statistical power may have favored the emergence of significant loneliness-age and loneliness-sex associations within BASE-II. It was somehow unexpected to find more feelings of loneliness among German men, as compared to their female counterparts, because older women usually report increased loneliness (Beal, 2006; Jarach et al., 2021). Despite this, it has been recently argued that loneliness is more associated with health, functional limitations and depression (Jarach et al., 2021), than with social isolation itself, and we have not controlled for physical health variables in the present study, neither have we taken into account other generational factors (alleged 'working men' versus 'stay-at-home women'), which may have contributed to the abovementioned associations found in BASE-II.

On the other hand, in the younger cohort, females reported feeling lonelier than males, a finding that is in line with a recent study, also with Danish adolescents (Eccles et al., 2020). Authors

pointed to the possibility of a sex-based stigma when reporting feelings of loneliness among male adolescents. No sex differences were seen for the BETULA cohort, but bearing in mind the few proportion of people feeling lonely and the smaller sample size as compared to BASE-II, it is plausible that sex differences, if any, would have emerged with a more representative sample of lonely people.

Regarding associations between loneliness and episodic memory over time, we found that for BETULA, those participants reporting medium to high levels of loneliness over time displayed more memory loss at follow-up examinations, even though memory performance was not associated with loneliness at baseline. Thus, self-perceived and constant feelings of loneliness among these participants co-occurred with a decrease in memory performance for this cohort of Swedish participants. It is worth mentioning, however, that the model was conducted in a way that loneliness at one time point predicted memory for that particular time point, without taking into account previous loneliness levels. Despite this, interpretations considering distinct trajectories of loneliness are possible and this model predicted that low baseline loneliness for BETULA participants, combined with increases of loneliness feelings in any subsequent evaluation would be equally associated with steeper episodic memory decline, for that particular time point. Therefore, persistent loneliness but also increases in loneliness perceptions, were associated with episodic memory loss among this cohort of participants. In a previous review discussing associations between loneliness and global cognitive function as well as episodic memory, Boss et al. (2015) suggested that some associations may disappear after adjusting for demographic and psychosocial factors that influence loneliness. In our study, the association between loneliness and memory performance over time among BETULA participants was seen while controlling for age, education, sex and depressive symptoms, which emphasizes the relevance of loneliness for memory maintenance in aging and reinforces the notion that loneliness can be considered a different entity from depression, as acknowledged also in previous studies examining loneliness associations with cognitive decline (Wilson et al., 2007). Nonetheless, the association was lost after having ruled out the eight cases of dementia with onset in any of the subsequent follow-up assessments. This fact reinforced previous evidence suggesting that there is an association between loneliness and cognitive impairment (Wilson et al., 2007; Sundström et al., 2020; Sutin et al., 2020). Notably, the fact that the association did not remain significant does not imply that there is no association between loneliness and memory decline, but rather that if cases of significant memory decline are excluded, then this association is critically diminished. It is important to point out the possibility of reverse causality. Loneliness may cause memory decline but memory decline might also contribute to increased feelings of loneliness. Previous data supported the notion of a bidirectional effect between cognitive ability and loneliness among older adults (Zhong et al., 2017; Okely and Deary, 2019).

As regards the lack of association between loneliness measures and memory change in the BASE-II cohort, even when using the original continuous measure of loneliness, we believe that two main factors could be contemplated. First, the verbal memory test (recognition accuracy) was not a comparable measure with the recall (free and cued) that was used for BETULA. While recall has been consistently associated with hippocampus (Aggleton and Shaw, 1996), the role of this structure on recognition has been more debated. Recognition can depend on recollection, familiarity or both processes (Yonelinas, 2002), and though some believe that the hippocampus exclusively subserves recollection, more recent data is available indicating that both recollection and familiarity are majorly supported by the hippocampus (Merkow et al., 2015). Yet, memory measures used for the two cohorts of older adults may not be completely comparable. Second, it is possible that the considerable number of drop-outs, particularly at the second wave of assessment, may have constrained our analyses. After a thorough analysis on missing data for BASE-II cohort, it was seen that a greater proportion of missing participants at time point 3, presented increased baseline loneliness (26.82% of missing volunteers at time point 3 had medium to high scores of loneliness at baseline versus 20.74% of not missing participants; χ^2: 6.58, $p = 0.037$). Likewise, they were also older at baseline examination [mean age: 71.24 (*SD*: 4.01) versus 70.16 (3.64), $t = 5.02$, $p < 0.0001$]. On the whole, if older and more lonely participants dropped out from the study, this could represent a 'missing not at random' (MNAR) case, where it is likely that these participants were experiencing some kind of cognitive impairment. Assuming that the drop-outs could contain a great proportion of cognitively impaired patients and bearing in mind that the association between loneliness and memory decline was stronger in BETULA when the eight cases of dementia were included, it is reasonable to think that the MNAR could be masking some associations.

