Word Count
Abstract: 243
Main Text: 3990
2 tables, 4 figures
43 references

Tracking brain development and dimensional psychiatric symptoms in children: a longitudinal population-based neuroimaging study

Ryan L. Muetzel, PhD[1,2], Laura M.E. Blanken, MD, PhD[1,2], Jan van der Ende, MSc[1], Hanan El Marroun, PhD[1,2], Philip Shaw, MD, PhD[3], Gustavo Sudre, PhD[3], Aad van der Lugt, MD, PhD[4], Vincent Jaddoe, MD, PhD[2,5], Frank C. Verhulst, MD, PhD[1], Henning Tiemeier, MD, PhD[1,6], Tonya White, MD, PhD[1,4]

[1]Department of Child and Adolescent Psychiatry, Erasmus MC-Sophia Children's Hospital, Rotterdam, the Netherlands, [2]The Generation R Study Group, Erasmus MC-Sophia Children's Hospital, Rotterdam, the Netherlands, [3]Neurobehavioral Clinical Research Section, Social and Behavioral Research Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA, Erasmus MC, Rotterdam, the Netherlands, [4]Department of Radiology, Erasmus MC, Rotterdam, the Netherlands, [5]Department of Pediatrics, Erasmus MC-Sophia Children's Hospital, Rotterdam, the Netherlands,[6]Department of Epidemiology, Erasmus MC, Rotterdam, the Netherlands.

Keywords: psychopathology, cross-lagged panel model, linear mixed-effects model, diffusion tensor imaging, fiber tractography, brain morphology

Corresponding Author
Tonya White
Department of Child and Adolescent Psychiatry/Psychology
Erasmus MC-Sophia Children's Hospital
Rotterdam, 3000CB
the Netherlands
t.white@erasmusmc.nl

Disclosures:
Dr. Verhulst is a contributing author of the Achenbach System of Empirically Based Assessment (ASEBA), from which he receives remuneration.

All other authors report no financial relationships with commercial interests.

Acknowledgements
This study was supported by the Sophia Children's Hospital Research Foundation (SSWO) Project 639, the Netherlands Organization for Health Research and Development (ZonMw) TOP project number 91211021 (awarded to T.W.), and ZonMw Vici project 016.VICI.170.200 (awarded to H.T.). MRI data acquisition was sponsored in part by the European Community's 7th Framework Program (FP7/2008-2013, 212652). Supercomputing resources were supported by the NWO Physical Sciences Division (Exacte Wetenschappen) and SURFsara (Cartesius compute cluster, www.surfsara.nl). V.J. received support from the European Union's Horizon 2020 Program (633595). The Generation R Study is conducted by the Erasmus Medical Center in close collaboration with the School of Law and Faculty of Social Sciences of the Erasmus University Rotterdam, the Municipal Health Service Rotterdam area, Rotterdam, the Rotterdam Homecare Foundation, Rotterdam and the Stichting Trombosedienst & Artsenlaboratorium Rijnmond (STAR-MDC), Rotterdam. We gratefully acknowledge the contribution of children and parents, general practitioners, hospitals, midwives and pharmacies in Rotterdam. The general design of Generation R Study is made possible by financial support from the Erasmus Medical Center, Rotterdam, the Erasmus University Rotterdam, ZonMw, the Netherlands Organisation for Scientific Research (NWO), and the Ministry of Health, Welfare and Sport.

Abstract

Objective: Psychiatric symptomatology during childhood predicts persistent mental illness later in life. While neuroimaging methodologies are routinely applied cross-sectionally to the study of child and adolescent psychopathology, the nature of the relationship between childhood symptoms and the underlying neurodevelopmental processes remains largely unclear. The current study delineates the longitudinal relationship between childhood psychiatric problems and brain development using a prospective population-based cohort. **Methods:** 845 children participated in the study. Psychiatric symptoms were measured using the parental report Child Behavior Checklist at ages 6 and 10 years. Magnetic resonance imaging data were also collected at two time-points, at 8 and 10 years of age. Cross-lagged panel models and linear mixed-effects models were used to determine the associations between psychiatric symptom ratings and quantitative anatomic and white matter microstructural measures over development. **Results:** Higher externalizing and internalizing symptoms at baseline predicted smaller increases in both subcortical gray matter volume and global fractional anisotropy over time. The reverse relationship did not hold: thus, baseline measures of gray matter and white matter were not significantly related to changes in symptom scores over time. **Conclusions:** Children presenting with behavioral problems at an early age show differential subcortical and white matter development. Most neuroimaging models tend to explain brain differences observed in psychopathology as an underlying (causal) neurobiological substrate. However, the present work suggests that future neuroimaging studies showing effects that are pathogenic in nature should additionally explore the possibility of the downstream effects of psychopathology on the brain.

Introduction

Given that children who experience psychiatric problems at a young age are at an increased risk for impaired functioning and continued psychopathology later in life (1, 2), characterization of any associated neurodevelopmental features is crucial. Neuroimaging offers a unique window into *in vivo* brain development and the associated features of mental illness (3, 4). As the maturation of morphological (5) and white matter microstructural (6) features has been demonstrated with neuroimaging during childhood, the importance of examining emerging psychopathology in the context of typical brain development has been highlighted (7). However, limited information exists on the exact interplay between the emergence of psychiatric problems and corresponding trajectories of macro- and microstructural neurodevelopment.

