



RESEARCH ARTICLE

REVISED Longitudinal trajectories of brain age in young individuals at familial risk of mood disorder from the Scottish Bipolar Family Study [version 3; peer review: 3 approved]

Previous title: Longitudinal trajectories of brain age in young individuals at familial risk of mood disorder

Laura de Nooij ^{1,2}, Mathew A. Harris¹, Emma L. Hawkins¹, Toni-Kim Clarke¹, Xueyi Shen¹, Stella W. Y. Chan³, Tim B. Ziermans ², Andrew M. McIntosh ¹, Heather C. Whalley ¹

¹Division of Psychiatry, University of Edinburgh, Edinburgh, UK

²Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

³Section of Clinical Psychology, University of Edinburgh, Edinburgh, UK

V3 First published: 17 Dec 2019, 4:206
<https://doi.org/10.12688/wellcomeopenres.15617.1>
 Second version: 30 Jul 2020, 4:206
<https://doi.org/10.12688/wellcomeopenres.15617.2>
 Latest published: 09 Nov 2020, 4:206
<https://doi.org/10.12688/wellcomeopenres.15617.3>

Abstract

Background: Within young individuals, mood disorder onset may be related to changes in trajectory of brain structure development. To date, however, longitudinal prospective studies remain scarce and show partly contradictory findings, with a lack of emphasis on changes at the level of global brain patterns. Cross-sectional adult studies have applied such methods and show that mood disorders are associated with accelerated brain aging. Currently, it remains unclear whether young individuals show differential brain structure aging trajectories associated with onset of mood disorder and/or presence of familial risk.

Methods: Participants included young individuals (15-30 years, 53%F) from the prospective longitudinal Scottish Bipolar Family Study with and without close family history of mood disorder. All were well at time of recruitment. Implementing a structural MRI-based brain age prediction model, we globally assessed individual trajectories of age-related structural change using the difference between predicted brain age and chronological age (brain-predicted age difference (brain-PAD)) at baseline and at 2-year follow-up. Based on follow-up clinical assessment, individuals were categorised into three groups: (i) controls who remained well (C-well, $n = 93$), (ii) high familial risk who remained well (HR-well, $n = 74$) and (iii) high familial risk who developed a mood disorder (HR-MD, $n = 35$).

Results: At baseline, brain-PAD was comparable between groups. Results showed statistically significant negative trajectories of brain-

Open Peer Review

Reviewer Status

| | Invited Reviewers | | |
|------------------|-------------------|--------|--------|
| | 1 | 2 | 3 |
| version 3 | | | |
| (revision) | | | |
| 09 Nov 2020 | report | | report |
| | | | |
| version 2 | | | |
| (revision) | | | |
| 30 Jul 2020 | report | report | report |
| | | | |
| version 1 | | | |
| 17 Dec 2019 | | | |
| | report | report | |

1. **Philip B. Mitchell** , University of New South Wales, Sydney, Australia
2. **Manpreet K. Singh** , Stanford University School of Medicine, Stanford, USA
3. **Janice Fullerton** , Neuroscience Research Australia, Sydney, Australia

Any reports and responses or comments on the

PAD between baseline and follow-up for HR-MD versus C-well ($\beta = -0.60$, $p_{\text{corrected}} < 0.001$) and HR-well ($\beta = -0.36$, $p_{\text{corrected}} = 0.02$), with a potential intermediate trajectory for HR-well ($\beta = -0.24$ years, $p_{\text{corrected}} = 0.06$).

Conclusions: These preliminary findings suggest that within young individuals, onset of mood disorder and familial risk may be associated with a deceleration in brain structure aging trajectories. Extended longitudinal research will need to corroborate findings of emerging maturational lags in relation to mood disorder risk and onset.

Keywords

depression, bipolar disorder, high risk, brain maturation, brain age, structural MRI, pattern recognition

.....
article can be found at the end of the article.

Corresponding author: Laura de Nooij (v1ldeno@ed.ac.uk)

Author roles: **de Nooij L:** Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Methodology, Project Administration, Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Harris MA:** Conceptualization, Formal Analysis, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Hawkins EL:** Data Curation, Project Administration, Writing – Review & Editing; **Clarke TK:** Methodology, Supervision, Writing – Review & Editing; **Shen X:** Methodology, Writing – Review & Editing; **Chan SWY:** Supervision, Writing – Review & Editing; **Ziermans TB:** Conceptualization, Supervision, Writing – Review & Editing; **McIntosh AM:** Conceptualization, Funding Acquisition, Methodology, Supervision, Writing – Review & Editing; **Whalley HC:** Conceptualization, Data Curation, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by IMAGEMEND, which received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) [grant number 602450], a Wellcome Trust Strategic Award [grant number 104036], and an Erasmus+ grant. All imaging aspects received financial support from the Dr Mortimer and Theresa Sackler Foundation. *The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

Copyright: © 2020 de Nooij L *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: de Nooij L, Harris MA, Hawkins EL *et al.* **Longitudinal trajectories of brain age in young individuals at familial risk of mood disorder from the Scottish Bipolar Family Study [version 3; peer review: 3 approved]** Wellcome Open Research 2020, 4:206 <https://doi.org/10.12688/wellcomeopenres.15617.3>

First published: 17 Dec 2019, 4:206 <https://doi.org/10.12688/wellcomeopenres.15617.1>

REVISED Amendments from Version 2

This new version includes minor changes:

- Correction of spelling inconsistency
- Updated extended data
- Reference to updated repository

Any further responses from the reviewers can be found at the end of the article

Introduction

Mood disorders are amongst the most common psychiatric disorders, with a life-time prevalence of around 15% (Kessler & Bromet, 2013). Globally, they are the greatest contributor to non-fatal ill-health (World Health Organization, 2017). However, underlying biological mechanisms remain unclear. It is known that mood disorders are highly heritable and share complex genetic architecture; individuals with a family history of Bipolar Disorder (BD) have >10-fold increased risk of developing BD or Major Depressive Disorder (MDD) (Smoller & Finn, 2003). Mood disorders often manifest during adolescence and young adulthood (de Girolamo *et al.*, 2012). During these life stages, age-related changes in brain structure contribute to cognitive development but also increase vulnerability to mental illness, including mood disorders (Andersen, 2003; Dahl, 2004).

From adolescence onward, decreases in brain grey matter and fine-tuning/stabilisation of synapses parallel changes in cognition and affect regulation (Giorgio *et al.*, 2010; Spear, 2000). For higher-order cortical areas these structural trajectories extend into young adulthood (Gogtay *et al.*, 2004; Wierenga *et al.*, 2014; Wierenga *et al.*, 2016). Previous prospective longitudinal studies including young individuals have shown inconsistent findings with regard to brain structure changes and mood disorder onset (Bos *et al.*, 2018; Ducharme *et al.*, 2014; Pappmeyer *et al.*, 2015a; Pappmeyer *et al.*, 2016; Whittle *et al.*, 2014). The most consistent results suggest that the frontal cortex (Ducharme *et al.*, 2014; Pappmeyer *et al.*, 2015a) and subcortical volumes (Whittle *et al.*, 2014) show decelerated brain structure aging trajectories. Theoretically, decelerated trajectories may contribute to vulnerability to mood disorders, particularly when frontal-limbic brain systems of cognitive control and emotional stability are affected. Previous studies in this field mostly focused on specific regions of interest, so that spatially unbiased comprehensive approaches investigating global patterns are currently lacking. We were therefore interested in determining from longitudinal prospective data of young individuals, whether a comprehensive and spatially unconstrained measure of brain structure aging trajectory across the brain was related to concurrent mood disorder onset and/or to familial risk.

A new framework that allows for global assessment of age-related patterns of structural change in the brain involves the estimation of an individual's "biological brain age" from an MRI scan. Subsequent comparison with chronological age provides the brain-predicted age difference (brain-PAD) as a

cross-sectional measure of brain aging. Conceptually, when brain aging trajectories that shape cognition and behaviour show individually different temporal dynamics, brain-PAD is expected to relate to relevant outcomes. Indeed, previous cross-sectional research with adults in their mid-later life showed that older appearing brains were associated with age-related diseases and mental illness (for overview see Cole *et al.*, 2019), including mood disorders (Han *et al.*, 2019; Koutsouleris *et al.*, 2014; but no effect in Nenadić *et al.*, 2017), and were furthermore predictive of mortality (Cole *et al.*, 2018). Interestingly, accelerated brain aging in mood disorders is in accordance with accelerated biological aging (Rizzo *et al.*, 2014; Sibille, 2013; Wolkowitz *et al.*, 2011) as well as increased risk of age-related disease and mortality (e.g., Mezuk *et al.*, 2008; Osby *et al.*, 2001; Pan *et al.*, 2011).

Within longitudinal designs including younger individuals, brain-PAD has potential to show the temporal origin of accelerated brain aging trajectories that have been observed in adult samples. Importantly, the brain-PAD approach has previously been applied and validated within samples of children and adolescents (Franke *et al.*, 2012). Furthermore, a previous cross-sectional study did not find differences in brain-PAD between young adults at high familial risk with mood disorder diagnosis, those at high familial risk who were well, and control subjects (Hajek *et al.*, 2019). To our knowledge, however, the current study is the first longitudinal study to apply brain-PAD methods within a sample of young individuals to investigate associations between mood disorder risk and onset, and age-related changes in brain structure.

Specifically, the current study investigated divergence of normative brain structure aging trajectories in young individuals by applying the brain-PAD framework within a prospective longitudinal design, starting before mood disorder onset. We used data from the Scottish Bipolar Family Study (SBFS), which included young individuals who were all initially well and some of whom had a close family history of BD. Within this cohort, our group previously identified differences in cortical thickness trajectories associated with high risk and mood disorder onset, including increased thickness of the left inferior frontal gyrus and left precentral gyrus in those at high risk who subsequently developed mood disorder versus cortical thickness reductions in those who remained well (Pappmeyer *et al.*, 2015a). By contrast, no subcortical volume markers of risk and illness were found (Pappmeyer *et al.*, 2016). Investigation of white matter structure at baseline furthermore revealed reduced white matter integrity associated with familial risk (Sprooten *et al.*, 2011), and follow-up data suggested that this finding was related to sub-clinical symptoms rather than predictive of clinical outcome (Ganzola *et al.*, 2018). The current study builds on previous research within the BFS cohort, which identified differences in specific grey matter regions and white matter abnormalities, by investigating global trajectories of grey matter structure associated with familial risk and onset of mood disorder.