We did not find any associations between loneliness and difference in memory performance for the HUBU cohort. While loneliness was measured at four time points, memory testing was only administered in two waves of assessments, and only one of them co-occurred with the loneliness measure. An increased number of cognitive measurements would have been ideal to better grasp memory trajectories and possibly capture putative associations with loneliness trajectories.

Contrary to expectations, we did not observe any statistically significant associations between loneliness and HPC volume in any of the cohorts, neither at baseline nor longitudinally. The same applied for our explorative investigations for associations between loneliness and vertex-wise cortical thickness. Our null findings are in apparent contrast with previous reports of cross-sectional loneliness-brain structure associations in healthy young and older adults (Kanai et al., 2012; Kong et al., 2015; Düzel et al., 2019). Studies with young adults pointed to compromised left frontotemporal regions (Kanai et al., 2012; Kong et al., 2015), whereas Düzel et al. (2019) found decreased volume of subcortical regions and cerebellum among lonely old adults. Although more studies are required to further investigate a putative cortical versus subcortical association of loneliness across aging groups, it is important to note that risk factors for loneliness are not comparable across life stages. For instance, a recent study with Finnish adolescents and young adults reveled that this negative feeling is related to social transitions and expectations, group differences, former

destructive experiences or negative self-image among others (Sundqvist and Hemberg, 2021); while for middle-aged and older adults the most important factors contributing to loneliness are the loss of a spouse, frequency of contact with significant friends or family and the number of voluntary groups one is engaged to Cacioppo et al. (2015). It is also important to note that while the abovementioned studies focusing on cross-sectional associations between loneliness and brain structure used Voxel-Based Morphometry (VBM), we had a dual and different focus with a longitudinal approach: first, the hypothesis driven approach targeting the hippocampus as a subcortical structure related to emotional processing and episodic memory; and second, the exploratory approach of vertex-wise cortical thickness. Thus, the association between loneliness and memory decline in one cohort and lack of structural brain correlates suggests that structural correlates of loneliness may either appear later than cognitive correlates, or that the techniques employed were not sensitive enough to capture structural changes that were possibly too small to be detected with the current sample size.

Cacioppo et al. (2017) suggested that focus of attention is changed among people with long-term feelings of loneliness, and, though speculative, we believe that the result obtained for BETULA regarding the association between loneliness and poorer memory over time, might be partially explained by decreased attentional resources that might in turn affect encoding and retrieval of memories, overall resulting in poorer performance over time, even in the absence of brain structural changes. However, we cannot rule out the possibility that factors interacting with loneliness such as personality traits, social network and empathy in both young (Kanai et al., 2012; Kong et al., 2015) and older adults (von Soest et al., 2020), which have not been considered here, might account for lack of results in the other cohorts or the absence of structural correlates of loneliness among our participants.

The link between loneliness and cognitive decline has been previously established (Holwerda et al., 2014), although the nature of this association is poorly understood (Boss et al., 2015). Likewise, loneliness has a notable association with global health with a complexity that is not yet fully grasped (Yanguas et al., 2018). As already stated, a variety of mediating intrinsic and extrinsic factors interact with one another to converge in a combination of social vulnerability and frailty that might lead to a particular feeling of loneliness. More recently, it has been pointed out that the complexity of associations between loneliness and adverse health outcomes depend on a combination of interlinked genetic, social behavioral, physical and socioeconomic factors (de Lange et al., 2021).