There is a vast literature on structural neuroimaging studies of psychopathology. Broadly speaking, externalizing disorders, such as attention deficit/hyperactivity disorder, have been frequently tied to anomalies in fronto-striatal/ fronto-cerebellar circuitry (8, 9), and mood and anxiety disorders have been associated with anomalies in cortico-limbic circuitry (10). However, inconsistencies exist and multiple psychiatric disorders have been found to display spatial overlap in alterations across a broad range of anatomical areas, including those with cortico-limbic and cortico-striatal components. While highly informative, there are two notable limitations to much prior work. First, most studies are cross-sectional, limiting the inferences that can be drawn about developmental processes. Specifically, it is unclear whether early neural anomalies are associated with later psychopathology, or if the reverse relationship also holds (i.e. early psychopathology

is tied to later neural anomalies). Longitudinal data, that combines the collection of clinical and imaging data at baseline and follow-up, can help disentangle the temporality of this relation (11). A second major limitation is that most studies have examined clinical samples, comparing cases and controls. However, many psychiatric symptoms exist on a continuum in the general population (12). Larger, population-based studies have demonstrated that symptoms, when considered dimensionally, vary with neurobiological features, lending further support for this framework (13, 14). However, to date, very few studies have examined child psychiatric symptoms along a continuum in relation to longitudinal brain development (15, 16).

The present study examined the bi-directional association of psychiatric problems with longitudinal gray and white matter microstructural development in a large sample of children from the general population. A dimensional approach was applied in the quantification of internalizing and externalizing problems, along with continuous measures of DSM symptom classes. We hypothesized that psychiatric problems along a continuum would be associated with altered anatomic and white matter microstructural development. In order to parse the precise direction of this relationship between brain and behavior over time, a cross-lagged panel model was utilized.

Methods

Participants

The current study is part of the Generation R Study, a population-based cohort study of maternal and child health from fetal life onwards, in Rotterdam, the Netherlands (17). In addition to an age 5-to-6 behavioral assessment (Figure 1, 18), a sub-sample of 1,070 children were recruited for MRI scanning (referred to as time 1)

(19). As part of the age-10 assessment (referred to as time 2), 520 children who had a scan at time 1 also visited our research-dedicated imaging facility for MRI scanning at time 2. Figure 1 outlines the timeline of the various data collection efforts, and the flow chart depicted in Supplemental Figure S1 illustrates the exclusion of MRI data for both time points. The final sample consisted of 845 usable T_1-weighted scans and 715 usable diffusion tensor imaging (DTI) datasets at time 1, and 480 T_1-weighted scans and 361 DTI datasets at time 2. The Medical Ethics Committee of the Erasmus Medical Center approved all study procedures, and all parents provided written informed consent.

Child psychiatric symptom assessment

Child psychopathology was assessed using the Child Behavior Checklist (CBCL). The CBCL is a widely-used 100-item inventory that provides parental report information on a wide array of behavioral problems in young children (20). The instrument is reliable, valid, and has been used internationally (20, 21). We utilized the broadband scales (internalizing and externalizing) and 4 DSM-oriented scales (affective problems and anxiety problems scales, which correspond to DSM internalizing disorders, and the attention deficit/hyperactivity problems and oppositional defiant problems scales, which correspond to DSM externalizing disorders). The DSM-oriented scales were developed to view the rated problems in the context of a formal diagnostic system (20), and have been shown to correspond to actual clinical diagnoses (22). We administered the preschool CBCL/1½-5, even though some children were older than 5 years at time 1, as the children had not yet received any formal schooling. Cronbach's alphas were the same in 5 year-old children and children older than 5 years, indicating problems were reliably measured

in all children (23). At time 2, the children completed the CBCL 6-18 (24). For all analyses, raw scores (square-root transformed) were used to preserve the natural variation in the data from this non-clinical sample, as T-scores require truncation of values. The average age at the time 1 behavioral assessment was 5.9 years, and the average age of assessment at time 2 was 9.7 years. At baseline, the CBCL was administered prior to the MRI scan in all children, and at time 2 a small number of participants (n=9) received the behavioral assessment after the MRI scan. The percentage of participants meeting borderline and clinical cutoff thresholds for the broadband and DSM-oriented scales is presented in Supplemental Table S1.

Image acquisition

Prior to scanning, all children underwent a 30-minute mock scanning session for acclimation to the MR-environment (19). Data were acquired on 3 Tesla General Electric scanners (at time 1: MR750, at time 2: MR750w, GE, Milwaukee, WI). Both systems utilized an 8-channel receive-only head coil. T_1-weighted structural images were acquired with an IR-prepared Fast Spoiled Gradient Recalled sequence. Diffusion MRI data were collected with 3 B=0 volumes and 35 non-colinear diffusion encoded volumes using an echo-planar imaging sequence (see Supplemental Material for details).

Morphological Image Processing

Structural MRI data were processed through the FreeSurfer analysis suite (v5.3, 25). Briefly, non-brain tissue was removed, voxels intensities were normalized for B_1 inhomogeneity, whole-brain tissue segmentation was performed, and a surface-based model of the cortex was reconstructed. Global metrics of volume were

extracted (e.g., total brain volume and subcortical volume), and a number of subcortical and cortical structures were automatically labeled (e.g., amygdala, orbitofrontal cortex, etc.).

Diffusion Image preprocessing

Image preprocessing was conducted using the Functional MRI of the Brain's Software Library (FSL, version 5.0.5, 26) and the Camino Diffusion Toolkit (27). Non-brain tissue was removed, and diffusion images were corrected for eddy current-induced artifacts and translations/rotations resulting from head motion. In order to account for rotations applied to the diffusion data, the resulting transformation matrices were used to rotate the diffusion gradient direction table. The diffusion tensor was fit at each voxel, and common scalar metrics (e.g., fractional anisotropy, mean diffusivity) were subsequently computed.

Probabilistic tractography was run on each subject's diffusion data using the fully automated FSL plugin, "AutoPtx" (see Supplemental Material, 28). Connectivity distributions were estimated for 12 fiber bundles (Figure 2). Using the connectivity distributions, average fractional anisotropy and mean diffusivity values were then computed for each fiber bundle (29).