Recognising similarities between BD and MDD in symptomatology and genetic architecture, as well as the difficulty of

defining a definitive stable diagnosis at young age, early-onset mood disorder was defined as having an onset of MDD or BD during adolescence or young adulthood. The longitudinal character of the study enabled the investigation of brain-PAD over two years, to assess differential brain structure aging trajectories for those who were at high risk for mood disorder and/or subsequently developed illness.

Based on previous research, we predicted that mood disorder onset in youth would be associated with differential trajectories of brain structural change, without a specific hypothesis relating to the direction of this effect. Previous results from longitudinal developmental studies are inconclusive, but show weak evidence of decelerated trajectories, which also corresponds to theoretical developmental perspectives. Conversely, early adulthood may represent the temporal origin of the accelerated brain aging observed in older adults, in which case we would expect an effect of mood disorder in this direction instead. We also hypothesised that the presence of familial risk would be associated with differences in brain-PAD trajectory.

Methods

Participants

Participants were adolescents and young adults ($N = 283$, age 15–30 years) recruited as part of the Scottish Bipolar Family Study (SBFS) (Chan *et al.*, 2016; Ganzola *et al.*, 2018; Sprooten *et al.*, 2011; Whalley *et al.*, 2015). Participants at high familial risk of mood disorder (HR-participants) had at least one first-degree relative or two second-degree relatives with BD type-I, and were thus at increased risk of developing a mood disorder (i.e., over 10-fold increased risk for both BD and MDD) (Smoller & Finn, 2003). Unrelated control participants without family history of BD or other mood disorder were recruited from the social networks of HR-participants, and were matched to the HR-group by age and sex. Details of familial structure within the groups are described in the *Extended data*, which are available online (de Nooij, 2020). Exclusion criteria ensured that, at the time of recruitment, all participants had no personal history of MDD, mania or hypomania, psychosis, or any other major neurological or psychiatric disorder, substance dependence, learning disability, or head injury that included loss of consciousness, and that they were without contraindications to MRI. Therefore, all individuals (HR and control) were considered well at the baseline imaging assessment.

The following additional exclusion criteria were applied in the context of the current study: (i) missing MRI or age data ($n = 40$), (ii) scans of insufficient image or segmentation quality ($n = 15$) (iii) unclear or other psychiatric diagnosis without mood disorder ($n = 5$) (see *Extended data* (de Nooij, 2020)), and (iv) high familial risk for mood disorder without follow-up measurement ($n = 9$). These criteria excluded 69 participants, reducing the sample size to a total of 214 participants at timepoint 1 (108 HR-participants), with follow-up timepoint 2 data available for 133 of these participants (78 HR-participants).

Procedure

Participants of the SBFS were invited every two years for a total of four assessments over six years (Whalley *et al.*, 2015). Participants were interviewed and screened with the Structured Clinical Interview for DSM-IV Axis-I Disorders (SCID) (First *et al.*, 2002) by two trained psychiatrists at timepoint 1 to ensure that they were all initially well, and at timepoint 2 to determine the presence of any mood disorder meeting diagnostic criteria since the previous assessment. Timepoint 2 clinical information was available for 93% of the included control participants, and for all included HR-participants. Of note, control participants with missing clinical information at timepoint 2 (7%) also had missing timepoint 2 MRI data, but their baseline data was retained to increase the training sample size and contributed to statistical modelling of the mean brain-PAD at baseline. Using the same procedure as previous studies on this cohort (Chan *et al.*, 2016; Whalley *et al.*, 2015), participants were categorised as well or diagnosed with mood disorder according to available clinical information. Individuals with well outcomes at the earlier two assessments were assumed to have remained well in the absence of further clinical information to the contrary at timepoint 3 (see Appendix A.3, Table S1 in *Extended data* (de Nooij, 2020)). Additionally, however, if individuals were subsequently found to have been diagnosed with mood disorder at further assessments ($n = 13$), they were then categorised in the mood disorder group. Including these participants in the mood disorder group enables the investigation of early disease mechanisms, while keeping the well-groups as pure as possible. Group categorisation resulted in the following groups: control participants who remained well (C-well, $n = 93$), HR-participants who remained well (HR-well, $n = 74$), and HR-participants who developed a mood disorder (HR-MD, $n = 35$, including 6 BD). Thus, for the control group, individuals were included if they had remained without being diagnosed with any psychiatric disorder throughout the period of the study (via assessments and/or GP records). Those that did become unwell were a small group ($n=12$, including 2 BD) and are not included in the current analysis. This approach was used to reduce heterogeneity in the control group. For the high-risk group, none met clinical criteria for any mood disorder at baseline. If at any assessment they did meet criteria for a mood disorder (at the time of assessment or over the intervening period since the earlier assessment) they were considered as being in the HR-MD group. So, once someone had a diagnosis, even though they may not have been actively symptomatic at the time of the assessment, they were still considered as being in the mood disorder group. This approach was based on the premise that once an individual had met diagnostic criteria at any time over the course of the study, they were no longer ‘high risk’ for mood disorder, but actually a ‘case’. The sample for our main analysis therefore consisted of 202 participants at baseline and 124 participants at follow-up.

The National Adult Reading Test (NART) (Nelson & Willison, 1991) and Hamilton Rating Scale for Depression (HRSD) (Hamilton, 1960) were administered at the time of

scanning. Mania was also determined at every assessment using the Young Mania Rating Scale (YMRS, ref), however previous studies indicated that there were no significant differences between the three groups in terms of the YMRS cross-sectionally or over time, and the median and interquartile ranges were low, hence we did not specifically examine mania ratings in the current analysis (Papmeyer *et al.*, 2015b; Papmeyer *et al.*, 2016; Young *et al.*, 2000). The participant's age at the time of each assessment was registered in years with a precision of two decimals. Assessments at timepoint 1, timepoint 2 and timepoint 4 included an MRI session, although only MRI measurements at timepoint 1 and timepoint 2 were considered within this study to restrict to a single scanner. The SBFS was approved by the Research Ethics Committee for Scotland (reference number 06/MRE00/9), and written informed consent including consent for data linkage via medical health records was acquired from all participants. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

MRI acquisition and pre-processing

Timepoint 1 and timepoint 2 MRI sessions, approximately two years apart, were carried out on a 1.5 T Signa scanner (GE Medical, Milwaukee, USA) at the Brain Research Imaging Centre in Edinburgh and included a structural T1 weighted sequence (180 contiguous 1.2 mm coronal slices; matrix = 192 × 192; fov = 24 cm; flip angle 8°).

Pre-processing of T1 weighted scans was done in Statistical Parametric Mapping (SPM) version 12. The [Computational Anatomy Toolbox \(CAT\) toolbox](#) (version CAT12.3 (r1318)), which runs on SPM12 software, was used to segment T1-weighted MRI scans into different tissue types (for details see Appendix A.4 in *Extended data* (de Nooij, 2020)). A cross-sectional segmentation approach was utilised in order to maximise the size of our training sample, and the longitudinal aspect of our data was handled with repeated measures linear mixed modelling (see [Comparison of brain maturation trajectories](#)). This approach avoided the exclusion of participants with incomplete MRI data. CAT12 Quality Assurance metrics were used in combination with manual checks to achieve an objective and comprehensive procedure to exclude scans with artefacts or of otherwise insufficient quality (see Appendix A.4 in *Extended data* (de Nooij, 2020)). Subsequently, modulated grey matter maps (GMM) were smoothed with a Gaussian kernel (FWHM = 8 mm). After loading the smoothed GMM (sGMM) into Python version 3.5.4, voxels were resampled into voxels of double the original voxel size, i.e. 3 × 3 × 3 mm³. This reduced the number of voxels without further loss of spatial information. The sGMM were then masked with a threshold of 0.01 to ensure that voxels outside the brain were represented by value zero. The resulting sGMM were used as input for the brain-PAD model.

Brain-PAD model

To initially train the brain age prediction model, the training sample included all control and HR-participants that remained

well ($n = 167$) in order to maximise the healthy sample size (a model including control participants only was considered underpowered, see Appendix A.5 in *Extended data* (de Nooij, 2020)). The current model was equally balanced across timepoint 1 and timepoint 2 measurements in order to maximise the age range. Specifically, each well-group participant provided one scan for the training sample: 48 timepoint 1 scans and 46 timepoint 2 scans (i.e. all available scans) for C-well, together with 37 scans per timepoint for HR-well. HR-well timepoint 2 scans were selected based on the highest chronological ages at follow-up so that the age range covered by the training sample was maximal ($M_{\text{age}} = 22.37$, $SD_{\text{age}} = 2.94$, age range = 15.2–28.1 years; for age distributions see Figure S1 in Appendix A.6 in *Extended data* (de Nooij, 2020)). Specifically, each well-group participant provided one scan for the training sample; this was a timepoint 2 scan for all 46 C-well participants with follow-up scan and the 37 HR-well participants with the highest chronological ages at follow-up.

Similar to previous studies (see Cole *et al.*, 2019), the sGMM and corresponding chronological ages of the training sample were used to train a brain age prediction model. This model was implemented in Python (version 2.6.6). Corresponding to recent recommendations (Smith *et al.*, 2019), this model initially consisted of dimension reduction of all sGMM voxels to 73 brain components (based on eigenvalue > 1) using principal component analysis (PCA) based on singular value decomposition (SVD) from scikit-learn (Pedregosa *et al.*, 2011). We subsequently used these brain components (X) and chronological age (y) as input for estimating a Relevance Vector Regression (RVR) model with linear kernel (Tipping, 2001); this was implemented using the publicly available [scikit-rvm](#) package (version 14 May 2017). The RVR algorithm was chosen because kernel-based methods have been most commonly implemented in brain age models (Cole *et al.*, 2019), because linear RVR was found to be the favourable algorithm in a previous brain-PAD study (Franke *et al.*, 2010) and because RVR does not require estimation of hyperparameters using cross-validation (a procedure that would limit our sample size).

The trained model was then applied to each participant's sGMM to predict their brain age, ensuring that the participant for whom the brain age was being predicted was left out of the training sample to prevent bias (leave-one-out cross-validation). A residuals approach was used to regress out chronological age and gender, and subsequently calculate brain-PAD (for details see Appendix A.7 in *Extended data* (de Nooij, 2020)), i.e. the gap between brain age prediction and chronological age. This residuals based approach is typically used to derive measures of accelerated aging (e.g. epigenetic aging; Chen *et al.*, 2016; Horvath, 2013) and is recommended for the brain-PAD approach (Smith *et al.*, 2019).

Regarding brain development, a positive brain-PAD reflected a brain-predicted age older than the chronological age of the participant, while a negative brain-PAD indicated a brain-predicted age younger than the participant's chronological age. Changes in brain-PAD over time indicated a relative acceleration in brain maturation if brain-PAD became more

positive (or less negative), or a relative deceleration in brain maturation if brain-PAD became more negative (or less positive).

Given the aim of the current study to specifically investigate brain structure aging trajectories within the SBFS cohort, as well as the demographics of our cohort (particularly the narrow age range, also including late adolescence), we achieved within-sample model evaluation based on the brain age predictions for the training sample, using leave-one-out cross-validation.

Comparison of brain maturation trajectories

Since the objective of this study was to investigate deviation of brain maturation trajectories in young individuals at high risk for mood disorder and the association with illness onset, participants were divided in three groups based on clinical information as described above. Clinical information from all available assessments was considered in group categorisation as described above.

In order to compare brain structure aging trajectories between groups, we applied a linear mixed model (LMM) to the brain-PAD measures, taking into account loss to follow-up as well as individual and family-related effects (Gueorguieva & Krystal, 2004). This was modelled using R (version 3.2.3) package `nlme` (version 3.1-122) with the formula: 'Brain-PAD ~ Timepoint × Group, random = ~1 | FamilyID / SubjectID'. Within this single pre-defined LMM model, we were interested in the following contrasts: group differences in brain-PAD at baseline (group effect), differential trajectories of brain-PAD between groups (group by timepoint interaction effect), and group differences at follow-up. For these contrasts we tested all three pairwise comparisons, and we multiple comparison corrected results ($n = 3$ pairwise comparisons) with the Holm-Bonferroni method (Holm, 1979) using R package `emmeans` (version 1.3.5.1).

Exploratory analyses were conducted to further explore group differences in Brain-PAD trajectories. Firstly, we tested a longitudinal model that considered the interaction effect between age (at baseline) and group on the difference in brain-PAD between baseline and follow-up; this was modelled in R using the formula: 'Brain-PAD_difference ~ Age_baseline × Group, random = ~1 | FamilyID/SubjectID'. A second exploratory analysis also modelled the brain-PAD trajectory for the group of control participants who developed a mood disorder (C-MD) within the LMM of the main analysis, considering the pairwise comparisons with control group C-well. In all of the analyses described above, continuous variables (brain-PAD, age) were transformed to Z-scores to retrieve standardised β -coefficients.

Results

Demographic and clinical variables

Sample sizes, demographic information and clinical measures are presented in Table 1. There were no significant differences between groups with regard to age at either timepoint, and no differences in gender, handedness and NART intelligence quotient score. Aggregate information is available for each participant as *Underlying data* (de Nooij, 2020).

However, HR-MD participants reported greater depression symptomatology on the HRSD as compared to the groups of participants who remained well (C-well and HR-well) at both timepoints (Table 1). At baseline, seven HR-MD participants (20%; $M_{\text{HRSD}} = 11.4$) reported subclinical symptoms of depression (defined as HRSD score > 7). At timepoint 2, ten HR-MD participants reported symptoms of depression (defined as HRSD score > 7). For two of these participants depression symptoms were at subclinical level, as they were not yet diagnosed with a mood disorder. In contrast, there was a very low prevalence of subclinical depression symptomatology within the well-groups: two participants at timepoint 1 (1.2%; $M_{\text{HRSD}} = 9.0$) and two other participants (with included follow-up scans) at timepoint 2 (2.6%; $M_{\text{HRSD}} = 9.0$).

Participants did not report any use of psychotropic medication at baseline. At follow-up, six included HR-MD participants reported the use of psychotropic medication, whereas none of the well-group participants received medication for the treatment of psychiatric symptoms.

HR-MD showed lower attrition (11.4%) than C-well (50.5%; $\chi^2(1) = 14.6$, $p < 0.001$) and HR-well (36.5%; $\chi^2(1) = 6.2$, $p = 0.01$), with no significant difference between C-well and HR-well ($\chi^2(1) = 2.8$, $p = 0.10$). None of the clinical or demographic variables at baseline differed between those individuals with and without a follow-up scan (see Table S2, Appendix B.1 in *Extended data* (de Nooij, 2020)).

Model evaluation

Our model showed a significant positive Pearson correlation between predicted brain age and chronological age ($r(165) = .40$, $p < 0.001$), and a mean absolute error (MAE) of 2.21 years (scaled MAE = MAE / age range = 0.17; see Appendix B.2 in *Extended data* (de Nooij, 2020)) within the training sample. For a discussion on model evaluation within the context of the current study, see *Discussion* and Appendix B.2 in *Extended data* (de Nooij, 2020).

The 73 brain components that were used as input for the brain age prediction algorithm indicated a mean total explained variance of 84.0% (SD = 0.0004) for all (leave-one-out) training sample dimension reduction iterations. These brain components showed loadings distributed across the brain, because dimension reduction was spatially unconstrained. This complicated unbiased interpretation (Smith *et al.*, 2019), and therefore, also given our aim to comprehensively assess global patterns of brain structure aging trajectories, these components were not further explored. However, we do present visualisation of these brain components in order to illustrate the method (Figure S3 in Appendix B.3 in *Extended data* (de Nooij, 2020)).

Comparison of brain maturation trajectories

Comparison at baseline

Group allocation based on diagnostic information resulted in mean brain-PADs of +0.04 (SD = 1.14, $n = 93$) for C-well, -0.36 (SD = 1.22, $n = 74$) for HR-well, and -0.01 (SD = 1.39, $n = 35$) for HR-MD. Results of the LMM (Table 2) suggested

Table 1. Demographic and clinical characteristics.

| | C-well | HR-well | HR-MD | p-value |
|--------------------------------|------------|------------|------------|------------------------|
| n, timepoint 1 | 93 | 74 | 35 | |
| n, timepoint 2 | 46 | 47 | 31 | |
| Age, timepoint 1 ^a | 21.6 (3.8) | 21.6 (4.9) | 21.4 (5.2) | 0.82 |
| Age range | 16.3-25.6 | 15.2-26.6 | 16.0-30.0 | |
| Age, timepoint 2 ^a | 23.4 (3.6) | 23.8 (5.2) | 23.7 (5.0) | 0.45 |
| Age range | 18.3-27.6 | 17.6-28.1 | 18.1-28.1 | |
| Gender ^b | | | | 0.51 |
| Male | 42 (45.2%) | 38 (51.4%) | 14 (40.0%) | |
| Female | 51 (54.8%) | 36 (48.6%) | 21 (60.0%) | |
| Handedness ^b | | | | 0.32 |
| Left | 5 (5.4%) | 7 (9.5%) | 2 (5.7%) | |
| Right | 86 (92.5%) | 65 (87.8%) | 33 (94.3%) | |
| Mixed | 0 (0%) | 2 (2.7%) | 0 (0%) | |
| Unknown | 2 (2.2%) | 0 (0%) | 0 (0%) | |
| NART score ^a | 111 (9.0) | 111 (10.7) | 107 (8.9) | 0.15 |
| HRSD, timepoint 1 ^a | 0 (1.0) | 0 (2.0) | 2 (5.5) | <0.001*** ^α |
| HRSD, timepoint 2 ^a | 0 (2.0) | 0 (1.0) | 5 (8.0) | <0.001*** ^α |

*** $p < 0.001$

Note: Individual NART scores were averaged over all completed assessments (max. 4).

^a Medians and interquartile ranges for variables not normally distributed (Kruskal-Wallis test)

^b Frequency and percentages for categorical variables (Chi-squared test).

^α HR-MD vs. C-well and HR-MD vs. HR-well (Kruskal-Wallis test, followed by Dunn's test for pairwise comparisons)

C-well, group of participants without family history who remained well; HR-MD, group of participants at high familial risk who developed a mood disorder; HRSD, Hamilton Rating Scale for Depression; HR-well, group of participants at high familial risk who remained well; NART, National Adult Reading Test.

Table 2. Fixed effects of linear mixed model applied to investigate group differences in the brain-predicted age difference (brain-PAD).

| Main model - Comparison to C-well | | | | | |
|------------------------------------|----------------------|------|-----|---------|-----------|
| Fixed effect | β -coefficient | SE | df | t-value | p-value |
| (Intercept) | 0.09 | 0.10 | 174 | 0.84 | 0.55 |
| Timepoint 2 | 0.37 | 0.09 | 121 | 4.10 | <0.001*** |
| HR-well | -0.37 | 0.16 | 25 | -2.33 | 0.03* |
| HR-MD | -0.07 | 0.20 | 25 | -0.35 | 0.72 |
| Timepoint 2*HR-well | -0.24 | 0.13 | 121 | -1.93 | 0.06 |
| Timepoint 2*HR-MD | -0.60 | 0.14 | 121 | -4.23 | <0.001*** |
| Main model - Comparison to HR-well | | | | | |
| Fixed effect | β -coefficient | SE | df | t-value | p-value |
| HR-MD | 0.30 | 0.21 | 25 | 1.14 | 0.16 |
| Timepoint 2*HR-MD | -0.36 | 0.14 | 121 | -2.52 | 0.01* |

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: β -coefficients are standardised following scaling of the outcome variable.

C-well, group of participants without family history who remained well; HR-MD, group of participants at high familial risk who developed a mood disorder; HR-well, group of participants at high familial risk who remained well.

lower baseline brain-PAD for HR-well compared to C-well (-0.42 years; $\beta = -0.37$, $p = 0.03$, $p_{\text{corrected}} = 0.08$), but statistical significance did not survive multiple comparison correction. There were no baseline differences in brain-PADs for HR-MD versus C-well (-0.05 years; $\beta = -0.07$, $p_{\text{corrected}} = 0.73$) or HR-MD versus HR-well (0.35 years; $\beta = 0.30$, $p_{\text{corrected}} = 0.24$).

Brain structure aging trajectories

Results showed a statistically significant timepoint by group interaction effect for HR-MD compared to C-well (-0.70 years; $\beta = -0.60$, $p_{\text{corrected}} < 0.001$) and HR-well (-0.43 years; $\beta = -0.36$, $p_{\text{corrected}} = 0.02$), indicating decelerating brain structure aging trajectories. Besides that, HR-well showed an intermediate trajectory (-0.28 years; $\beta = -0.24$, $p_{\text{corrected}} = 0.06$) which was not statistically significant. Figure 1 displays brain maturation trajectories per group as modelled by unstandardised LMM fixed effects; for clarity, these trajectories are displayed relative to the control group following correction for the effects observed in C-well (i.e., intercept and the significant timepoint coefficient, see Table 2). Figure 2 shows the heterogeneity in observed brain maturation trajectories by displaying the participants' individual changes in brain-PAD over time.

Comparison at follow-up

At follow-up, two years later, the mean brain-PADs were +0.36 (SD = 1.01) for C-well, -0.08 (SD = 1.04) for HR-well, and -0.30 (SD = 1.30) for HR-MD. Results indicated a statistical significant difference in brain-PAD between HR-MD and C-well (-0.69 years; $\beta = -0.61$, $p = 0.02$, $p_{\text{corrected}} = 0.06$), and between HR-well and C-well (-0.54 years; $\beta = -0.48$, $p = 0.04$, $p_{\text{corrected}} = 0.06$), although these results not survive multiple comparison correction. We found no evidence for a difference between HR-MD and HR-well at follow-up (-0.15 years; $\beta = -0.13$, $p_{\text{corrected}} = 0.57$).

Exploratory findings

Our exploratory longitudinal model showed similar group trajectories as our main model. Furthermore, this model suggested

that within our sample, younger individuals from the HR-MD group showed greater deceleration in their structural brain trajectory than older HR-MD individuals (HR-MD, $\beta = -0.90$, $p = 0.002$; with Age \times HR-MD, $\beta = 0.53$, $p = 0.03$), whereas no such effects were found for the HR-well group (HR-well, $\beta = -0.38$, $p = 0.06$; with Age \times HR-MD, $\beta = 0.53$, $p = 0.94$) (see Figure S4 and Table S4, Appendix B.4 in

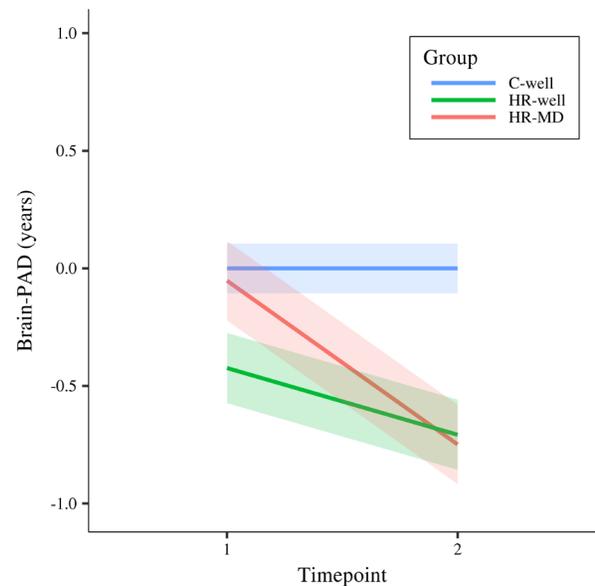


Figure 1. Modelled fixed effects of the brain-predicted age difference (brain-PAD) per group, for clarity corrected for effects in C-well (i.e., the intercept and timepoint coefficients) as this group functions as control group. Shaded areas display standard errors of the timepoint by group interaction effects. Brain-PAD, brain-predicted age difference; C-well, group of participants without family history who remained well; HR-MD, group of participants at high familial risk who developed a mood disorder; HR-well, group of participants at high familial risk who remained well.

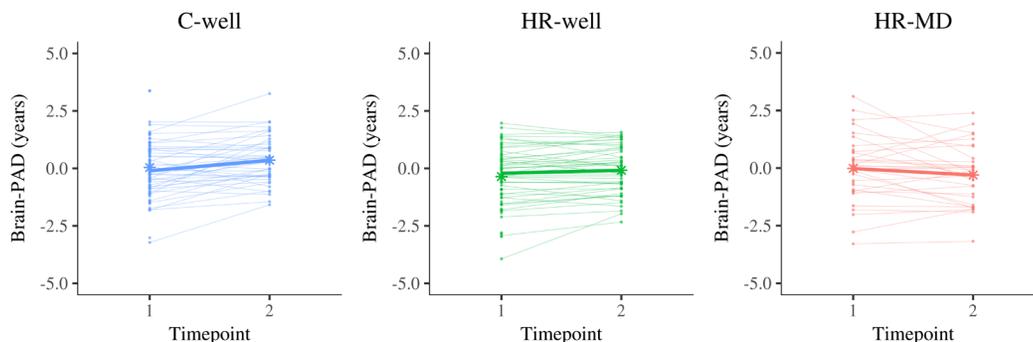


Figure 2. Display of brain structure aging trajectories per participant, reflecting a changing brain-predicted age difference (brain-PAD) between timepoint 1 and timepoint 2 (two years apart). Each panel contains the trajectories of one group in thin line graphs, whereas the thicker line graph represents the average trajectory of that group (of complete cases). The star dots display the mean brain-PAD at each timepoint. Left panel, C-well; Middle panel, HR-well; Right panel, HR-MD. Brain-PAD, brain-predicted age difference; C-well, group of participants without family history who remained well; HR-MD, group of participants at high familial risk who developed a mood disorder; HR-well, group of participants at high familial risk who remained well.

Extended data (de Nooij, 2020)). Exploratory findings of the model including the C-MD group showed a non-significant negative trajectory of brain-PAD for this group (-0.42 years; $\beta = -0.36$, $p = 0.11$; see Table S5 and Figure S5, Appendix B.5 in *Extended data (de Nooij, 2020)*).

Discussion

The results of the current study showed that in young individuals at familial risk the onset of mood disorder was associated with differences in brain structure changes over time. Statistically significant reductions in brain-PAD indicated decelerated brain structure aging trajectories in young HR individuals who developed a mood disorder as compared to control and HR individuals who remained well. Intermediate effect sizes indicated that young individuals who were at risk but remained well showed intermediate trajectories. These preliminary findings suggest genetic predisposition to mood disorder is accompanied by changes in adolescent brain structural development trajectories that are increased with the onset of mood disorder.

Further research will be necessary to disentangle the role of genetic predisposition and additional environmental risk factors (e.g. adverse life events) on global age-related brain structure changes. As development of mood disorder was associated with a more decelerating trajectory, differences observed for the mood disorder group may also partly reflect prodromal symptoms or early-disease mechanisms of psychological stress. Further, in the familial risk group who became ill, we cannot disentangle separate effects of risk and depressive symptoms. Notably, all groups showed considerable heterogeneity in the direction, size and emergence of the individual brain-PADs. Additional research is therefore required to substantiate the hypothesis that the emergence of a lag in brain structure aging in youth indicates mood disorder onset and familial risk.

The current findings correspond with previous neuroimaging studies using different methods that also indicated deceleration, in the same as well as independent prospective longitudinal cohorts (Ducharme *et al.*, 2014; Pappmeyer *et al.*, 2015a; Whittle *et al.*, 2014). According to empirical-based neural models, dysfunctions in medial prefrontal networks and limbic areas underlie disturbances in emotion regulation and cognitive control (e.g. Drevets *et al.*, 2008) which are proposed to play a causal role in the development of mood disorder (Nolen-Hoeksema *et al.*, 2008; Phillips *et al.*, 2008). Correspondingly, previous findings within the same cohort have revealed that illness risk and onset were associated with differential cortical thickness trajectories in prefrontal areas (Pappmeyer *et al.*, 2015a) as well as differential patterns of brain activation during emotional tasks in cortico-thalamic-limbic regions (Chan *et al.*, 2016; Whalley *et al.*, 2015), and neurocognitive performance was found to be a trait-marker of familial risk (Pappmeyer *et al.*, 2015b). Although the current study adopted a global approach (thus refraining from investigation of regional brain structural or functional development), we speculate about a potential neural mechanism by which decelerated trajectories of brain structural change in young individuals potentially disrupt frontal and limbic

brain networks that underly emotion regulation and cognitive control, consequently increasing vulnerability to mood disorder. Inferences of causality however should be drawn with caution. It is important to consider that though prospective longitudinal studies are one approach to examining causal processes, interpretation is complex. In the current study for example, individuals who subsequently developed a mood disorder also showed higher mean subclinical depression symptomatology at baseline. This could be interpreted as a predictor of subsequent illness, or indeed as prodromal or early stages of the illness itself.

Importantly, our findings suggest disease-related brain aging deceleration may emerge in young individuals, in contrast to previous findings of accelerated aging in association with mood disorder (Koutsouleris *et al.*, 2014; Sibille, 2013; Wolkowitz *et al.*, 2011). However, we note the majority of these studies are conducted in older adult samples, rather than younger adolescent cohorts. Given the non-linear trajectory of regional brain maturation / aging over the life-course, particularly over periods where there is significant developmental change (e.g. adolescence), it may not be the case that there is a simple retrospective linear trajectory of accelerated biological aging from these adult studies to periods earlier in life (Giedd *et al.*, 1999; Scahill *et al.*, 2003; Shaw *et al.*, 2008; Tamnes *et al.*, 2010; Wierenga *et al.*, 2014). Further, this finding is consistent with wider theories of adolescent psychopathology, for example the ‘dual system’ model where delayed maturation in higher order cortical regions in relation to limbic regions is proposed to underlie difficulties in emotional regulation, cognitive function and social behaviour, increasing vulnerability to mood disorder (Casey *et al.*, 2011).

The current study applied a novel pattern recognition method that, to our knowledge, has not been previously applied to a longitudinal cohort of young individuals at risk of mood disorder. This approach derives a global measure of brain structure, which captures the complexity of spatial and temporal dynamics of brain aging. Challenges in collecting clinical data from young individuals mean that large cohorts are scarce; the SBFS provided a unique opportunity to investigate the dynamics of brain structure aging trajectories in relation to mood disorder. Development of mood disorder was found to be associated with decelerated age-related changes in brain grey matter, which could not have been identified within a cross-sectional design. Clinical information was also available for up to six years, which produced some heterogeneity in the HR-MD group, due to a range in times before onset, but also provided confidence that those classified as HR-well were not in the early stages of mood disorder at the time of imaging assessments. Overall, the current study of the SBFS shows unique strengths for youth mental health research. Ongoing work on the sample seeks to implement data linkage at 10+ years to obtain more definitive, stable diagnoses.

One limitation of our study was the low correlation between brain age prediction and chronological age, reflecting suboptimal performance of the brain age prediction model, although this is probably also related to the tight age range of

our sample as well as individual differences within this life stage. Performance of the model was constrained by the limited cohort size, as for the purpose of the current study, we adopted a within-sample approach. In order to maximise the use of available data in building our brain age prediction model, we adhered to a cross-sectional segmentation approach and included all individuals who remained well, also those at high familial risk for mood disorder. Brain-PAD slightly increased over time within the control group, indicating that a brain-PAD of zero did not indicate normative brain maturation in the current study (for details see Appendix B.2 in *Extended data* (de Nooij, 2020)). This suggests that the brain age prediction model was biased by familial risk related structural differences following inclusion of HR-well participants within the training sample. Although this explanation would not invalidate our results, as it would suggest valid comparison of relative group differences in brain-PAD, it is considered a limitation that we cannot reliably tease apart the familial risk effect from the normative trajectory. Importantly, we directly addressed potential threats to the validity of the brain age prediction model, to the extent possible within the current sample, with a deliberate pre-defined approach consisting of dimension reduction, sparse RVR modelling and a residuals approach, in order to prevent overfitting and thereby optimise the overall model validity. However, our prediction model was not validated for generalisability to other samples because of challenges related to scanner heterogeneity, so that transferability of the model remains uncertain. Further limitations of the current study are that we were unable to investigate differences between MDD and BD, and that we cannot exclude the possibility of medication effects, although use of psychotropic medications was limited within the current sample (see *Demographic and clinical variables*). Additionally, out-of-sample predictions often show more prediction error than within-sample predictions achieved using cross-validation (Varoquaux *et al.*, 2017). Although participants from all three groups belonged to the same cohort and were recruited and assessed according to the same procedures, brain age predictions for individuals with mood disorder onset (i.e., HR-MD) were out-of-training-sample predictions, whereas predictions for individuals who remained well (i.e., HR-well and C-well) required leave-one-out cross-validation. Although these differential prediction procedures may have led to increased prediction error for HR-MD, our finding of intermediate trajectories in the HR-well group suggests that results in HR-MD are unlikely to be driven entirely by increased random error. Further, additional testing of out of training sample scans of well-group participants indicated that prediction error was not significantly increased compared to within training sample estimates, conferring further confidence in our findings. However, taken together, the findings of the current study are considered preliminary as they should be interpreted in the context of these limitations. We also note that in the HR-MD group, the depression severity scores are relatively low. However, this group categorisation was based on the presence of a previous or current diagnosis, rather than on the severity of current symptoms at the time of the scan. One final limitation is the use of a 1.5T scanner. Though this is not ideal given current technology, other

studies had successfully applied similar methods to 1.5T brain MRI scans previous to the current study (Franke & Gaser, 2012; Gaser *et al.*, 2013). The choice of scanner strength was determined at the beginning of this 10 year prospective study. We considered the value of having prospective longitudinal data, and together with detailed QC this provided sufficient confidence in the quality of data used in the current study.

In order to resolve the above limitations, future research should aim to replicate our results within a larger sample (Button *et al.*, 2013; Jollans *et al.*, 2019). Large-scaled and extended MRI follow-up assessments would furthermore allow the application of a longitudinal brain age prediction model, which will provide a more nuanced understanding of individual developmental trajectories. A sufficient sample size would also allow for investigation of MDD and BD separately, and could account for potential medication effects. Spatial interpretability of the current model's brain age prediction was limited, but with a larger sample methods such as orthonormal projective non-negative matrix factorisation (OPNMF) could provide information about specific regions or networks involved in associations between brain age and mood disorder (Sotiras *et al.*, 2015; Sotiras *et al.*, 2017; Varikuti *et al.*, 2018). Additionally, our exploratory findings suggest it may be useful to investigate associations between brain structure aging trajectories and mood disorder in younger samples. This along with leveraging opportunities from other statistical approaches to determine causal directionality, could be implemented to attempt to understand causal relationships between imaging findings and mood symptoms. For now, the present study lays a theoretical and empirical foundation for the field to build upon, and will hopefully encourage further longitudinal studies of clinical youth cohorts. In the future, replication and further investigation of the association between mood disorder and decelerated brain structure aging trajectories may provide important insights into the prediction of mood disorder onset in young individuals.

Data availability

Underlying data

Open Science Framework: Brain age trajectories and mood disorders (SBFS). <https://doi.org/10.17605/OSF.IO/QKCYD> (de Nooij, 2020).

This project contains the following underlying data:

- SBFS_Data (containing raw demographic information and brain ages for all participants)

For reasons of confidentiality, we are unable to openly share underlying MRI data. Specifically, Research Ethics Committee approval for this project requires raw MRI data to remain stored within the computer system of University of Edinburgh. MR images and other types of data from the SBFS cohort can only be shared in anonymized form, and only with other organisations within the European Union, using a secure data transfer. For this reason, directly accessible underlying data provided via OSF only consists of the unidentifiable underlying data that is needed to reproduce the results of this manuscript, including

the intermediate brain age prediction imaging phenotype. Access to MR images and other data requires the arrangement of a data access contract, please contact data holder Dr Heather Whalley (Heather.Whalley@ed.ac.uk) for more information and arranging access.

Extended data

Open Science Framework: Open Science Framework: Brain age trajectories and mood disorders (SBFS). <https://doi.org/10.17605/OSF.IO/QKCYD> (de Nooij, 2020).

This project contains the following extended data:

- Supplementary information
 - Appendix A: Supplementary Methods.
 - Appendix B: Supplementary Results.
 - Appendix C: Supplementary References.
 - Supplementary figures and corresponding code¹.

¹ These figures were created in response to peer review comments and are referred to in the author's response to peer review.

- SBFS_Code (containing annotated Python code for derivation of brain ages, R Markdown integrated code for statistical analysis).

Accessible data are available under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/) (CC-BY 4.0).

Acknowledgements

We would like to thank all of the participants who took part in the study and the radiographers who acquired the MRI scans. We also acknowledge the support of National Health Service (NHS) Research Scotland, through the Scottish Mental Health Research Network (www.smhrn.org.uk), who provided assistance with subject recruitment and cognitive assessments, and health professionals who cooperated in accessing clinical health record information. This study was conducted at the Brain Research Imaging Centre (<http://www.bric.ed.ac.uk>), which is supported by SINAPSE (Scottish Imaging Network, a Platform for Scientific Excellence, <http://www.sinapse.ac.uk>).

An earlier version of this article can be found on bioRxiv (DOI: <https://doi.org/10.1101/537951>).

References

- Andersen SL: **Trajectories of brain development: point of vulnerability or window of opportunity?** *Neurosci Biobehav Rev.* 2003; **27**(1–2): 3–18. [PubMed Abstract](#) | [Publisher Full Text](#)
- Bos MGN, Peters S, van de Kamp FC, et al.: **Emerging depression in adolescence coincides with accelerated frontal cortical thinning.** *J Child Psychol Psychiatry.* 2018; **59**(9): 994–1002. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Button KS, Ioannidis JP, Mokrysz C, et al.: **Power failure: why small sample size undermines the reliability of neuroscience.** *Nat Rev Neurosci.* 2013; **14**(5): 365–376. [PubMed Abstract](#) | [Publisher Full Text](#)
- Casey B, Jones RM, Somerville LH: **Braking and accelerating of the adolescent brain.** *J Res Adolesc.* 2011; **21**(1): 21–33. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Chan SW, Sussmann JE, Romaniuk L, et al.: **Deactivation in anterior cingulate cortex during facial processing in young individuals with high familial risk and early development of depression: fMRI findings from the Scottish Bipolar Family Study.** *J Child Psychol Psychiatry.* 2016; **57**(11): 1277–1286. [PubMed Abstract](#) | [Publisher Full Text](#)
- Chen BH, Marioni RE, Colicino E, et al.: **DNA methylation-based measures of biological age: meta-analysis predicting time to death.** *Aging (Albany NY).* 2016; **8**(9): 1844–1865. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Cole JH, Marioni RE, Harris SE, et al.: **Brain age and other bodily 'ages': implications for neuropsychiatry.** *Mol Psychiatry.* 2019; **24**(2): 266–281. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Cole JH, Ritchie SJ, Bastin ME, et al.: **Brain age predicts mortality.** *Mol Psychiatry.* 2018; **23**(5): 1385–1392. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Dahl RE: **Adolescent brain development: a period of vulnerabilities and opportunities. Keynote address.** *Ann N Y Acad Sci.* 2004; **1021**: 1–22. [PubMed Abstract](#) | [Publisher Full Text](#)
- de Girolamo G, Dagan J, Purcell R, et al.: **Age of onset of mental disorders and use of mental health services: needs, opportunities and obstacles.** *Epidemiol Psychiatr Sci.* 2012; **21**(1): 47–57. [PubMed Abstract](#) | [Publisher Full Text](#)
- de Nooij L: **Brain age trajectories and mood disorders (SBFS).** 2020. <https://www.doi.org/10.17605/OSF.IO/QKCYD>
- Drevets WC, Price JL, Furey ML: **Brain structural and functional abnormalities in mood disorders: implications for neurocircuitry models of depression.** *Brain Struct Funct.* 2008; **213**(1–2): 93–118. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Ducharme S, Albaugh MD, Hudziak JJ, et al.: **Anxious/depressed symptoms are linked to right ventromedial prefrontal cortical thickness maturation in healthy children and young adults.** *Cereb Cortex.* 2014; **24**(11): 2941–2950. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- First MB, Spitzer RL, Gibbon M, et al.: **Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition with Psychotic Screen.** *Biometrics Research.* New York State Psychiatric Institute, New York. 2002.
- Franke K, Luders E, May A, et al.: **Brain maturation: predicting individual BrainAGE in children and adolescents using structural MRI.** *Neuroimage.* 2012; **63**(3): 1305–1312. [PubMed Abstract](#) | [Publisher Full Text](#)
- Franke K, Ziegler G, Klöppel S, et al.: **Estimating the age of healthy subjects from T₁-weighted MRI scans using kernel methods: exploring the influence of various parameters.** *Neuroimage.* 2010; **50**(3): 883–892. [PubMed Abstract](#) | [Publisher Full Text](#)
- Franke K, Gaser C: **Longitudinal Changes in Individual BrainAGE in Healthy Aging, Mild Cognitive Impairment, and Alzheimer's Disease.** *GeroPsych.* 2012; **25**(4): 235–245. [PubMed Abstract](#) | [Publisher Full Text](#)
- Ganzola R, McIntosh AM, Nickson T, et al.: **Diffusion tensor imaging correlates of early markers of depression in youth at high-familial risk for bipolar disorder.** *J Child Psychol Psychiatry.* 2018; **59**(8): 917–927. [PubMed Abstract](#) | [Publisher Full Text](#)
- Gaser C, Franke K, Klöppel S, et al.: **BrainAGE in Mild Cognitive Impaired Patients: Predicting the Conversion to Alzheimer's Disease.** *PLoS One.* 2013; **8**(6): e67346. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Giedd JN, Blumenthal J, Jeffries NO, et al.: **Brain development during childhood and adolescence: a longitudinal MRI study.** *Nat Neurosci.* 1999; **2**(10): 861–863. [PubMed Abstract](#) | [Publisher Full Text](#)
- Giorgio A, Watkins KE, Chadwick M, et al.: **Longitudinal changes in grey and white matter during adolescence.** *Neuroimage.* 2010; **49**(1): 94–103. [PubMed Abstract](#) | [Publisher Full Text](#)
- Gogtay N, Giedd JN, Lusk L, et al.: **Dynamic mapping of human cortical**

- development during childhood through early adulthood. *Proc Natl Acad Sci U S A*. 2004; **101**(21): 8174–8179.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Gueorguieva R, Krystal JH: **Move over ANOVA: progress in analyzing repeated-measures data and its reflection in papers published in the Archives of General Psychiatry.** *Arch Gen Psychiatry*. 2004; **61**(3): 310–317.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Hajek T, Franke K, Kolenic M, et al.: **Brain Age in Early Stages of Bipolar Disorders or Schizophrenia.** *Schizophr Bull*. 2019; **45**(1): 190–198.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Hamilton M: **A rating scale for Depression.** *J Neurol Neurosurg Psychiatry*. 1960; **23**(1): 56–62.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Han LK, Dinga R, Hahn T, et al.: **Brain Aging in Major Depressive Disorder: Results from the ENIGMA Major Depressive Disorder working group.** *bioRxiv [pre-print]*. 2019; 1–33.
[Publisher Full Text](#)
- Holm S: **A simple sequentially rejective multiple test procedure.** *Scand J Stat*. 1979; **6**: 65–70.
[Publisher Full Text](#)
- Horvath S: **DNA methylation age of human tissues and cell types.** *Genome Biol*. 2013; **14**(10): R115.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Jollans L, Boyle R, Artiges E, et al.: **Quantifying performance of machine learning methods for neuroimaging data.** *NeuroImage*. 2019; **199**: 351–365.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kessler RC, Bromet EJ: **The epidemiology of depression across cultures.** *Annu Rev Public Health*. 2013; **34**: 119–38.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Koutsouleris N, Davatzikos C, Borgwardt S, et al.: **Accelerated brain aging in schizophrenia and beyond: a neuroanatomical marker of psychiatric disorders.** *Schizophr Bull*. 2014; **40**(5): 1140–53.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Mezuk B, Eaton WW, Albrecht S, et al.: **Depression and type 2 diabetes over the lifespan: a meta-analysis.** *Diabetes Care*. 2008; **31**(12): 2383–90.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Nelson HE, Willison JRT: **National Adult ReadingTest (NART)**, manual (2nd ed.), Test manual, 2nd ed. Windsor: NFER-Nelson. 1991.
[Publisher Full Text](#)
- Nenadić I, Dietzek M, Langbein K, et al.: **BrainAGE score indicates accelerated brain aging in schizophrenia, but not bipolar disorder.** *Psychiatry Res Neuroimaging*. 2017; **266**: 86–89.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Nolen-Hoeksema S, Wisco BE, Lyubomirsky S: **Rethinking rumination.** *Perspect Psychol Sci*. 2008; **3**(5): 400–424.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Osby U, Brandt L, Correia N, et al.: **Excess mortality in bipolar and unipolar disorder in Sweden.** *Arch Gen Psychiatry*. 2001; **58**(9): 844–850.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Pan A, Sun Q, Okereke OI, et al.: **Depression and risk of stroke morbidity and mortality: A meta-analysis and systematic review.** *JAMA*. 2011; **306**(11): 1241–1249.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Papmeyer M, Giles S, Sussmann JE, et al.: **Cortical thickness in individuals at high familial risk of mood disorders as they develop Major Depressive Disorder.** *Biol Psychiatry*. 2015a; **78**(1): 58–66.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Papmeyer M, Sussmann JE, Hall J, et al.: **Neurocognition in individuals at high familial risk of mood disorders with or without subsequent onset of depression.** *Psychol Med*. 2015b; **45**(15): 3317–3327.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Papmeyer M, Sussmann JE, Stewart T, et al.: **Prospective longitudinal study of subcortical brain volumes in individuals at high familial risk of mood disorders with or without subsequent onset of depression.** *Psychiatry Res Neuroimaging*. 2016; **248**: 119–125.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Pedregosa F, Varoquaux G, Gramfort A, et al.: **Scikit-learn: Machine Learning in Python.** *J Mach Learn Res*. 2011; **12**: 2825–2830.
[Reference Source](#)
- Phillips ML, Ladouceur CD, Drevets WC: **A neural model of voluntary and automatic emotion regulation: implications for understanding the pathophysiology and neurodevelopment of bipolar disorder.** *Mol Psychiatry*. 2008; **13**(9): 829–857.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Rizzo LB, Costa LG, Mansur RB, et al.: **The theory of bipolar disorder as an illness of accelerated aging: Implications for clinical care and research.** *Neurosci Biobehav Rev*. 2014; **42**: 157–169.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Scahill RI, Frost C, Jenkins R, et al.: **A Longitudinal Study of Brain Volume Changes in Normal Aging Using Serial Registered Magnetic Resonance Imaging.** *Arch Neurol*. 2003; **60**(7): 989–994.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Shaw P, Kabani NJ, Lerch JP, et al.: **Neurodevelopmental trajectories of the human cerebral cortex.** *J Neurosci*. 2008; **28**(14): 3586–3594.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Sibille E: **Molecular aging of the brain, neuroplasticity, and vulnerability to depression and other brain-related disorders.** *Dialogues Clin Neurosci*. 2013; **15**(1): 53–65.
[PubMed Abstract](#) | [Free Full Text](#)
- Smith SM, Vidaurre D, Alfaro-Almagro F, et al.: **Estimation of Brain Age Delta from Brain Imaging.** *Neuroimage*. 2019; **200**: 528–539.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Smoller JW, Finn CT: **Family, twin, and adoption studies of bipolar disorder.** *Am J Med Genet C Semin Med Genet*. 2003; **123C**(1): 48–58.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Sotiras A, Resnick SM, Davatzikos C: **Finding imaging patterns of structural covariance via Non-Negative Matrix Factorization.** *Neuroimage*. 2015; **108**: 1–16.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Sotiras A, Toledo JB, Gur RE, et al.: **Patterns of coordinated cortical remodeling during adolescence and their associations with functional specialization and evolutionary expansion.** *Proc Natl Acad Sci U S A*. 2017; **114**(13): 3527–3532.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Spear LP: **The adolescent brain and age-related behavioral manifestations.** *Neurosci Biobehav Rev*. 2000; **24**(4): 417–63.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Sprooten E, Sussmann JE, Clugston A, et al.: **White matter integrity in individuals at high genetic risk of bipolar disorder.** *Biol Psychiatry*. 2011; **70**(4): 350–356.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Tamnes CK, Ostby Y, Fjell AM, et al.: **Brain maturation in adolescence and young adulthood: regional age-related changes in cortical thickness and white matter volume and microstructure.** *Cereb Cortex*. 2010; **20**(3): 534–548.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Tipping ME: **Sparse Bayesian Learning and the Relevance Vector Machine.** *J Mach Learn Res*. 2001; 211–244.
[Reference Source](#)
- Varikuti DP, Genov S, Sotiras A, et al.: **Evaluation of non-negative matrix factorization of grey matter in age prediction.** *Neuroimage*. 2018; **173**: 394–410.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Varoquaux G, Raamana RP, Engemann DA, et al.: **Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines.** *Neuroimage*. 2017; **145**(Pt B): 166–179.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Whalley HC, Sussmann JE, Romaniuk L, et al.: **Dysfunction of emotional brain systems in individuals at high risk of mood disorder with depression and predictive features prior to illness.** *Psychol Med*. 2015; **45**(6): 1207–1218.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Whittle S, Lichten R, Dennison M, et al.: **Structural brain development and depression onset during adolescence: A prospective longitudinal study.** *Am J Psychiatry*. 2014; **171**(5): 564–571.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Wierenga LM, Langen M, Oranje B, et al.: **Unique developmental trajectories of cortical thickness and surface area.** *Neuroimage*. 2014; **87**: 120–126.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Wierenga LM, van den Heuvel MP, van Dijk S, et al.: **The development of brain network architecture.** *Hum Brain Mapp*. 2016; **37**(2): 717–729.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Wolkowitz OM, Reus VI, Mellon SH: **Of sound mind and body: Depression, disease, and accelerated aging.** *Dialogues Clin Neurosci*. 2011; **13**(1): 25–39.
[PubMed Abstract](#) | [Free Full Text](#)
- World Health Organization: **Depression and Other Common Mental Disorders: Global Health Estimates**. 2017.
[Reference Source](#)
- Young R, Biggs J, Ziegler V, et al.: **Young Mania Rating Scale.** In: American Psychiatric Association (Ed.), *Handbook of Psychiatric Measures*. American Psychiatric Association, Washington DC, 2000; 540–542.
[Reference Source](#)

Open Peer Review

Current Peer Review Status:   

Version 3

Reviewer Report 11 November 2020

<https://doi.org/10.21956/wellcomeopenres.18057.r41346>

© 2020 Fullerton J. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Janice Fullerton 

Neuroscience Research Australia, Sydney, NSW, Australia

I thank the authors for addressing the concerns I raised in my peer review report, and including the additional supplementary graphs which clearly indicate no effect of MRI quality on brain-PAD.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Young people at increased genetic risk of bipolar disorder

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 11 November 2020

<https://doi.org/10.21956/wellcomeopenres.18057.r41345>

© 2020 Mitchell P. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Philip B. Mitchell 

School of Psychiatry, University of New South Wales, Sydney, NSW, Australia

I approve this revision.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Young people at increased genetic risk of bipolar disorder

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 2

Reviewer Report 14 September 2020

<https://doi.org/10.21956/wellcomeopenres.17743.r39781>

© 2020 Mitchell P. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Philip B. Mitchell 

School of Psychiatry, University of New South Wales, Sydney, NSW, Australia

The authors have addressed my concerns. No further work required.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Young people at increased genetic risk of bipolar disorder

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 08 September 2020

<https://doi.org/10.21956/wellcomeopenres.17743.r40196>

© 2020 Fullerton J. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Janice Fullerton 

Neuroscience Research Australia, Sydney, NSW, Australia

This work explores an interesting and topical research area in a cohort of young individuals at risk of mental illness, a resource the likes of which are scarcely available elsewhere. Structural imaging data are available at two timepoints approximately 2 years apart, while clinical outcome data is available across a maximum of 6 years follow-up enabling the stratification of participants on the basis of clinical outcomes (including outcomes which develop beyond the timepoint of the second MRI scan). Overall this is an interesting and unique study which is well written and of interest to the field.

I was slightly disappointed by the application of cross-sectional image processing for this longitudinally acquired sample. Some of the “considerable heterogeneity in the direction, size and emergence of the individual brain-PADs” (raised by the authors in discussion) across the two timepoints could potentially be ameliorated by longitudinal processing, which I hope will be applied to this dataset in future. I note that the linear mixed models employed here did not take into account the time interval between scans, which while acquired ~2 years apart on average, there is variability in this timing between scans (the range was 1.0-3.8 years, $\text{mean} \pm \text{SD} = 2.12 \pm 0.35$ according to the supplied supplementary datafiles) which together with the lack of inclusion of baseline age in the main exploratory model and non-linear maturation trajectories not being assessed, may impact the interpretation of the findings. Notwithstanding, the analysis reported here using these simpler statistical models are valid and provide interesting new data worthy of publication, as long as the limitations are clear. Incorporation of more complex statistical models which allow the simultaneous modelling of age, age², time (between scans), and their interactions with diagnostic group – plus incorporation of time of diagnosis (in relation to date of scan acquisition) – will allow for greater resolution of the biological processes underpinning cortical trajectory differences in future analyses. This may be particularly important given the relationship between age at baseline and brain-PAD observed here, which suggests that the differences in cortical trajectories are more apparent at earlier time points.

The discussion describes and appropriately acknowledges the limitations of the study, particularly in regards to the low correlation with chronological age, the potential of medication effects (not examined but small exposed group), the potential for increased error in the HR-MD group due to exclusion of those individuals in the training set, and model transferability, the low resolution 1.5T scanner.

In regards to the aging model, there are a few limitations of which to be particularly cognisant. The aging model developed employed training on scans which were also used as targets for prediction, though I note that leave-one-out prediction was employed (for HR-well and C-well scans which were employed as training data) to ameliorate circularity of the model prediction. Of course, in a perfect world, it would be much better to train the model using an independent dataset to ensure the training and test sets are unique and to reduce potential for bias. However employing a training dataset acquired on a different scanner/s may also introduce bias and reduce model fit, so there are advantages to the approach taken by the authors in this regard, particularly if no alternative dataset with equivalent acquisition parameters was available for training. The authors' discussion of possible contributors to the low correlation reported here between chronological age and brain-age raised important and relevant points, and I agree that the tight age window (15-30 years) of the training dataset may result in a poorer correlation coefficient than models trained on samples with greater age variability due to subtler age-related brain differences over a smaller developmental window. It would be interesting to see whether the findings reported here are robust with the application of different (independent) brain-age estimators, such as that developed by the ENIGMA MDD working group (Han et al, *Mol Psychiatry*. 2020¹), although I do not suggest that as necessary for the publication of the current work. Furthermore, I note that the training set for the current study did not stratify based on gender, which may also contribute to poorer model fit due to sex-differences in brain maturation (nor did it appear to exclude relatives which may introduce artifact due to heritability of brain structures), and may warrant a further note in the discussion.

Can the authors please clarify: 1) whether relatives were included in the training dataset for model

predictor; 2) at what stage the brain-PAD values were transformed to z-scores (i.e. was this done for all values generated for scan 1 and scan 2 simultaneously, or separately for scan 1 and then scan 2 data?); 3) the relationship between CAT12 quality assurance value and accuracy of predicted brain-PAD, particularly in regards to lower bounds for 'satisfactory', 'good' and 'excellent' are 70%, 80%, and 90% respectively.

Minor comment: there are still a few instances of "ageing" rather than "aging" remaining in the document (abstract background, intro para 4 (twice), intro para 8, & Fig 2 title).

References

1. Han LKM, Dinga R, Hahn T, Ching CRK, et al.: Brain aging in major depressive disorder: results from the ENIGMA major depressive disorder working group. *Mol Psychiatry*. 2020. [PubMed Abstract](#) | [Publisher Full Text](#)

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Young people at increased genetic risk of bipolar disorder

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 25 Oct 2020

Laura de Nooij, University of Edinburgh, Edinburgh, UK

We thank the reviewer for the time and effort to evaluate our manuscript and for her valuable comments and suggestions.

Regarding the statistical analysis we acknowledge the additional limitations raised by the reviewer, including cross-sectional image processing, linear age predictors and variation in scanning interval. Future studies should apply more complex and more detailed statistical models, preferably on larger datasets in order to prevent overfitting and decrease the impact of data loss when applying longitudinal processing.

We appreciate that the reviewer acknowledges not only the limitations but also the advantages of our leave-one-out approach. We would like to add that we had also tested a different, externally validated brainage model (brainageR version 1, <https://github.com/james-cole/brainageR>) for this sample. This model was developed using a large adult sample ($n=2001$, age mean age = 36.95 ± 18.12 , age range 18–90 years). This model however appeared more biased than our previous model and showed structural overestimations for our specific sample, and therefore less optimal for our purposes. To be specific, performance metrics with brainageR (as assessed within our model's training sample) were $r = 0.39$ for the correlation between brain age prediction and chronological age; MAE = 8.2 years; M brain age prediction = 30.1 (with sample M age = 22.4), SD brain age prediction = 6.3 years (with sample SD age = 3.0 years). This supports our notion that the demographics of our dataset complicated the use of externally validated models, hence our current approach.

The reviewer furthermore requested clarification on three matters, which we provide below:

- 1) Were relatives included in the training dataset for model predictor? In the model training phase we only left out scans of the subject for whom the brain age was being predicted. So, in some cases the scans of a sibling remained in the training set. We acknowledge this limitation but would also like to note that relatedness within the sample is limited and that we did take it into account within subsequent statistical models (i.e. as random effect).
- 2) At what stage the brain-PAD values were transformed to z-scores? This was done when running the statistical models (i.e. R formula: $\text{lme}(\text{scale}(\text{BrainPAD}) \sim \text{Wave} * \text{Group}, \text{random} = \sim 1 | \text{PedID}/\text{ID}, \text{data}=\dots)$). Given the long data format, scaling was done over all observations (i.e. across timepoints/waves). This is consistent with the approach of a single brain age prediction model (balanced with regard to timepoints). Furthermore, when one would scale separately per timepoint, the different scaling factor per timepoint would complicate interpretation.
- 3) What is the relationship between CAT12 quality assurance value and accuracy of predicted brain-PAD, particularly in regards to lower bounds for 'satisfactory', 'good' and 'excellent' are 70%, 80%, and 90% respectively? This is an interesting question to raise. We have now explored the relationship between Brain-PAD and image quality. We have created plots, which can be viewed in the OSF repository (<https://osf.io/pz2x5/>) under 'Supplementary information' > 'Plots_peer_review'. The plots and corresponding statistics show no correlation between image quality and Brain-PAD (T1: $r = 0.056$, $p = 0.42$; T2: $r = -0.086$, $p = 0.32$). To elaborate further, visual inspection suggest a minor, non-statistically significant trend in which higher image quality is associated with slightly more accurate brain age predictions / less bias, because (i) at the right end the regression lines approach brain-PAD = 0, and (ii) the SD of brain-PAD appears smaller when image quality is better.

This logically reflects slightly (not significantly) more noise in the model with lower quality images.

As described above, we have updated our OSF repository by adding additional plots (and code for these plots). Furthermore, following the minor comment of inconsistent spelling (ageing/aging) we have updated our manuscript accordingly.

Competing Interests: No competing interests were disclosed.

Reviewer Report 12 August 2020

<https://doi.org/10.21956/wellcomeopenres.17743.r39782>

© 2020 Singh M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Manpreet K. Singh 

Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, USA

Thank you for addressing reviewer comments.

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Partly

Are sufficient details of methods and analysis provided to allow replication by others?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Imaging in youth at risk for mood disorders

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 27 May 2020

<https://doi.org/10.21956/wellcomeopenres.17108.r38282>

© 2020 Singh M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Manpreet K. Singh**

Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, USA

Summary: This is an interesting and well-written article examining structural trajectories in brain age in young individuals at familial risk for mood disorder. That the study was insufficiently sampled to assess the contribution of family history as an independent factor is a significant limitation to this study. Given that the sample involves young people with a family history of bipolar disorder, consider that as a more accurate representation of the title.

Methods: Please specify more clearly in the methods that the time to follow up is 2 years. The design of the study is confusing and circuitous that it doesn't become apparent until the results that the time elapsed between baseline and follow up scans is 2 years. The difference between a 15 year old two years later is significantly different than a 30 year old 2 years later in terms of brain maturation. By which standard is 15-30 years of age considered to be a "tight" age range?

Methods: Please clarify the temporal reference for group classifications – is it at 2 year follow up or at 6 year follow up? Please provide justification for approach.

Methods: Was mania severity assessed in this sample on clinical follow up?

Results/Discussion: The HAMD scores upon follow up in the individuals categorized as HR-MD are quite low and in the upper limit of the mild range, with a mean that doesn't even reach the lower bound of the mild range. How do authors conceptualize these individuals relative to the healthy individuals in terms of clinical staging in the context of brain age? Is deceleration of brain age likely to be associated with an increased risk for clinical progression or precede early clinical progression? What kind of study design would be able to tease this apart? It would be helpful to include rational future directions beyond increasing sample size.

Scanning at 1.5 tesla significantly limits spatial resolution and should be noted as a limitation.

Minor comment: unless specified by the journal style, the term "ageing" spelled in this way is

distracting.

Is the work clearly and accurately presented and does it cite the current literature?

No

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

No

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

No

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Imaging in youth at risk for mood disorders

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 14 Jul 2020

Laura de Nooij, University of Edinburgh, Edinburgh, UK

We thank the reviewer for her time and effort in reviewing our article and her constructive comments. Based on these comments, we have improved the manuscript and have uploaded a revised version of this article. Here, we would also like to address the comments on a point-by-point basis.

"This is an interesting and well-written article examining structural trajectories in brain age in young individuals at familial risk for mood disorder. That the study was insufficiently sampled to assess the contribution of family history as an independent factor is a significant limitation to this study."

As above, we are happy that these particular limitations of the study are clearly portrayed and our rationale transparent. In particular, we provide detailed explanations for the choice of within-sample prediction, our concerns about just using the control sample alone, further, we thoroughly tested performance of the model and overfitting per group, we tested other available algorithms, and tested the method against other widely available

brain age prediction software derived in independent samples (BrainAge). We therefore consider that though the study has limitations, and we acknowledge the need for replication, it represents an important foundation for further research in the field.

"Given that the sample involves young people with a family history of bipolar disorder, consider that as a more accurate representation of the title"

We agree with the reviewer and have added the cohort name to the title of the manuscript for clarity, this now reads "Longitudinal trajectories of brain age in young individuals at familial risk of mood disorder from the Scottish Bipolar Family Study"

"Methods: Please specify more clearly in the methods that the time to follow up is 2 years. The design of the study is confusing and circuitous that it doesn't become apparent until the results that the time elapsed between baseline and follow up scans is 2 years. The difference between a 15 year old two years later is significantly different than a 30 year old 2 years later in terms of brain maturation. By which standard is 15-30 years of age considered to be a "tight" age range?"

We have now clarified in the methods that the time between scans was ~ 2 years. We also agree with the reviewer that the change over time for a 15 year old will be different than that for someone aged 30 years. For this reason, we performed an age interaction model, investigating the potential effect of age on these relationships. These results were initially included as exploratory findings in supplemental information and indeed demonstrated larger decelerations for younger participants in the HR-MD group. These results are now detailed in the main body of the manuscript.

"Methods: Please clarify the temporal reference for group classifications - is it at 2 year follow up or at 6 year follow up? Please provide justification for approach."

This information has now been added to the manuscript. In most cases individuals were seen every 2 years for a total of 6 years (x3 assessments, but only 2 scans)

For the control group, individuals were included if they had remained without being diagnosed with any psychiatric disorder throughout the period of the study (via assessments and/or GP records). Those that did become unwell were a small group (n=12) and are not included in the current analysis. This approach was used to reduce heterogeneity in the control group.

For the high-risk group, none met clinical criteria for any mood disorder at baseline. If at any assessment they did meet criteria for a mood disorder (at the time of assessment or over the intervening period since the earlier assessment) they were considered as being in the HR-MD group. So, once someone had a diagnosis, even though they may not have been actively symptomatic at the time of the assessment, they were still considered as being in the mood disorder group. This approach was based on the premise that once an individual had met diagnostic criteria at any time over the course of the study, they were no longer 'high risk' for mood disorder, but actually a 'case'.

"Methods: Was mania severity assessed in this sample on clinical follow up?"

Yes, mania was also assessed at every assessment using the Young Mania Rating Scale (YMRS), however previous studies indicated that there were no significant differences between the three groups in terms of the YMRS cross sectionally or over time, and the median and interquartile ranges were low, hence we did not specifically examine mania ratings in the current analysis (Pappmeyer et al, Psychiatry Res. 2016 Feb 28; 248: 119–125; Pappmeyer et al, Psychol Med. 2015 Nov; 45(15): 3317–3327.). This information has now been added to the manuscript.

"Results/Discussion: The HAMD scores upon follow up in the individuals categorized as HR-MD are quite low and in the upper limit of the mild range, with a mean that doesn't even reach the lower bound of the mild range. How do authors conceptualize these individuals relative to the healthy individuals in terms of clinical staging in the context of brain age? Is deceleration of brain age likely to be associated with an increased risk for clinical progression or precede early clinical progression? What kind of study design would be able to tease this apart? It would be helpful to include rational future directions beyond increasing sample size."

The reviewer is correct that the HAMD scores are relatively low, we however note that the depression status was not defined according to the severity of current symptoms at the time of the scan, but based on lifetime depression, a note to this effect has been added to the discussion.

Regarding the second point, it is not possible to disentangle whether the differences in brain age are a cause or consequence of clinical progression. In future studies, earlier scans prior to prodromal phases of disease, along with leveraging opportunities from other statistical approaches to determine causal directionality could be implemented to attempt to understand these relationships. This has also now been added to the discussion.

"Scanning at 1.5 tesla significantly limits spatial resolution and should be noted as a limitation."

This is now explicitly mentioned as a limitation in the manuscript. As above we also now acknowledge that though this is not ideal given current technology, the choice of scanner strength was determined at the beginning of this 10 year prospective study. We considered the value of having prospective longitudinal data, together with detailed QC (using software quality assurance metrics in combination with manual checks excluding scans with artefacts or of otherwise insufficient quality) provided sufficient confidence in the data quality used in the current study.

"Minor comment: unless specified by the journal style, the term "ageing" spelled in this way is distracting."

All spellings of the term "ageing" have now been changed to "aging"

Competing Interests: No competing interests

Reviewer Report 20 January 2020

<https://doi.org/10.21956/wellcomeopenres.17108.r37434>

© 2020 Mitchell P. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Philip B. Mitchell 

School of Psychiatry, University of New South Wales, Sydney, NSW, Australia

This is the first study of cross-sectional and longitudinal changes in brain aging in young people at high genetic risk of bipolar disorder. As such, it is of interest to the field.

The main result is the somewhat counter-intuitive finding that those at risk who go on to develop a mood disorder demonstrate reduced ('decelerated') brain aging. This contrasts with findings in older patients in which depression is related to increased brain aging. The explanation for this unexpected finding is less than convincing.

The authors are frank about the limitations of this study, the major one being that the 'training' dataset was a combination of controls who stayed well, and at risk subjects who stayed well - thereby confounding risk and clinical status.

Issues to be addressed:

1. There needs to be more frank acknowledgement that the brain aging difference in the risk group that developed mood episodes may have been a result of the depression in addition to the risk state per se.
2. The study used a 1.5T MRI scanner. Is it possible that this did not provide sufficient resolution for such a study?

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

No source data required

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Young people at increased genetic risk of bipolar disorder

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 14 Jul 2020

Laura de Nooij, University of Edinburgh, Edinburgh, UK

We thank the reviewer for his time and effort in reviewing our article and his supportive and positive comments. Based on these comments, we have improved the manuscript and have uploaded a revised version of this article. Here, we would also like to address the comments on a point-by-point basis.

"The main result is the somewhat counter-intuitive finding that those at risk who go on to develop a mood disorder demonstrate reduced ('decelerated') brain aging. This contrasts with findings in older patients in which depression is related to increased brain aging. The explanation for this unexpected finding is less than convincing."

We have now expanded on this finding and simplified the discussion, adding that the majority of studies reporting accelerated brain aging in mood disorder are conducted in older adult samples, rather than younger adolescent cohorts. Given the non-linear trajectory of regional brain maturation / aging over the life-course, particularly over periods where there is significant developmental change (e.g. adolescence), it may not be the case that there is a simple retrospective linear trajectory of accelerated biological aging from these adult studies to periods earlier in life. Further, this finding fits with wider theories of adolescent psychopathology, for example the 'dual system' model where delayed maturation ('decelerated brain age') in higher order cortical regions in relation to limbic regions is proposed to underlie difficulties in emotional regulation, cognitive function and social behaviour, increasing vulnerability to mood disorder.

"The authors are frank about the limitations of this study, the major one being that the 'training' dataset was a combination of controls who stayed well, and at risk subjects who stayed well - thereby confounding risk and clinical status."

We are happy that these particular limitations of the study are clearly portrayed and our rationale transparent. In particular, we provide detailed explanations for the choice of within-sample prediction, our concerns about just using the control sample alone, further, we thoroughly tested performance of the model and overfitting per group, we tested other available algorithms, and tested the method against other widely available brain age prediction software derived in independent samples (BrainAge). We therefore consider that though the study has limitations, and we acknowledge the need for replication, it

represents an important foundation for further research in the field.

"There needs to be more frank acknowledgement that the brain aging difference in the risk group that developed mood episodes may have been a result of the depression in addition to the risk state per se."

We have now added a section to the discussion addressing this point, explaining that the brain aging differences may have resulted from depression and risk state in combination, and that we cannot disentangle effects of risk and depressive symptoms in the current study.

"The study used a 1.5T MRI scanner. Is it possible that this did not provide sufficient resolution for such a study?"

We also now acknowledge this point in the limitations section. Though this is not ideal given current technology, the choice of scanner strength was determined at the beginning of this 10 year prospective study. We considered the value of having prospective longitudinal data, together with detailed QC (using software quality assurance metrics in combination with manual checks excluding scans with artefacts or of otherwise insufficient quality) provided sufficient confidence in the data quality used in the current study.

To elaborate from a technical point of view, 1.5T MRI scans indeed provide more noisy data and lower resolution than 3T MRI data. Smoothing is a recommended pre-processing step that reduces the noise on voxel level and was also applied in the pipeline of the current study. Subsequently, in order to appropriately deal with this high-dimensional data, voxels were resampled into greater voxel sizes in order to reduce the number of variables (features) for the machine learning approach (note however that due to previous smoothing, this step did not further reduce the resolution). We find that these recommended pre-processing steps reduce noise and resolution, but that the resolution of our pattern recognition approach is still higher than with standard sMRI analysis approaches in which voxels are often averaged across areas. The last step of the pipeline was the extraction of features with PCA. We thus applied multiple pre-processing steps, and although derived and pre-processed variables will also be more precise when input variables are more precise (e.g. due to higher scanner strength), we would argue that the low scanner strength in itself does not comprise the methods of this study. In addition, we reference several other studies that applied similar approaches also including 1.5T sMRI data (e.g. Franke et al., NeuroImage, 2012, 63, 1305–1312; Gaser et al, PLOS ONE, 2013, 8(6), e67346).

Competing Interests: No competing interests