Another issue that should be noted is the fact that only a low proportion of older participants from our study reported high levels of loneliness. While this is comparable with previous studies (Victor and Yang, 2012), it may have constrained our analyses. Differences in loneliness across European countries have been described and the European Commission's Joint Research Centre concluded from surveys conducted in 2010, 2012, and 2014 that Eastern and Southern Europeans feel lonelier and are more socially isolated than Western and Northern Europeans (D'Hombres et al., 2018). Notably, Northern Europeans appear to tolerate greater rates of social isolation without having an impact on subjective measures of loneliness. Greater satisfaction with the social network has also been reported in the Scandinavian countries (Sundström et al., 2009). The three cohorts included in the present study come from Sweden, Germany, and Denmark, and this may account for the fact that only a very low percentage of our participants were categorized as experiencing 'high loneliness.' Another explanation would be the fact that depression was an exclusion criterion in the present study, and even though depression is considered a different entity from loneliness, they are mutually reinforcing (Cacioppo et al., 2006).

Taken together, abovementioned data may suggest different associations of loneliness and brain health across European regions and this study only included cohorts from Northern Europe. To understand differences between European regions it is important to note that the association between well-being and social network is complex and may include not only quantity and quality of relationships with family and friends, but also perceived social support (Berkman et al., 2000), as well as access to health resources, which may vary across European countries.

Several limitations should be considered in addition to the low number of participants with high feelings of loneliness already mentioned. Differences among cohorts, particularly, old adults should be also considered. As shown in **Figure 1**, there is a difference in age to consider, with BETULA participants being slightly younger than BASE-II's at baseline examination. It is possible, in accordance with previous reports associating loneliness with increased risk of dementia (Sutin et al., 2020), that these associations are more evident in earlier stages of cognitive decline, which would favor BETULA's association and not BASE-II's. There was also a considerable time lag for loneliness evaluation for BETULA and BASE-II. In this line, BETULA started assessments in 1998, a period with no influence of digital technology, and were extended until 2014. As for BASE-II, first assessment took place in 2011. The way that one interacts with other people has considerably changed these past 15 years even though this may have mainly affected young people, for whom the use of social media has become an important part of their daily lives (Ryan et al., 2017). However, we cannot rule out the possibility that the feelings of loneliness measured during the last decade have been somehow influenced by digital revolution also among old adults. Further, heterogeneity of measures of episodic memory and loneliness should also be taken into account, as mentioned above. While BETULA's memory measures included a composite of verbal free and cued recall, BASE-II's was based on a recognition memory test. Though normality was not met for this subtest at baseline, this measure presented the best goodness of fit with age as a linear effect with the expected negative slope of memory decrease with time as compared to other memory measures such as delayed recall. It is, however, possible that this measure of memory was not sensitive enough to detect a clear memory change, as expected for age. Additionally, it would have been ideal to focus on other cognitive domains other than memory, particularly executive functions, with specific focus on attention, which may be influenced by stress-related processes (Girotti et al., 2018). As regards loneliness, measures were based on a personality or depression scale item for two cohorts, whereas the remaining cohort included a more quantitative measure, with

a short form of the UCLA-loneliness scale. However, these are all accepted measures of loneliness (Yanguas et al., 2018) and were accordingly harmonized. Also, interval periods between evaluations, particularly for MRI, might be too short for BASE-II and HUBU cohorts, and this may have also been a disadvantage when exploring structural brain changes. Finally, two main points should be considered for the youngest cohort. First, while loneliness and MRI were conducted periodically at four time points, only two assessments of memory were available. Second, sample size for HUBU was not comparable to the older adults', a fact that may have limited representativeness.

CONCLUSION

We found associations between longitudinal measures of loneliness and verbal episodic memory change within one of two cohorts of healthy older adults, which was dependent on participants' cognitive status. After excluding dementia cases, the previous association was no longer significant, strengthening previous findings associating loneliness with cognitive impairment and dementia. This study suggests that the association between loneliness and memory decline might be independent of hippocampal volume change or changes in cortical thickness. Likewise, while incidence of loneliness was increased for the younger cohort, no correlates of memory or brain structure were evidenced. We believe this is a first step toward other longitudinal approaches examining both cognitive and structural brain correlates of loneliness among healthy individuals at different life stages. Forthcoming studies would benefit from including other European countries in an ideal design able to identify 'stable' versus 'changing' feelings of loneliness and with the possibility to examine further cultural differences of self-perceived social isolation, resilience, and association of loneliness with brain health.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: data may be made available upon reasonable request, given appropriate ethical and data protection approvals. Requests to access these datasets should be directed to https://www.lifebrain.uio.no/about/contact.