Image quality assurance

FreeSurfer reconstructions were visually inspected using a protocol similar to previously reported methods (30). Raw and processed diffusion image quality was assessed using a combination of automated and manual methods. Further details of the quality assurance procedure are available in the Supplemental Material, and the flow chart in Supplemental Figure S1 outlines the number of datasets excluded.

Statistical Analysis

Statistical analyses were conducted using the R Statistical Computing software (v3.2.3, 'Wooden Christmas Tree', 31). The Lavaan package (32) was used to fit cross-lagged panel models to study the associations between longitudinal measures of brain and behavior. Cross-lagged panel models allow associations between two or more repeatedly measured variables to be investigated contemporaneously. Figure 3A depicts the general modeling strategy utilized. The first cross-lagged coefficient $\beta_{CL\text{-}1}$ represents the association between psychiatric problems measured at time 1 and brain metrics measured at time 2 that have been adjusted for baseline brain metrics measured at time 1. Similarly, the other cross-lagged coefficient, $\beta_{CL\text{-}2}$, represents the association between brain metrics measured at time 1 and psychiatric problems measured at time 2 that have been adjusted for baseline psychiatric problems measured at time 1. Cross-sectional associations between brain metrics and psychiatric problems are also modeled, though only the association at time 1 (coefficient $\beta_{CL\text{-}Baseline}$) is reported, given the association at time 2 represents the correlation in the residual terms and is not straightforward to interpret. Lastly, autoregressive coefficients $\beta_{AR\text{-}CBCL}$ and $\beta_{AR\text{-}MRI}$, representing the stability of psychiatric problems and brain metrics from time 1-to-time 2, respectively, are modeled. While cross-lagged models are able to provide information on associations that relate to inter-individual variability in two repeatedly measured variables, they do not provide an explicit metric of change. Thus, significant cross-lagged associations were followed up with linear mixed-effects models in order to obtain an explicit interpretation of within-subject change. These models are described in detail in the

Supplemental Material. Further, the correlations between behavioral and MRI metrics for time 1 and time 2 are presented in Supplemental Tables S2 and S3.

A hierarchical approach was imposed where broad/global metrics were first examined using cross-lagged panel models, in order to gain a comprehensive picture of the time 1 and time 2 associations. These models were subsequently followed by more refined metrics, honing in on the associations of interest using linear mixed-effects models (33). First, we tested whether broadband behavioral measures were associated with global MRI metrics (e.g., total brain volume and global fractional anisotropy). If an association was observed with one of the broadband behavioral measures ($p < 0.05$), follow-up associations between the corresponding DSM-oriented sub-scales and MRI metrics were then tested. This approach was utilized in order to determine whether specific psychiatric traits account for any observed association between a broadband scale and brain metrics. Second, along similar lines, in order to determine whether the effects on the brain were global (i.e., wide-spread in the brain) or focal (i.e., limited to a particular set of regions/tracts), behavioral broadband scores showing a relationship with global MRI metrics were also tested for associations with changes in MRI metrics from more focal regions of interest. When examining individual regions rather than global metrics, each region of interest was z-transformed within the time-point to control for confounding effects of MR-scanner. Further, given the number of statistical tests examined with individual regions of interest, a false discovery rate (FDR) correction was applied to control for Type-I error (34).

Global macro- and microstructural brain metrics

The anatomic metrics used were provided by FreeSurfer and included total brain, total cortical, subcortical, lateral ventricular and white matter volume. For DTI data, rather than computing a simple average, latent factors were used to represent global DTI metrics (i.e., across multiple tracts) within the cross-lagged panel model (29, 32). Thus, all tract metrics were summarized by a single, 'global' factor. For each DTI scalar metric, the tracts depicted in Figure 2 were used to model the latent factors. Latent global DTI factors were modeled separately for time 1 and for time 2, given the two waves were acquired on different MR-scanners, and were normally distributed. A visual depiction of the standardized factor loadings for the time 1 and time 2 global fractional anisotropy metric is available in Supplemental Figure S2.

Covariates

All models were adjusted for age at assessment (behavioral and MRI, as the two assessments were conducted at different times), sex, and ethnicity (reference-coded with the Dutch population as the reference group).

Results

Sample Characteristics

Table 1 presents information on the characteristics of the sample. Children were approximately 8 years of age at the time 1 MRI, and 10 years of age at the time 2 MRI.

Neuroanatomical Macrostructure

Table 2 outlines associations between global cortical brain metrics and psychiatric problem scores. Coefficient labels in the model illustration (see Figure

3A) correspond to the headings used in Table 2. Cross-sectionally at baseline, higher broadband externalizing scores were associated with smaller total brain volume, cortical gray matter volume, white matter volume and subcortical volume (Table 2). Higher externalizing scores at time 1 were related to smaller subcortical volumes at time 2, after adjusting for time 1 volumes (Table 2, Figure 3B). Baseline externalizing scores predicted smaller increases in subcortical volume over time with linear mixed-effects models, and results remained consistent after adjusting for intracranial volume (Supplemental Table S4). Interestingly, the path testing whether baseline neuroanatomical features predicted time 2 externalizing scores was non-significant (Table 2).

For broadband internalizing scores, cross-sectionally at baseline there were no significant associations with any of the macrostructural features. However, higher internalizing scores at time 1 were associated with smaller subcortical volumes at time 2, after adjusting for volumes at time 1 (Table 2). Similar to externalizing, the subcortical association was consistent in linear mixed-effects models, even after further adjusting for intracranial volume (Supplemental Table S4). The path testing associations between global metrics of cortical morphology at time 1 and internalizing scores at time 2 was non-significant (Table 2).