ETHICS STATEMENT

Local ethical approvals for data sharing were acquired for each participating site (Walhovd et al., 2018). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

CS-P contributed to conceptualization, data curation, formal analysis, and writing (original draft). DM and MA contributed to software, formal analysis, and visualization. MS, SP, and SD contributed to resources and data curation. EZ, KE, and WB contributed to conceptualization and writing (review and editing). JB and CD contributed to writing (review and editing). AB and AF contributed to validation and writing (review and editing). AM contributed to resources and software. KM contributed to data curation and writing (review and editing). UL contributed to resources. LN contributed to validation, resources, and writing (review and editing). KW and DB-F contributed to conceptualization, supervision, writing (review and editing), and project administration. All authors contributed to the article and approved the submitted version.

FUNDING

The Lifebrain project was funded by the EU Horizon, 2020 Grant: "Healthy minds 0–100 years: Optimizing the use of European brain imaging cohorts ('Lifebrain')." Grant agreement number: 732592 (Lifebrain). Call: Societal challenges: Health, demographic change and well-being. Some of the Freesurfer analyses were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at HPC2N in Umeå, partially funded by the Swedish Research Council through grant agreement no. 2018-05973. DB-F was awarded by ICREA, 2019 Academia, a program supported by the Catalan Government. DM was partially funded by Spanish Ministry of Science and Innovation and State Research Agency through the "Centro de Excelencia Severo Ochoa 2019–2023" Program (CEX2018-000806-S), and support from the Generalitat de Catalunya through the CERCA Program.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2022.795764/full#supplementary-material

REFERENCES

Aggleton, J. P., and Shaw, C. (1996). Amnesia and recognition memory: a re-analysis of psychometric data. *Neuropsychologia* 34, 51–62. doi: 10.1016/0028-3932(95)00150-6

Allen, J. L., Rapee, R. M., and Sandberg, S. (2012). Assessment of maternally reported life events in children and adolescents: A comparison of interview and checklist methods. *J. Psychopathol. Behav. Assess.* 34, 204–215. doi: 10.1007/s10862-011-9270-5

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Beal, C. (2006). Loneliness in older women: a review of the literature. *Issues Ment. Health Nurs.* 27, 795–813. doi: 10.1080/01612840600781196

Berkman, L. F., Glass, T., Brissette, I., and Seeman, T. E. (2000). From social integration to health: Durkheim in the new millennium. *Soc. Sci. Med.* 51, 843–857. doi: 10.1016/s0277-9536(00)00065-4

Bernal-Rusiel, J. L., Reuter, M., Greve, D. N., Fischl, B., and Sabuncu, M. R. (2013). Spatiotemporal linear mixed effects modeling for the mass-univariate analysis of longitudinal neuroimage data. *Neuroimage* 81, 358–370. doi: 10.1016/j.neuroimage.2013.05.049

Bertram, L., Böckenhoff, A., Demuth, I., Düzel, S., Eckardt, R., Li, S. C., et al. (2014). Cohort profile: the Berlin Aging Study II (BASE-II). *Int. J. Epidemiol.* 43, 703–712. doi: 10.1093/ije/dyt018

Beutel, M. E., Klein, E. M., Brähler, E., Reiner, I., Jünger, C., Michal, M., et al. (2017). Loneliness in the general population: prevalence, determinants and relations to mental health. *BMC Psychiatry* 17:97. doi: 10.1186/s12888-017-1262-x

Boss, L., Kang, D. H., and Branson, S. (2015). Loneliness and cognitive function in the older adult: a systematic review. *Int. Psychogeriatr.* 27, 541–553. doi: 10.1017/S1041610214002749

Burgess, N., Maguire, E. A., and O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron* 35, 625–641. doi: 10.1016/s0896-6273(02)00830-9