In order to better characterize the regionally specificity of the significant subcortical volume analyses outlined above, linear mixed-effects models were run predicting change in the individual regions that comprise the total subcortical volume. Following adjustment for multiple comparisons, psychiatric problems at time 1 were not related to changes in any of the regions over time.

White matter microstructure

Externalizing scores at time 1 were not associated cross-sectionally with global DTI measures at time 1. However, externalizing scores at time 1 were negatively associated with global fractional anisotropy at time 2, after adjusting for global fractional anisotropy at time 1 (Table 2). Linear mixed-effects analyses were consistent, with higher baseline externalizing scores predicting smaller increases in global fractional anisotropy over time (Figure 4, Supplemental Table S5). Baseline global DTI measures were not associated with externalizing scores at time 2.

Internalizing scores at time 1 were also not cross-sectionally associated with global DTI measures at time 1. Broadband internalizing scores were negatively associated with global fractional anisotropy at time 2, after adjusting for global fractional anisotropy at time 1 (Table 2, Figure 3C). Further, linear mixed-effects models showed higher baseline internalizing scores predicted smaller increases in global fractional anisotropy over time (Figure 4, Supplemental Table S5). Similar to what was observed with the neuroanatomical features, time 1 global fractional anisotropy did not predict time 2 internalizing problems (Table 2).

Given associations were observed between time 1 behavioral measures and time 2 DTI metrics, individual tracts were examined to determine whether there was any regional specificity in the white matter associations. Higher externalizing scores at time 1 were associated smaller increases in fractional anisotropy in the superior longitudinal fasciculus (Supplemental Table S6). Higher broadband internalizing scores at time 1 were associated with smaller increases in fractional anisotropy in the right cingulum and bilateral superior longitudinal fasciculus (Supplemental Table S6). Significant associations presented above were adjusted further for potential confounders (e.g., IQ and motion), and results are presented in the Supplemental Data.

DSM-oriented subscales

In order to further characterize the global associations, we examined the individual DSM-oriented subscales when broadband scales showed significant associations with changes in global MRI metrics. Higher baseline scores on both the DSM-oriented attention and oppositional defiant disorder subscales predicted smaller increases in total subcortical volume and global fractional anisotropy over time (Supplemental Table S7). Further, higher baseline affective subscale scores were associated with smaller increases in global fractional anisotropy (Supplemental Table S7).

Stability of psychiatric scores and brain metrics over time

Auto-regressive coefficients in Table 2 demonstrate that psychiatric scores measured at time 1 are positively associated with those measured at time 2 (roughly 0.65 for externalizing and 0.56 for internalizing), suggesting stability in the measure and also some inter-individual variability over time. Similarly, brain metrics at time 1 were positively associated with those measured at time 2, suggesting a relatively high stability though some inter-individual variability over time, particularly in global fractional anisotropy.

Discussion

In this large population imaging study of children, we demonstrate a link between psychiatric problems along a continuum and a differential pattern of brain changes over time. Even in the general population, psychiatric problems were related to altered trajectories of both macro- and microstructural brain development.

Interestingly, baseline brain metrics measured during childhood did not predict changes in psychiatric symptom ratings over time; instead, psychiatric problems at a young age predicted an altered pattern of brain changes over a 2½-year interval.

Consistent with existing literature, this study shows facets of psychopathology are related to smaller subcortical volumes and lower fractional anisotropy. The subcortex has been implicated to some degree in nearly all psychiatric disorders, and brain imaging studies have revealed evidence for its involvement in, for example, depression, attention deficit/hyperactivity disorder, and obsessive compulsive disorder (8, 35, 36). Similarly, numerous studies have also demonstrated the potential role of white matter microstructure in psychopathology (10, 37). The present study expands upon this existing literature by demonstrating that dimensionally-assessed psychiatric problems are related to smaller changes in brain volume and fractional anisotropy over time. Most subcortical structures show increases in volume that peak during late childhood and into adolescence, followed by decreases in volume into adulthood (38). The age-range of the present sample shows an association that is consistent with this, and additionally shows attenuated increases that are related to both externalizing and internalizing problems. Fractional anisotropy in white matter has shown largely linear increases over time within this restricted age-range (6). Data from the present study fit with the literature, again additionally showing that these trajectories are potentially modified by the presence of psychiatric symptoms at an early age. Fractional anisotropy, but not mean diffusivity, showed a differential pattern of change. This could suggest differences in, for example myelination or axonal packing.

Limited specificity was observed, as both externalizing and internalizing symptoms were associated with differential macro- and microstructural changes over time. Further, three of the four DSM scales were associated with changes in white

matter microstructure. As the two broadband domains are relatively highly correlated (20), it is possible that the overlapping variance in these scales best explains the differential changes in the brain. Alternatively, given the relatively young age of this sample, it is also possible that domain-specific patterns of development become more apparent at older ages. In terms of symptom assessment, this study utilized a parental report of children's psychiatric symptoms. It is likely that combining data from multiple informants (i.e., the parents, teachers and children themselves) will provide a more robust and accurate picture of their symptom profile, which may in turn help to elucidate more specific correlates with neuroimaging features (39). Along similar lines, though this study utilized high-resolution anatomical data and diffusion tensor imaging data, the two modalities were examined separately. It is possible that a multimodal approach where both pieces of information are elegantly combined into a single analysis could help to disentangle unique neurobiological features in both categories of disorders (40).

In the context of psychopathology, the most widely applied model in the field of neuroimaging has focused on identifying associated, underlying neurobiological substrates of potential etiological significance. However, we did not find that early brain metrics predicted changes in psychiatric problems over time. There are several potential explanations for this. As this is not a clinical sample, it is possible that such processes are not part of a continuum and are not present in sub-clinical presentation of psychiatric constructs in the general population. It is thus a priority to see if these findings hold in large cohorts that are enriched for psychiatric disorders. Alternatively, the underlying neurobiological predictors of the emergence and development of psychopathology may be spatially or temporally heterogeneous, particularly during brain development, and that more sophisticated image-analysis

methods are needed to better characterize them (e.g., machine learning)(41). It is also possible that the responsible underlying brain features are detectable at higher spatial resolution or perhaps using alternative imaging modalities (e.g., MR-spectroscopy, MR-perfusion, task-based or resting-state functional MRI).

The bulk of the findings in the present study consisted of early measures of psychiatric problems predicting differential patterns of both macro- and microstructural brain development. This suggests that, in addition to the standard model of the '*brain shaping behavior*' discussed above, perhaps '*behaviors also shape the brain'* (42). While there are undoubtedly underlying neurobiological explanations for the emergence of psychopathology in children, the data in the present study suggest psychiatric symptoms in children may also *contribute* to some of the macro- and microstructural abnormalities reported in the literature; a potentially cascading interaction between psychopathology and the brain. Take the example obsessive-compulsive disorder where the symptomatology originates in a particular set of brain areas (e.g., fronto-striatal circuitry), but over time the symptoms themselves could modulate structural brain development (e.g., repetitive motor behaviors leading alterations in motor cortex) (42). A similar extension can be drawn to, in for example anxiety disorders, the potential effect of increased hypothalamic-pituitary-adrenal axis activity on the brain through downstream hormonal exposure (43). A child with psychopathology may also experience his or her environment differently as a result of the disorder, which could consequently influence how the brain is shaped during development (e.g., reduced novelty seeking). The cofounding effect of psychiatric treatment could also be playing a role (i.e., children with more psychiatric symptoms seek treatment, which subsequently influences brain development over time), however additional analyses in this study did not show that

psychiatric medication status explained the association (see Supplementary Data). Future studies should not rule out the possibility that, in addition to a preceding/causal factor, observed neuroimaging associations with a given disorder could also be a downstream effect of the disorder.

Despite the large population-based sample, longitudinal design, and dimensional assessment of psychiatric symptoms, this study has important limitations. First, though data from this study are part of a larger population-based study, representativeness is important to discuss. While many features of this sample are representative of the catchment area (e.g., ethnicity), some potentially are less so. For example, mean non-verbal IQ in this sample was a few points above the population average. Though models adjusted for IQ proved to be similar to unadjusted models (see Supplemental Data), IQ remains an important consideration in child psychiatric studies as it is often intertwined with clinical diagnosis and is related to many neuroimaging features. Next, as a hierarchical approach was implemented, some associations at a more focal/specific level could have been missed and future work will benefit from analyzing brain and behavior at a more focal level. Regarding data acquisition, imaging data were acquired on two separate MRI scanners, possibly introducing problems with the longitudinal interpretation of results. However, acquisitions were made as similar as possible (e.g., gradient table, head coil, etc.) and a number of steps were introduced to mitigate such problems, including within-scanner normalization of MRI metrics (e.g., latent factor modeling within scanner, and Z-score standardization), which should attenuate the effects of scanner difference. Additionally, the statistical models used in the present study to test the association between psychiatric problems and change in macro- and microstructural brain features operate on a relative, rather than absolute, scale further ameliorating concerns

over the effect of scanner. Lastly, only linear models were tested as data were only acquired at two points in time. Gray matter morphological studies have previously shown the importance of nonlinear trajectories, and future work with additional time-points should address this in more detail.

In conclusion, we demonstrate internalizing and externalizing problems are related to altered macro- and microstructural changes in a large sample of children from the general population. Tracking the emergence of psychopathology in children, both in terms of symptomatology and neurobiology, may help guide not only diagnosticians, but also improve the selection and timing of treatments. It is important to appreciate that this study does not challenge the causal role of neural changes in the pathogenesis of psychiatric disorders. Nonetheless, our findings raise the intriguing possibility that some emergent neuroimaging features of psychopathology may be partly a consequence of the disorder.

Figure captions

Figure 1. Timeline of the study data collection points

Figure indicates the age ranges of study participants during each type of assessment

Figure 2. Depiction of the tracts used in the global DTI metric

Note: Tracts are group average representations in standard MNI coordinate space. Blue indicates the cingulum bundle, gray the forceps major, tan the forceps minor, red the inferior longitudinal fasciculus, orange the superior longitudinal fasciculus, and green the uncinate fasciculus.

Figure 3. Cross-lagged panel models.

*A.) The general modeling strategy used for cross-lagged panel models, B.) Cross-lagged panel model where total subcortical volume was associated with broadband externalizing problems, and C.) Cross-lagged panel model where global fractional anisotropy was associated with broadband internalizing problems. FA = Fractional anisotropy and CBCL = Child behavior checklist. *p < 0.05, **p < 0.01*

Figure 4. Association between high and low levels of psychopathology and changes in white matter microstructure

Figure represents predicted model estimates derived from linear mixed effects models. The top panel shows broadband internalizing problems and the bottom panel shows broadband externalizing problems. Separate lines for 1 standard deviation above the mean problem score ("high") and 1 standard deviation below the mean problem score ("low"), with the Y-axis representing the predicted global FA value based on model estimates. Note: Fractional anisotropy is a unit-less measure.

References

1. Hofstra MB, van der Ende J, Verhulst FC. Child and adolescent problems predict DSM-IV disorders in adulthood: a 14-year follow-up of a Dutch epidemiological sample. J Am Acad Child Adolesc Psychiatry. 2002;41:182-189.

2. Pine DS, Cohen E, Cohen P, Brook J. Adolescent depressive symptoms as predictors of adult depression: moodiness or mood disorder? Am J Psychiatry. 1999;156:133-135.

3. Mana S, Paillere Martinot ML, Martinot JL. Brain imaging findings in children and adolescents with mental disorders: a cross-sectional review. Eur Psychiatry. 2010;25:345-354.

4. Giedd JN, Raznahan A, Alexander-Bloch A, Schmitt E, Gogtay N, Rapoport JL. Child psychiatry branch of the National Institute of Mental Health longitudinal structural magnetic resonance imaging study of human brain development. Neuropsychopharmacology. 2015;40:43-49.

5. Lenroot RK, Giedd JN. Brain development in children and adolescents: insights from anatomical magnetic resonance imaging. Neurosci Biobehav Rev. 2006;30:718-729.

6. Schmithorst VJ, Yuan W. White matter development during adolescence as shown by diffusion MRI. Brain Cogn. 2010;72:16-25.

7. Di Martino A, Fair DA, Kelly C, Satterthwaite TD, Castellanos FX, Thomason ME, Craddock RC, Luna B, Leventhal BL, Zuo XN, Milham MP. Unraveling the miswired connectome: a developmental perspective. Neuron. 2014;83:1335-1353.

8. Shaw P, Rabin C. New insights into attention-deficit/hyperactivity disorder using structural neuroimaging. Curr Psychiatry Rep. 2009;11:393-398.

9. Nakao T, Radua J, Rubia K, Mataix-Cols D. Gray matter volume abnormalities in ADHD: voxel-based meta-analysis exploring the effects of age and stimulant medication. Am J Psychiatry. 2011;168:1154-1163.

10. Hulvershorn LA, Cullen K, Anand A. Toward dysfunctional connectivity: a review of neuroimaging findings in pediatric major depressive disorder. Brain Imaging Behav. 2011;5:307-328.

11. Kraemer HC, Yesavage JA, Taylor JL, Kupfer D. How can we learn about developmental processes from cross-sectional studies, or can we? Am J Psychiatry. 2000;157:163-171.

12. Garvey M, Avenevoli S, Anderson K. The National Institute of Mental Health Research Domain Criteria and Clinical Research in Child and Adolescent Psychiatry. J Am Acad Child Adolesc Psychiatry. 2016;55:93-98.

13. Blanken LM, Mous SE, Ghassabian A, Muetzel RL, Schoemaker NK, El Marroun H, van der Lugt A, Jaddoe VW, Hofman A, Verhulst FC, Tiemeier H, White T. Cortical morphology in 6- to 10-year old children with autistic traits: a population-based neuroimaging study. Am J Psychiatry. 2015;172:479-486.

14. Mous SE, Muetzel RL, El Marroun H, Polderman TJ, van der Lugt A, Jaddoe VW, Hofman A, Verhulst FC, Tiemeier H, Posthuma D, White T. Cortical thickness and inattention/hyperactivity symptoms in young children: a population-based study. Psychol Med. 2014;44:3203-3213.

15. Albaugh MD, Ducharme S, Karama S, Watts R, Lewis JD, Orr C, Nguyen TV, McKinstry RC, Botteron KN, Evans AC, Hudziak JJ, Brain Development Cooperative G. Anxious/depressed symptoms are related to microstructural maturation of white matter in typically developing youths. Dev Psychopathol. 2016:1-8.

16. Ducharme S, Albaugh MD, Hudziak JJ, Botteron KN, Nguyen TV, Truong C, Evans AC, Karama S, Brain Development Cooperative G. Anxious/depressed symptoms are linked to right ventromedial prefrontal cortical thickness maturation in healthy children and young adults. Cereb Cortex. 2014;24:2941-2950.

17. Kooijman MN, Kruithof CJ, van Duijn CM, Duijts L, Franco OH, van IMH, de Jongste JC, Klaver CC, van der Lugt A, Mackenbach JP, Moll HA, Peeters RP, Raat H, Rings EH, Rivadeneira F, van der Schroeff MP, Steegers EA, Tiemeier H, Uitterlinden AG, Verhulst FC, Wolvius E, Felix JF, Jaddoe VW. The Generation R Study: design and cohort update 2017. Eur J Epidemiol. 2016;31:1243-1264.

18. Tiemeier H, Velders FP, Szekely E, Roza SJ, Dieleman G, Jaddoe VW, Uitterlinden AG, White TJ, Bakermans-Kranenburg MJ, Hofman A, Van Ijzendoorn MH, Hudziak JJ, Verhulst FC. The Generation R Study: A review of design, findings to date, and a study of the 5-HTTLPR by environmental interaction from fetal life onward. Journal of the American Academy of Child and Adolescent Psychiatry. 2012;51:1119-1135 e1117.

19. White T, El Marroun H, Nijs I, Schmidt M, van der Lugt A, Wielopolki PA, Jaddoe VW, Hofman A, Krestin GP, Tiemeier H, Verhulst FC. Pediatric population-based neuroimaging and the Generation R Study: the intersection of developmental neuroscience and epidemiology. European journal of epidemiology. 2013;28:99-111.

20. Achenbach TM, Rescorla LA: Manual for ASEBA preschool forms & profiles. Burlington, VT, University of Vermont, Reseach Center for Children, Youth & Families; 2000.

21. Ivanova MY, Achenbach TM, Rescorla LA, Harder VS, Ang RP, Bilenberg N, Bjarnadottir G, Capron C, De Pauw SS, Dias P, Dobrean A, Doepfner M, Duyme M, Eapen V, Erol N, Esmaeili EM, Ezpeleta L, Frigerio A, Goncalves MM, Gudmundsson HS, Jeng SF, Jetishi P, Jusiene R, Kim YA, Kristensen S, Lecannelier F, Leung PW, Liu J, Montirosso R, Oh KJ, Plueck J, Pomalima R, Shahini M, Silva JR, Simsek Z, Sourander A, Valverde J, Van Leeuwen KG, Woo BS, Wu YT, Zubrick SR, Verhulst FC. Preschool psychopathology reported by parents in 23 societies: testing the seven-syndrome model of the child behavior checklist for ages 1.5-5. J Am Acad Child Adolesc Psychiatry. 2010;49:1215-1224.

22. van Lang ND, Ferdinand RF, Oldehinkel AJ, Ormel J, Verhulst FC. Concurrent validity of the DSM-IV scales Affective Problems and Anxiety Problems of the Youth Self-Report. Behav Res Ther. 2005;43:1485-1494.

23. Basten MM, Althoff RR, Tiemeier H, Jaddoe VW, Hofman A, Hudziak JJ, Verhulst FC, van der Ende J. The dysregulation profile in young children: empirically defined classes in the Generation R study. J Am Acad Child Adolesc Psychiatry. 2013;52:841-850 e842.

24. Achenbach TM, Rescorla LA: Manual for the ASEBA School-Age Forms & Profiles. Burlington, VT, University of Vermont, Reseach Center for Children, Youth & Families; 2001.

25. Fischl B, van der Kouwe A, Destrieux C, Halgren E, Segonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, Caviness V, Makris N, Rosen B, Dale AM. Automatically parcellating the human cerebral cortex. Cereb Cortex. 2004;14:11-22.

26. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. Fsl. Neuroimage. 2012;62:782-790.

27. Cook PA, Bai Y, Nedjati-Gilani S, Seunarine KK, Hall MG, Parker GJ, Alexander DC: Camino: Open-Source Diffusion-MRI Reconstruction and Processing. in 14th Scientific Meeting of the International Society for Magnetic Resonance in Medicine. Seattle, WA, USA2006. pp. 2759.

28. de Groot M, Ikram MA, Akoudad S, Krestin GP, Hofman A, van der Lugt A, Niessen WJ, Vernooij MW. Tract-specific white matter degeneration in aging: The Rotterdam Study. Alzheimers Dement. 2015;11:321-330.

29. Muetzel RL, Mous SE, van der Ende J, Blanken LM, van der Lugt A, Jaddoe VW, Verhulst FC, Tiemeier H, White T. White matter integrity and cognitive performance in school-age children: A population-based neuroimaging study. Neuroimage. 2015;119:119-128.

30. Hibar DP, Stein JL, Renteria ME, Arias-Vasquez A, Desrivieres S, Jahanshad N, Toro R, Wittfeld K, Abramovic L, Andersson M, Aribisala BS, Armstrong NJ, Bernard M, Bohlken MM, Boks MP, Bralten J, Brown AA, Chakravarty MM, Chen Q, Ching CR, Cuellar-Partida G, den Braber A, Giddaluru S, Goldman AL, Grimm O, Guadalupe T, Hass J, Woldehawariat G, Holmes AJ, Hoogman M, Janowitz D, Jia T, Kim S, Klein M, Kraemer B, Lee PH, Olde Loohuis LM, Luciano M, Macare C, Mather KA, Mattheisen M, Milaneschi Y, Nho K, Papmeyer M, Ramasamy A, Risacher SL, Roiz-Santianez R, Rose EJ, Salami A, Samann PG, Schmaal L, Schork AJ, Shin J, Strike LT, Teumer A, van Donkelaar MM, van Eijk KR, Walters RK, Westlye LT, Whelan CD, Winkler AM, Zwiers MP, Alhusaini S, Athanasiu L, Ehrlich S, Hakobjan MM, Hartberg CB, Haukvik UK, Heister AJ, Hoehn D, Kasperaviciute D, Liewald DC, Lopez LM, Makkinje RR, Matarin M, Naber MA, McKay DR, Needham M, Nugent AC, Putz B, Royle NA, Shen L, Sprooten E, Trabzuni D, van der Marel SS, van Hulzen KJ, Walton E, Wolf C, Almasy L, Ames D,

Arepalli S, Assareh AA, Bastin ME, Brodaty H, Bulayeva KB, Carless MA, Cichon S, Corvin A, Curran JE, Czisch M, de Zubicaray GI, Dillman A, Duggirala R, Dyer TD, Erk S, Fedko IO, Ferrucci L, Foroud TM, Fox PT, Fukunaga M, Gibbs JR, Goring HH, Green RC, Guelfi S, Hansell NK, Hartman CA, Hegenscheid K, Heinz A, Hernandez DG, Heslenfeld DJ, Hoekstra PJ, Holsboer F, Homuth G, Hottenga JJ, Ikeda M, Jack CR, Jr., Jenkinson M, Johnson R, Kanai R, Keil M, Kent JW, Jr., Kochunov P, Kwok JB, Lawrie SM, Liu X, Longo DL, McMahon KL, Meisenzahl E, Melle I, Mohnke S, Montgomery GW, Mostert JC, Muhleisen TW, Nalls MA, Nichols TE, Nilsson LG, Nothen MM, Ohi K, Olvera RL, Perez-Iglesias R, Pike GB, Potkin SG, Reinvang I, Reppermund S, Rietschel M, Romanczuk-Seiferth N, Rosen GD, Rujescu D, Schnell K, Schofield PR, Smith C, Steen VM, Sussmann JE, Thalamuthu A, Toga AW, Traynor BJ, Troncoso J, Turner JA, Valdes Hernandez MC, van 't Ent D, van der Brug M, van der Wee NJ, van Tol MJ, Veltman DJ, Wassink TH, Westman E, Zielke RH, Zonderman AB, Ashbrook DG, Hager R, Lu L, McMahon FJ, Morris DW, Williams RW, Brunner HG, Buckner RL, Buitelaar JK, Cahn W, Calhoun VD, Cavalleri GL, Crespo-Facorro B, Dale AM, Davies GE, Delanty N, Depondt C, Djurovic S, Drevets WC, Espeseth T, Gollub RL, Ho BC, Hoffmann W, Hosten N, Kahn RS, Le Hellard S, Meyer-Lindenberg A, Muller-Myhsok B, Nauck M, Nyberg L, Pandolfo M, Penninx BW, Roffman JL, Sisodiya SM, Smoller JW, van Bokhoven H, van Haren NE, Volzke H, Walter H, Weiner MW, Wen W, White T, Agartz I, Andreassen OA, Blangero J, Boomsma DI, Brouwer RM, Cannon DM, Cookson MR, de Geus EJ, Deary IJ, Donohoe G, Fernandez G, Fisher SE, Francks C, Glahn DC, Grabe HJ, Gruber O, Hardy J, Hashimoto R, Hulshoff Pol HE, Jonsson EG, Kloszewska I, Lovestone S, Mattay VS, Mecocci P, McDonald C, McIntosh AM, Ophoff RA, Paus T, Pausova Z, Ryten M, Sachdev PS, Saykin AJ, Simmons A,

Singleton A, Soininen H, Wardlaw JM, Weale ME, Weinberger DR, Adams HH, Launer LJ, Seiler S, Schmidt R, Chauhan G, Satizabal CL, Becker JT, Yanek L, van der Lee SJ, Ebling M, Fischl B, Longstreth WT, Jr., Greve D, Schmidt H, Nyquist P, Vinke LN, van Duijn CM, Xue L, Mazoyer B, Bis JC, Gudnason V, Seshadri S, Ikram MA, Alzheimer's Disease Neuroimaging I, Consortium C, Epigen, Imagen, Sys, Martin NG, Wright MJ, Schumann G, Franke B, Thompson PM, Medland SE. Common genetic variants influence human subcortical brain structures. Nature. 2015;520:224-229.

31. R Core Team: R: A Language and Environment for Statistical Computing. 3.1.0 ed. Vienna, Austria, R foundation for Statistical Computing; 2014.

32. Rosseel Y. lavaan: An R package for structural equation modeling. Journal of Statistical Software. 2012;48:1-36.

33. Bates D, Maechler M, Bolker B, Walker S. lme4: Linear mixed-effects models using Eigen and S4. https://cranr-projectorg/package=lme4. 2015.

34. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. 1995;57:289-300.

35. Hamilton JP, Siemer M, Gotlib IH. Amygdala volume in major depressive disorder: a meta-analysis of magnetic resonance imaging studies. Mol Psychiatry. 2008;13:993-1000.

36. Boedhoe PS, Schmaal L, Abe Y, Ameis SH, Arnold PD, Batistuzzo MC, Benedetti F, Beucke JC, Bollettini I, Bose A, Brem S, Calvo A, Cheng Y, Cho KI, Dallaspezia S, Denys D, Fitzgerald KD, Fouche JP, Gimenez M, Gruner P, Hanna GL, Hibar DP, Hoexter MQ, Hu H, Huyser C, Ikari K, Jahanshad N, Kathmann N, Kaufmann C, Koch K, Kwon JS, Lazaro L, Liu Y, Lochner C, Marsh R, Martinez-

Zalacain I, Mataix-Cols D, Menchon JM, Minuzzi L, Nakamae T, Nakao T, Narayanaswamy JC, Piras F, Piras F, Pittenger C, Reddy YC, Sato JR, Simpson HB, Soreni N, Soriano-Mas C, Spalletta G, Stevens MC, Szeszko PR, Tolin DF, Venkatasubramanian G, Walitza S, Wang Z, van Wingen GA, Xu J, Xu X, Yun JY, Zhao Q, Group EOW, Thompson PM, Stein DJ, van den Heuvel OA. Distinct Subcortical Volume Alterations in Pediatric and Adult OCD: A Worldwide Meta- and Mega-Analysis. Am J Psychiatry. 2016:appiajp201616020201.

37. White T, Nelson M, Lim KO. Diffusion tensor imaging in psychiatric disorders. Top Magn Reson Imaging. 2008;19:97-109.

38. Raznahan A, Shaw PW, Lerch JP, Clasen LS, Greenstein D, Berman R, Pipitone J, Chakravarty MM, Giedd JN. Longitudinal four-dimensional mapping of subcortical anatomy in human development. Proc Natl Acad Sci U S A. 2014;111:1592-1597.

39. Neumann A, Pappa I, Lahey BB, Verhulst FC, Medina-Gomez C, Jaddoe VW, Bakermans-Kranenburg MJ, Moffitt TE, van IMH, Tiemeier H. Single Nucleotide Polymorphism Heritability of a General Psychopathology Factor in Children. J Am Acad Child Adolesc Psychiatry. 2016;55:1038-1045 e1034.

40. Sui J, Adali T, Yu Q, Chen J, Calhoun VD. A review of multivariate methods for multimodal fusion of brain imaging data. J Neurosci Methods. 2012;204:68-81.

41. Arbabshirani MR, Plis S, Sui J, Calhoun VD. Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls. Neuroimage. 2016.

42. van den Heuvel OA, van Wingen G, Soriano-Mas C, Alonso P, Chamberlain SR, Nakamae T, Denys D, Goudriaan AE, Veltman DJ. Brain circuitry of compulsivity. Eur Neuropsychopharmacol. 2016;26:810-827.

43. Lupien SJ, McEwen BS, Gunnar MR, Heim C. Effects of stress throughout the lifespan on the brain, behaviour and cognition. Nat Rev Neurosci. 2009;10:434-445.