Cacioppo, J. T., Chen, H. Y., and Cacioppo, S. (2017). Reciprocal Influences Between Loneliness and Self-Centeredness: A Cross-Lagged Panel Analysis in a Population-Based Sample of African American, Hispanic, and Caucasian Adults. *Pers. Soc. Psychol. Bull.* 43, 1125–1135. doi: 10.1177/0146167217705120

Cacioppo, J. T., Hughes, M. E., Waite, L. J., Hawkley, L. C., and Thisted, R. A. (2006). Loneliness as a specific risk factor for depressive symptoms: cross-sectional and longitudinal analyses. *Psychol. Aging* 21, 140–151. doi: 10.1037/0882-7974.21.1.140

Cacioppo, S., Capitanio, J. P., and Cacioppo, J. T. (2014). Toward a neurology of loneliness. *Psychol. Bull.* 140, 1464–1504. doi: 10.1037/a0037618

Cacioppo, S., Grippo, A. J., London, S., Goossens, L., and Cacioppo, J. T. (2015). Loneliness: clinical import and interventions. *Perspect. Psychol. Sci.* 10, 238–249. doi: 10.1177/1745691615570616

Campagne, D. M. (2019). Stress and perceived social isolation (loneliness). *Arch. Gerontol. Geriatr.* 82, 192–199. doi: 10.1016/j.archger.2019.02.007

D'Orazio, M., Di Zio, M., and Scanu, M. (2006). Statistical matching for categorical data: Displaying uncertainty and using logical constraints. *J. Off. Stat.* 22, 137–157.

de Lange, A. G., Kaufmann, T., Quintana, D. S., Winterton, A., Andreassen, O. A., Westlye, L. T., et al. (2021). Prominent health problems, socioeconomic deprivation, and higher brain age in lonely and isolated individuals: A population-based study. *Behav. Brain Res.* 24:113510. doi: 10.1016/j.bbr.2021.113510

Demuth, I., Banszerus, V., Drewelies, J., Düzel, S., Seeland, U., Spira, D., et al. (2021). Cohort profile: follow-up of a Berlin Aging Study II (BASE-II) subsample as part of the GendAge study. *BMJ Open* 11:e045576. doi: 10.1136/bmjopen-2020-045576

D'Hombres, B., Sylke, S., Barjaková, M., Mendonca, F. T. (2018). *Data from: European Commission. Loneliness-an unequally shared burden in Europe*. Belgium: European Commission. doi: 10.13140/RG.2.2.21745.33128

Düzel, S., Drewelies, J., Gerstorf, D., Demuth, I., Steinhagen-Thiessen, E., Lindenberger, U., et al. (2019). Structural Brain Correlates of Loneliness among Older Adults. *Sci. Rep.* 9:13569. doi: 10.1038/s41598-019-49888-2

Eccles, A. M., and Qualter, P. (2021). Review: Alleviating loneliness in young people - a meta-analysis of interventions. *Child Adolesc. Ment. Health* 26, 17–33. doi: 10.1111/camh.12389

Eccles, A. M., Qualter, P., Madsen, K. R., and Holstein, B. E. (2020). Loneliness in the lives of Danish adolescents: Associations with health and sleep. *Scand. J. Public Health* 48, 877–887. doi: 10.1177/1403494819865429

Ehlers, D. K., Daugherty, A. M., Burzynska, A. Z., Fanning, J., Awick, E. A., Chaddock-Heyman, L., et al. (2017). Regional brain volumes moderate, but do not mediate, the effects of group-based exercise training on reductions in loneliness in older adults. *Front. Aging Neurosci.* 9, 110. doi: 10.3389/fnagi.2017.00110

Eysenck, S. B. (1965). *Junior Eysenck Personality Inventory*. London: Hodder and Stoughton.

Gennatas, E. D., Avants, B. B., Wolf, D. H., Satterthwaite, T. D., Ruparel, K., Ciric, R., et al. (2017). Age-Related Effects and Sex Differences in Gray Matter Density, Volume, Mass, and Cortical Thickness from Childhood to Young Adulthood. *J. Neurosci.* 37, 5065–5073. doi: 10.1523/JNEUROSCI.3550-16.2017

Gerstorf, D., Bertram, L., Lindenberger, U., Pawelec, G., Demuth, I., Steinhagen-Thiessen, E., et al. (2016). Editorial. *Gerontology* 62, 311–315. doi: 10.1159/000441495

Girotti, M., Adler, S. M., Bulin, S. E., Fucich, E. A., Paredes, D., and Morilak, D. A. (2018). Prefrontal cortex executive processes affected by stress in health and disease. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 85, 161–179. doi: 10.1016/j.pnpbp.2017.07.004

Goddings, A. L., Mills, K. L., Clasen, L. S., Giedd, J. N., Viner, R. M., and Blakemore, S. J. (2014). The influence of puberty on subcortical brain development. *NeuroImage* 88, 242–251. doi: 10.1016/j.neuroimage.2013.09.073

Gorbach, T., Pudas, S., Bartrés-Faz, D., Brandmaier, A. M., Düzel, S., Henson, R. N., et al. (2020). Longitudinal association between hippocampus atrophy and episodic-memory decline in non-demented *APOE* ε4 carriers. *Alzheimers Dement.* 12:e12110. doi: 10.1002/dad2.12110

Hawkley, L. C., and Cacioppo, J. T. (2010). Loneliness matters: a theoretical and empirical review of consequences and mechanisms. *Ann. Behav. Med.* 40, 218–227. doi: 10.1007/s12160-010-9210-8

Helmstaedter, C., and Durwen, H. F. (1990). The Verbal Learning and Retention Test. A useful and differentiated tool in evaluating verbal memory performance. *Schweiz. Arch. Neurol. Psychiatr.* 141, 21–30.

Hjordt, L. V., Ozenne, B., Armand, S., Dam, V. H., Jensen, C. G., Köhler-Forsberg, K., et al. (2020). Psychometric Properties of the Verbal Affective Memory Test-26 and Evaluation of Affective Biases in Major Depressive Disorder. *Front. Psychol.* 5:961. doi: 10.3389/fpsyg.2020.00961

Holt-Lunstad, J., and Smith, T. B. (2016). Loneliness and social isolation as risk factors for CVD: implications for evidence-based patient care and scientific inquiry. *Heart* 102, 987–989. doi: 10.1136/heartjnl-2015-309242

Holwerda, T. J., Deeg, D. J., Beekman, A. T., van Tilburg, T. G., Stek, M. L., Jonker, C., et al. (2014). Feelings of loneliness, but not social isolation, predict dementia onset: results from the Amsterdam Study of the Elderly (AMSTEL). *J. Neurol. Neurosurg. Psychiatry* 85, 135–142. doi: 10.1136/jnnp-2012-302755

Holwerda, T. J., van Tilburg, T. G., Deeg, D. J., Schutter, N., Van, R., Dekker, J., et al. (2016). Impact of loneliness and depression on mortality: results from the Longitudinal Ageing Study Amsterdam. *Br. J. Psychiatry* 209, 127–134. doi: 10.1192/bjp.bp.115.168005

Jarach, C. M., Tettamanti, M., Nobili, A., and D'avanzo, B. (2021). Social isolation and loneliness as related to progression and reversion of frailty in the Survey of Health Aging Retirement in Europe (SHARE). *Age Ageing* 50, 258–262. doi: 10.1093/ageing/afaa168

Kanai, R., Bahrami, B., Duchaine, B., Janik, A., Banissy, M. J., and Rees, G. (2012). Brain structure links loneliness to social perception. *Curr. Biol.* 22, 1975–1979. doi: 10.1016/j.cub.2012.08.045

Kim, J. J., and Diamond, D. M. (2002). The stressed hippocampus, synaptic plasticity and lost memories. *Nat. Rev. Neurosci.* 3, 453–462. doi: 10.1038/nrn849

Kong, X., Wei, D., Li, W., Cun, L., Xue, S., Zhang, Q., et al. (2015). Neuroticism and extraversion mediate the association between loneliness and the dorsolateral prefrontal cortex. *Exp. Brain Res.* 233, 157–164. doi: 10.1007/s00221-014-4097-4

Kuiper, J. S., Zuidersma, M., Oude Voshaar, R. C., Zuidema, S. U., van den Heuvel, E. R., Stolk, R. P., et al. (2015). Social relationships and risk of dementia: A systematic review and meta-analysis of longitudinal cohort studies. *Ageing Res. Rev.* 22, 39–57. doi: 10.1016/j.arr.2015.04.006

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *J. Stat. Softw.* 82, 1–26. doi: 10.18637/jss.v082.i13

Lam, J. A., Murray, E. R., Yu, K. E., Ramsey, M., Nguyen, T. T., Mishra, J., et al. (2021). Neurobiology of loneliness: a systematic review. *Neuropsychopharmacology* 6, 1–15. doi: 10.1038/s41386-021-01058-7

Lasgaard, M., Friis, K., and Shevlin, M. (2016). "Where are all the lonely people?" a population- based study of high-risk groups across the life-span. *Soc. Psychiatry Psychiatr. Epidemiol.* 51, 1373–1384. doi: 10.1007/s00127-016-1279-3

Lewinsohn, P. M., Seeley, J. R., Roberts, R. E., and Allen, N. B. (1997). Center for Epidemiologic Studies Depression Scale (CES-D) as a screening instrument for depression among community-residing older adults. *Psychol. Aging* 12, 277–287. doi: 10.1037//0882-7974.12.2.277

Madsen, K. S., Jernigan, T. L., Vestergaard, M., Mortensen, E. L., and Baaré, W. F. C. (2018). Neuroticism is linked to microstructural left-right asymmetry of fronto-limbic fibre tracts in adolescents with opposite effects in boys and girls. *Neuropsychologia* 114, 1–10. doi: 10.1016/j.neuropsychologia.2018.04.010

Merkow, M. B., Burke, J. F., and Kahana, M. J. (2015). The human hippocampus contributes to both the recollection and familiarity components of recognition memory. *Proc. Natl. Acad. Sci. U. S. A.* 112, 14378–14383. doi: 10.1073/pnas.1513145112

Nakagawa, S., Takeuchi, H., Taki, Y., Nouchi, R., Sekiguchi, A., Kotozaki, Y., et al. (2015). White matter structures associated with loneliness in young adults. *Sci. Rep.* 20:17001. doi: 10.1038/srep17001

Nilsson, L. G., Bäckman, L., Erngrund, K., Nyberg, L., Adolfsson, R., Bucht, G., et al. (1997). The BETULA prospective cohort study: memory, health, and aging. *Aging Neuropsychol. Cogn.* 4, 1–32. doi: 10.1080/13825589708256633

Okely, J. A., and Deary, I. J. (2019). Longitudinal Associations Between Loneliness and Cognitive Ability in the Lothian Birth Cohort 1936. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 74, 1376–1386. doi: 10.1093/geronb/gby086

Ong, A. D., Uchino, B. N., and Wethington, E. (2016). Loneliness and Health in Older Adults: A Mini-Review and Synthesis. *Gerontology* 62, 443–449. doi: 10.1159/000441651

Radloff, L. S. (1977). The CES-D scale: A self report depression scale for research in the general population. *Appl. Psychol. Meas.* 1, 385–401.

Reuter, M., Schmansky, N. J., Rosas, H. D., and Fischl, B. (2012). Within-subject template estimation for unbiased longitudinal image analysis. *Neuroimage* 61, 1402–1418. doi: 10.1016/j.neuroimage.2012.02.084

RStudio Team (2020). *RStudio: Integrated Development for R*. Boston: RStudio

Russell, D., Cutrona, C. E., Rose, J., and Yurko, K. (1984). Social and emotional loneliness: an examination of Weiss's typology of loneliness. *J. Pers. Soc. Psychol.* 46, 1313–1321. doi: 10.1037//0022-3514.46.6.1313

Ryan, T., Allen, K., Gray, D., and McInerney, D. (2017). How Social Are Social Media? A Review of Online Social Behaviour and Connectedness. *J. Relationsh. Res.* 8:E8. doi: 10.1017/jrr.2017.13

Santangelo, V., Cavallina, C., Colucci, P., Santori, A., Macri, S., Mcgaugh, J. L., et al. (2018). Enhanced brain activity associated with memory access in highly superior autobiographical memory. *Proc. Natl. Acad. Sci. U. S. A.* 115, 7795–7800. doi: 10.1073/pnas.1802730115

Sapolsky, R. M. (1996). Stress, glucocorticoids, and damage to the nervous system: the current state of confusion. *Stress* 1, 1–19. doi: 10.3109/10253899609001092

Singh, A., and Misra, N. (2009). Loneliness, depression and sociability in old age. *Ind. Psychiatry J.* 18, 51–55. doi: 10.4103/0972-6748.57861

Squire, L. R. (1991). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychol. Rev.* 99, 195–231.

Sundqvist, A., and Hemberg, J. (2021). Adolescents' and young adults' experiences of loneliness and their thoughts about its alleviation. *Int. J. Adolesc. Youth* 26, 238–255. doi: 10.1080/02673843.2021.1908903

Sundström, A., Adolfsson, A. N., Nordin, M., and Adolfsson, R. (2020). Loneliness increases the risk of all-cause dementia and Alzheimer's disease. *J. Gerontol. Ser. B Psychol. Sci. Soc. Sci.* 75, 919–926. doi: 10.1093/geronb/gbz139

Sundström, G., Fransson, E., Malmberg, B., and Davey, A. (2009). Loneliness among older Europeans. *Eur. J. Ageing* 6:267. doi: 10.1007/s10433-009-0134-8

Sutin, A. R., Stephan, Y., Luchetti, M., and Terracciano, A. (2020). Loneliness and Risk of Dementia. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 75, 1414–1422. doi: 10.1093/geronb/gby112

Tamnes, C. K., Bos, M. G. N., van de Kamp, F. C., Peters, S., and Crone, E. A. (2018). Longitudinal development of hippocampal subregions from childhood to adulthood. *Dev. Cogn. Neurosci.* 30, 212–222. doi: 10.1016/j.dcn.2018.03.009

Victor, C. R., and Yang, K. (2012). The prevalence of loneliness among adults: a case study of the United Kingdom. *J. Psychol.* 146, 85–104. doi: 10.1080/00223980.2011.613875

von Soest, T., Luhmann, M., Hansen, T., and Gerstorf, D. (2020). Development of loneliness in midlife and old age: Its nature and correlates. *J. Pers. Soc. Psychol.* 118, 388–406. doi: 10.1037/pspp0000219

Walhovd, K. B., Fjell, A. M., Westerhausen, R., Nyberg, L., Ebmeier, K. P., Lindenberger, U., et al. (2018). Healthy minds 0-100 years: optimising the use of European brain imaging cohorts ("Lifebrain"). *Eur. Psychiatry* 50, 47–56. doi: 10.1016/j.eurpsy.2017.12.006

Wiedenmayer, C. P., Bansal, R., Anderson, G. M., Zhu, H., Amat, J., Whiteman, R., et al. (2006). Cortisol levels and hippocampus volumes in healthy preadolescent children. *Biol. Psychiatry* 60, 856–861. doi: 10.1016/j.biopsych.2006.02.011

Wilson, R. S., Krueger, K. R., Arnold, S. E., Schneider, J. A., Kelly, J. F., Barnes, L. L., et al. (2007). Loneliness and risk of Alzheimer disease. *Arch. Gen. Psychiatry* 64, 234–240. doi: 10.1001/archpsyc.64.2.234

Yanguas, J., Pinazo-Henandis, S., and Tarazona-Santabalbina, F. J. (2018). The complexity of loneliness. *Acta Biomed.* 89, 302–314. doi: 10.23750/abm.v89i2.7404

Yesavage, J. A. (1988). Geriatric Depression Scale. *Psychopharmacol. Bull.* 24, 709–711.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *J. Mem. Lang.* 46, 441–517. doi: 10.1006/jmla.2002.2864

Zhong, B. L., Chen, S. L., Tu, X., and Conwell, Y. (2017). Loneliness and cognitive function in older adults: Findings from the Chinese longitudinal healthy longevity survey. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 72, 120–128. doi: 10.1093/geronb/gbw037

Conflict of Interest: CD is a founder, stockowner, board member, and consultant in Vitas Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership