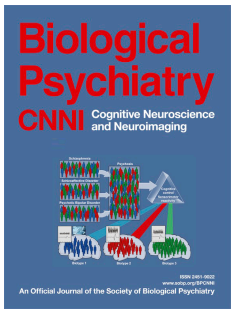
Accepted Manuscript



Stability of cortical thinning in persons at increased familial risk for major depression across eight years

Xuejun Hao, PhD, Ardesheer Talati, PhD, Stewart A. Shankman, PhD, Jun Liu, PhD, Jurgen Kayser, PhD, Craig E. Tenke, PhD, Virginia Warner, DrPh, David Semanek, BA, Priya J. Wickramaratne, PhD, Myrna M. Weissman, PhD, Jonathan Posner, MD

PII: S2451-9022(17)30093-9

DOI: 10.1016/j.bpsc.2017.04.009

Reference: BPSC 156

To appear in: *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*

Received Date: 14 February 2017

Revised Date: 12 April 2017

Accepted Date: 24 April 2017

Please cite this article as: Hao X., Talati A., Shankman S.A, Liu J., Kayser J., Tenke C.E, Warner V., Semanek D., Wickramaratne P.J, Weissman M.M & Posner J., Stability of cortical thinning in persons at increased familial risk for major depression across eight years, *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* (2017), doi: 10.1016/j.bpsc.2017.04.009.

Stability of cortical thinning in persons at increased familial risk for major depression across eight years

Xuejun Hao, PhD[1,2]*, Ardesheer Talati, PhD[1,2]*, Stewart A Shankman, PhD[6], Jun Liu, PhD[1,2], Jurgen Kayser, PhD[1,2], Craig E Tenke, PhD[1,2], Virginia Warner, DrPh[1,2], David Semanek, BA[3], Priya J Wickramaratne, PhD[1,2,5], Myrna M Weissman, PhD[1,2,4], Jonathan Posner, MD[1,3]

[1]Department of Psychiatry, Columbia University Medical Center, New York, NY
Divisions of [2]Epidemiology and [3]Child Psychiatry, New York State Psychiatric Institute, New York, NY
Departments of [4]Epidemiology & [5]Biostatistics, Mailman School of Public Health, Columbia University, New York, NY
[6]Department of Psychology and Psychiatry, University of Illinois at Chicago, Chicago, IL

**These authors contributed equally to this work and share first-authorship; they are listed alphabetically*

Short Title:	8-year stability of cortical thinning in depression
Keywords:	high-risk, biomarker, depression, MRI, Freesurfer, stability
Word Count (Abstract):	244 words
Word Count (Manuscript):	3,384 words
Number of Tables:	2
Number of Figures:	2
Number of Supplementary Files:	1

Corresponding Author:
Ardesheer Talati, Ph.D
Department of Psychiatry,
Columbia University and New York State Psychiatric Institute,
1051 Riverside Dr, Unit 24,
New York, NY 10032
Phone: (646) 774 6421
Email: talatia@nyspi.columbia.edu

Abstract

Background: A biological marker of vulnerability should precede onset of illness and be independent of disease course. We previously reported that cortical thinning may serve as a potential biomarker for risk for familial depression. We now test stability of the cortical thinning across 8 years, and whether thinning mediates associations between familial risk and depressive traits.

Method: Participants were from a 3-generation family study of depression, where 2nd and 3rd generation offspring were characterized as being at high- or low-risk for depression based on the presence/absence of major depression in the 1st generation. The analysis includes 82 offspring with anatomical MRI scans across two assessment waves, 7.8 (S.D.1.3, range: 5.2-10.9) years apart.

Results: High-risk offspring had thinner bilateral superior and middle frontal gyri, and left inferior parietal lobule, at both time-points. High intra-subject correlation ($0.60<r<0.91$) and intra-class correlation (0.72–0.78) of thickness measures across time points was detected within the above regions; rank order by effect size and region was also preserved across time. The thinning was stable despite changes in scanning platform (Siemens Sonata vs. GE Signa), field-strength (1.5 vs. 3T), and participant age and clinical course. Thinning at the first time-point predicted anger and hostility at the second, and mediated the relationship between familial risk and these traits.

Conclusion: The study provides evidence for cortical thinning as a stable biomarker for familial vulnerability for depressive illness, which supports the ability to detect persistent and clinically relevant anatomical findings irrespective of MRI platform.

Introduction

A central goal in psychiatric neuroscience is to identify abnormalities in brain structure and function that predispose to mental illnesses. To represent true markers of risk, such anomalies (also referred to as 'endophenotypes' (1, 2)) should *precede* onset of the disorder itself, be distinct from changes in the brain that occur as a result of the illness, and be stable over time. This last criterion is cumbersome to demonstrate, as it requires individuals to be scanned at more than one time point while keeping methodological variation to a minimum. Most longitudinal studies have focused on tracking *changes* in the brain as a function of illness; few have explicitly tested stability of brain phenotypes.

We previously reported on a potential biomarker for depression vulnerability in a three generation family study of major depressive disorder (MDD) (3). We found that 2nd and 3rd generation offspring at high- compared to low- familial risk for depression (where we defined familial risk based on the presence or absence of MDD in the 1st-generation probands) had thinner cortices, particularly in the lateral surface of the right hemisphere. The thinning was present even in offspring who were at risk, but had never had an episode of MDD, suggesting it was unlikely to be a consequence of the illness. We thus hypothesized that cortical thinning may represent an endophenotype for the familial form of MDD.

What we could not test in our prior report was whether the thinning represented a stable trait, as participants had only been imaged at one time point. We have since rescanned the population, on average 7.8 (range, 5.2 to 10.9) years later, which allows us to examine the extent to which differences in cortical morphology between the high- and low-risk groups are conserved over time. In testing this, we addressed three goals. First, we tested if cortical thinning was still present 8 years later. If so, this stability would increase confidence that cortical thinning reflects a stable biomarker rather than a transient phenotype or a clinical state. Second, we tested whether the cortical thinning was robust to methodological variation, such as the MRI scanning site, platform and magnetic field strength. For the development of biomarkers, it is critical to test potential effects

of methodological advances, and to disentangle heterogeneity attributable to biologically relevant processes from that due to methodological sources. Third, our original findings were reported using proprietary anatomical MRI methods that are not widely available. To foster greater reproducibility, we now use FreeSurfer (4, 5), a freely available, open-source software that is among the most frequently used techniques for examining cortical morphology. Freesurfer is also optimized for longitudinal analyses, as its algorithms allow for greater control over segmentation differences across scans than do other packages (6).

The primary goal of this study was to **test whether differences in cortical thickness related to familial risk for depression are stable over an 8-year period**, where stability was defined by test-retest consistency of anatomy (i.e., do brain regions showing significant differences in cortical thickness at the first time point also do so at the second?) and rank order (do participants with the greatest degree of thinning at the first time point continue to do so at the second?). Finally, to better understand the clinical implications of the cortical thinning, we explored whether cortical thinning identified at the first scan predicted clinical phenotypes relevant to depression ~8 years later.

Methods

Participants

The sample has been detailed in several prior publications (2, 7-9). Briefly, the study began in 1982 with the simultaneous recruitment of two groups of probands (Generation 1, G1). Depressed probands were selected from outpatient psychiatric clinics for the treatment of mood disorders in the New Haven, CT area and were required to have moderate-to-severe MDD. Non-depressed probands were selected concurrently from the same community, and were required to have no lifetime history of psychiatric illness, based on several interviews. All probands were of European ancestry. Their biological children (G2), and subsequently, grandchildren (G3) were followed prospectively over time. The offspring of the depressed probands formed the "high-risk" group, and those of the non-depressed probands, the "low risk" group (8, 9).

Assessments

Diagnostic Interviews were conducted using the adult (10) or child (6-17 years) (11) version of the semi-structured Schedule for Affective Disorders and Schizophrenia–Lifetime interview by doctoral- and masters-level mental health professionals (reliability was high, as documented elsewhere (8, 9)). The first interview assessed the lifespan to that point; follow-up interviews assessed catch-up periods; diagnoses are therefore cumulative until latest interview. Each family member was interviewed independently and blind to the clinical status of other family members. Final diagnoses were made by one or more experienced clinicians, using the best-estimate procedure (12). Current depressive symptoms at the time of each scan were measured using the Hamilton Depression Rating Scale (13) or Children's Depression Rating Scale-Revised (14) (for adults and children respectively); anxiety symptoms were measured with the Hamilton Anxiety Rating Scale (15) and revised Children's Manifest Anxiety Scale (16), respectively. Child and adult scores were first each converted to *z*-scores, which were then combined to create a single measure. In addition to state measures, we also collected trait measures of impulsivity (Barratt Impulsiveness Scale-11 (17)), and anger and hostility (Buss Perry Aggression Questionnaire (18, 19)) These measures were assessed because our primary hypothesis

suggests that cortical thinning represents a stable biomarker of risk for depression, and thus its clinical correlates should likewise be stable, trait markers of depression risk.

Analytic Sample

We obtained MRI scans from 158 G2 and G3 offspring, aged 7-55 years at W5; and from 114 offspring, aged 11-68 at W6. MRI scans from 8 individuals at W5 and 1 individual at W6 were excluded due to severe head motion. Out of the 150 usable W5 scans and 113 usable W6 scans, there were 82 common individuals (43 from high-risk, and 39 from low-risk families). The interval between W5 and W6 scans for the same individual varied from 5.2 to 10.9 years, with a mean of 7.8 and SD of 1.3 years. Inter-scan intervals were similar for the 2nd (7.6±1.4 yrs) and 3rd (8.1±1.2 yrs) generations.

Analyses on clinical correlates of cortical thinning were based on the 110 (of 150) individuals with a usable MRI scan at W5 who also had clinical measures at W6. These 110 participants did not differ on measures of age, sex, or risk status than the primary sample of N=82 with usable brain scans at both waves (not shown).

MRI Scanning

W5 MRI scans were obtained using a Siemens Sonata 1.5 Tesla scanner using a 3D MP-RAGE sequence (TR:24msec, TE:2.96msec, Flip angle:45°, FOV:30×30cm, phase FOV:100%, 2 excitations, Slice thickness:2mm, Matrix 256×192, 128 slices, voxel dimension:1.17×1.17×1.2mm). W6 MRI scans were obtained using a GE Signa 3 Tesla whole-body scanner equipped with an 8-channel, phased array head coil using 3D Fast Spoiled Gradient Recall (FSPGR) sequence (TR:4.7msec, TE:1.3msec, Flip angle:110°, Bandwidth:41.67MHz, FOV:25×25cm, ASSET factor:2, Slice thickness:1.0mm, Matrix: 256×256, 128 slices, voxel dimension:0.98×0.98×1.0mm, 1 NEX images×2).

Processing Pipeline

Methods for estimating cortical thickness can be broadly categorized as surface-based or voxel-based. We used FreeSurfer (4, 5), a surface-based approach to estimate cortical thickness at each point on the pial surface. The use of explicit surface models enables sub-voxel accuracy, high sensitivity, and robustness to different field strengths, scanner upgrade and scanner manufacturer (20).

The processing pipeline was as follows. First, large-scale variations in image intensity are corrected using bias field estimation and bias removal. Second, extra cerebral tissues are removed using an automated tool. Third, an initial segmentation separates grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF). The WM/GM boundary is referred to as the WM boundary, and the GM/CSF boundary as the pial boundary. With this initial segmentation, FreeSurfer constructs a surface of triangulated mesh based on the WM boundary, which is then deformed to find the opposing pial boundary, correcting topological defects using smoothness and self-intersection constraints. With the cortex closed at the brain stem, the resulting surface was topologically equivalent to a sphere. The closest distance from the white surface to the pial surface at each surface's vertex is then defined as the cortical thickness.

Processed data were visually inspected and manually edited by two independent technicians (one per wave), who were well-trained in neuroanatomy to correct for errors from the automated pipeline, such as skull strip errors, segmentation errors, intensity normalization failure, etc. Data processing was blind to the subject's risk status. Thus, inter-rater differences are unlikely to generate the observed group differences (thinning in high risk offspring) at either wave, or the corresponding of group differences across waves.

Statistical Analyses

We examined the relationship between cortical thickness and risk status (high or low) using a general linear model with each vertex across the pial boundary as the dependent variable, risk status as the independent

variable, and age at current scan and sex as covariates. P-values were color coded and plotted for each vertex at the pial boundary and results controlled for the false discovery rate (21).

We explored clinical correlates of cortical thinning using path analysis. Specifically, we tested whether cortical thinning mediates the relationship between familial risk and behavioral traits associated with depression. Following established methods for path analysis(22), we conducted a series of linear regression analyses. We first tested whether the mediating variable, cortical thinning, was associated with risk status. We then tested whether cortical thinning correlated with depressive symptoms and/or traits, while controlling for risk status. Statistical significance was then determined based a 95% confidence interval (CI) derived from the Sobel test, and confirmed by bias-corrected bootstrapping with 2,000 samples (not shown). Age at first scan, age difference between the two scans, and sex, were included as covariates in all of the path linear regression models and in the Sobel test and bootstrapping.

Results

Participants at high compared to low familial risk for depression were approximately 8 years older at each scan (range 5-11, **Table 1**). Within each generation of offspring, age did not vary by risk status; however, a greater proportion of 2nd generation offspring, as compared to third, were from high risk families. Individuals from high-risk families had higher lifetime rates of depressive, but not anxiety disorders.

---------- **Table 1 here** ----------

Stable cortical thinning in offspring at high risk for depression across two assessments

Statistically significant thinning was found in the high compared to low risk group in bilateral superior frontal and caudal middle frontal cortices, and left inferior parietal lobule **(Figure 1a,** thinning in blue). Thinning patterns were consistent across the first (Wave 5[W5]; left) and second (W6, right) scans. As shown in **Table 2**, not only were the same cortical regions identified at W5 and W6, but the rank order of findings, that is, the relative effect sizes of the group thickness differences (betas), were preserved.

---------- **Figure 1 and Table 2 here** ----------

Within regions in which significant group effect were detected, we plotted within-participant correlations of cortical thickness from W5 to W6. As shown in **Figure 1b,** high intra-subject correlation (r=0.60–0.91) and intra-class correlation (absolute agreement, 0.72–0.78) from W5 to W6 further supports the stability of cortical thinning in high risk individuals across waves.

Cortical thinning mediates the relationships between familial risk and anger and hostility

We explored associations between cortical thinning at W5 and depression-related traits measured at W6 (**Figure 2**). Familial risk status predicted both anger (left, beta=4.87, p=0.017) and hostility (right, beta=3.50, p=0.021) but not impulsivity or other facets of aggression (physical or verbal aggression). Within each

hemisphere of the superior frontal gyrus, thinner cortices predicted greater anger and hostility, and the thinning mediated the relationship between familial risk and trait measures (mediating paths shown in red). Cortical thinning was not associated with current depressive symptom severity, lifetime history of depression, or impulsivity.

---------- **Figure 2 here** ----------

Supplementary Analyses

Because there were a number of co-occurring changes from W5 to W6 (e.g., MRI scanner and field strength, development stage and age, diagnostic status, substance use, or medication), we conducted sensitivity analyses to test whether any of these changes could confound the main results. The findings, shown in **Figures S1-S5**, indicate that the group differences in cortical thickness remained stable and could not be attributed to any of the above factors.

Discussion

Summary

We show that familial risk for depression is related to cortical thinning in bilateral superior and caudal middle frontal gyri, and the left inferior parietal lobule. The thinning is stable across approximately 8 years, consistent with the hypothesis that cortical thinning reflects a stable trait. The stability was robust to methodological variation including changes in MRI scanner (Siemens and GE) or field strength (1.5 and 3T), and to individual variation across participants (age and clinical trajectories). Finally, thinning in bilateral frontal cortices mediated the relationship between familial risk and trait measures of anger and hostility, suggesting a potential pathway by which familial risk for depression may contribute to the biological (cortical thinning) and clinical (anger/hostility) presentation of offspring. This is to our knowledge the largest and longest study of its kind, demonstrating the stability of a biomarker of risk for depression.

Stability across 8 years

The implications of the consistency of findings across 8 years are multi-fold. From a clinical perspective, the stability, coupled with our prior (3) and current (E-Table 4) observations that the thinning is present even in offspring who are at risk but never develop depression, further validates cortical thinning as a *risk* marker. At the same time, the high reliability of findings across different scanning platforms and field strengths suggest that detection is resistant to methodological variation. The consistency of findings across 1.5T and 3T scanning platforms is particularly noteworthy, given the lower resolution with the former (voxel size and in-plane area were 71% and 43% larger, respectively). Many early MRI studies on depression were performed on 1.5T scanners (23-26). In recent years, 3T scanners are the norm, and because the technology for neuroimaging is constantly improving, methods and field-strengths will continue to evolve. Showing stability across different platforms will be particularly important for meta-analyses and systematic reviews, where consistency of data quality across space and time is essential, yet often assumed rather than empirically investigated.

Establishing reliability has been problematic in psychiatric research. This, in part, may be due to the substantial diagnostic heterogeneity within psychiatric nosology (27), and in MRI studies is exacerbated by the risks for false positive findings which in some cases may exceed 50% (28, 29). These hurdles against identifying reliable biomarkers underscore the importance of *within*-study, test-retest reliability. Our use of open-source software will permit external groups to replicate and extend findings. Although caution is always necessary when attempting to integrate data acquired through different platforms, our findings suggest that data acquired on 1.5T scanners may be at least qualitatively comparable with newer data (see limitations, however).

The thinning observed here was only partially consistent with the thinning we previously reported using different voxel-based methods (3). Although we found that offspring at high familial risk had significant thinning of the cortical mantle in both reports, the thinning that we previously reported was more globally distributed through parietal, temporal and frontal cortices, and predominantly in the lateral right and medial left hemisphere, an asymmetry we do not find here. Differences could be due to the only partially overlapping samples, or due to different approaches implemented. Surface-based estimations (as implemented here with FreeSurfer) may be preferable for longitudinal analyses because of deformable model processing, which optimizes correction against segmentation differences across scans (6). Conversely, voxel-based methods, which measure thickness directly from initial segmentation, are more prone to segmentation errors. Our use of open-source software will now allow other investigators to compare patterns of cortical thinning in independent samples using parallel methods.

Cortical thinning in depressive disorders

A number of prior studies have reported reductions in cortical thickness in frontal and temporal lobes in depressive disorders (30-33) [reviews (27, 34)]. The thinning has been further associated with atypical functional connectivity(33) and executive functioning(32, 35), and with depression phenotypes in other disorders (e.g., Alzheimer's dementia(36), diabetes(35)). Most studies to date however have only compared

individuals with and without depression, precluding disentanglement of brain differences that predispose to the disorder from those that may arise as a consequence of the illness. By identifying subjects based on *risk for*, rather than, *presence of* the outcome of interest, high-risk designs allow us to target biological processes that emerge prior to onset of disease (37). Few imaging studies have applied this approach to depression. A study by Papmeyer *et al* followed early adult offspring (mean age, 21 yrs) at high- and low-familial risk for depression, but only across a two-year period (31). Offspring at high-risk for depression had thinner right temporal cortex, regardless of whether they developed the disorder themselves. Whereas high-risk offspring who did not develop depression showed thinning across the 2 years in inferior frontal cortex, those who did develop depression showed corresponding thickening, which the authors suggest may be a function of insufficient synaptic pruning.

Cortical thinning mediates associations between familial risk and depression phenotypes

Frontal cortical thinning mediated the association between familial risk for depression and two adulthood traits associated with depression, hostility and anger. As illustrated in Fig.2, frontal thinning partially explained the relationship between familial risk and anger and hostility. Interestingly, thinning did not explain associations between familial risk and offspring depression. This could be because depressive symptoms present episodically and thus may be more likely to associate with similarly episodic biological measures. Expression also depends on the environment, and many offspring who may be at high-familial risk and have the cortical vulnerability still never go on to develop symptoms.

Though not core clinical symptoms of depression, anger and hostility are reported in about a third of depressed patients (15, 38), and are associated with worse course of the disorder, including poorer response to pharmacological treatment, and higher rates of suicide attempts, accidents, and cardiovascular disease (39, 40). Family history may thus lead to *some* depressions by increasing proneness to aggression traits via dysregulation of circuits involving regions showing frontal thinning. This hypothesis is consistent with studies showing higher levels of aggression in adolescents with thinner middle frontal cortices or disruptions in

fronto-limbic circuitry (41). Future research can test this hypothesis more directly by following teenage offspring who have both frontal thinning and elevated aggression traits, and test whether the aggression mediates an association between the thinning and later onset of clinical symptoms.

Limitations

The sample is based on families at high- or low-risk for major depression, neither of which are completely reflective of the general U.S. population. The sample was also of European ancestry, so findings may not generalize to other groups. Although we show that cortical thinning predicts measures of aggression ~8 years later, temporal causality should not be inferred, as the onset for development of the cortical thinning versus behavioral traits is not known. The effect size was also moderate-to-large: for subtler regional differences or more sensitive imaging modalities (e.g., fMRI), the signal-to-noise ratio may be insufficient to document correspondence across scans. On the other hand, the method variation across scans could have inflated Type II errors, leaving us with a conservative estimate of stability. Associations between cortical thickness and impulsivity could be accounted for by third variables (e.g., trauma) that could alter both thickness and trait measures. Finally, the study focused on biomarkers in offspring at high *familial* risk for depression. Thus, it is unclear whether similar markers would be found for other types of depression that are not family based. Similarly, as the study did not include offspring of probands with other disorders, we cannot empirically test whether the markers are specific to depression versus reflective of broader psychiatric risk.

Conclusions

Identifying biomarkers of vulnerability is a critical step in treatment development. A true vulnerability marker must represent *risk* for illness: simply serving as a one-time correlate of illness is insufficient (1). Using a high-risk family design, we have identified a stable biomarker for depression risk (bilateral frontal thinning) that meets the above criteria. The marker first needs to be reproduced and validated independently. Subsequent research can then address predictive validity by following individuals with frontal thinning over time and

examining the long-term clinical trajectories. At the same time, functional imaging approaches can be applied to examine brain-wide circuitry implications of the thinning.

Acknowledgements:

This work was supported in part by the National Institute of Mental Health (R01MH036197, Weissman, P.I., 1P50MH090966, Gingrich, P.I), the Sackler Institute for Developmental Psychobiology, and the Templeton Foundation. The funding sources had no role in the design of the study; collection, management, analysis, or interpretation of the data; preparation, review, or approval of the manuscript; or the decision to submit the manuscript for publication.

Disclosures:

In the last three years, Dr Weissman has reported receiving royalties from Oxford University Press, Perseus Books Group, American Psychiatric Association Publishing, and Multi-Health Systems. Dr Posner has received funding from Shire Pharmaceuticals. None of these present any conflict with the present work. All other authors report no biomedical financial interests or potential conflicts of interest.

References

1. Gottesman, II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. Am J Psychiatry 2003; 160(4): 636-45.
2. Gould TD, Gottesman, II. Psychiatric endophenotypes and the development of valid animal models. Genes, brain, and behavior 2006; 5(2): 113-9.
3. Peterson BS, Warner V, Bansal R, Zhu H, Hao X, Liu J, *et al.* Cortical thinning in persons at increased familial risk for major depression. Proceedings of the National Academy of Sciences of the United States of America 2009; 106(15): 6273-8.
4. Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. Segmentation and surface reconstruction. NeuroImage 1999; 9(2): 179-94.
5. Fischl B, Sereno MI, Dale AM. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. NeuroImage 1999; 9(2): 195-207.
6. Fisher LB, Fava M, Doros GD, Alpert JE, Henry M, Huz I, *et al.* The Role of Anger/Hostility in Treatment-Resistant Depression: A Secondary Analysis From the ADAPT-A Study. The Journal of nervous and mental disease 2015; 203(10): 762-8.
7. Mills KL, Tamnes CK. Methods and considerations for longitudinal structural brain imaging analysis across development. Developmental cognitive neuroscience 2014; 9: 172-90.
8. Weissman MM, Wickramaratne P, Nomura Y, Warner V, Pilowsky D, Verdeli H. Offspring of depressed parents: 20 years later. Am J Psychiatry 2006; 163(6): 1001-8.
9. Weissman MM, Wickramaratne P, Nomura Y, Warner V, Verdeli H, Pilowsky DJ, *et al.* Families at high and low risk for depression: a 3-generation study. Arch Gen Psychiatry 2005; 62(1): 29-36.
10. Mannuzza S, Fyer AJ, Klein DF, Endicott J. Schedule for Affective Disorders and Schizophrenia--Lifetime Version modified for the study of anxiety disorders (SADS-LA): rationale and conceptual development. Journal of psychiatric research 1986; 20(4): 317-25.
11. Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, *et al.* Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. Journal of the American Academy of Child and Adolescent Psychiatry 1997; 36(7): 980-8.
12. Fried PA, Watkinson B, Dillon RF, Dulberg CS. Neonatal neurological status in a low-risk population after prenatal exposure to cigarettes, marijuana, and alcohol. Journal of developmental and behavioral pediatrics : JDBP 1987; 8(6): 318-26.
13. Hamilton M. Development of a rating scale for primary depressive illness. The British journal of social and clinical psychology 1967; 6(4): 278-96.
14. Poznanski E, Mokros HB, Grossman J, Freeman LN. Diagnostic criteria in childhood depression. Am J Psychiatry 1985; 142(10): 1168-73.
15. Fava M, Rosenbaum JF. Anger attacks in patients with depression. The Journal of clinical psychiatry 1999; 60 Suppl 15: 21-4.
16. Perrin S, Last CG. Do childhood anxiety measures measure anxiety? Journal of abnormal child psychology 1992; 20(6): 567-78.
17. Patton JH, Stanford MS, Barratt ES. Factor structure of the Barratt impulsiveness scale. Journal of clinical psychology 1995; 51(6): 768-74.
18. Buss AH, Perry M. The aggression questionnaire. J Pers Soc Psychol 1992; 63(3): 452-9.

19. Gerevich J, Bacskai E, Czobor P. The generalizability of the Buss-Perry Aggression Questionnaire. Int J Methods Psychiatr Res 2007; 16(3): 124-36.

20. Han X, Jovicich J, Salat D, van der Kouwe A, Quinn B, Czanner S, *et al.* Reliability of MRI-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer. NeuroImage 2006; 32(1): 180-94.

21. Banjamini Y YD. The control of the false discovery rate in multiple testing under dependency. Annals of Statistics 2001; 29: 1165-88.

22. Rucker DD, Preacher KJ, Tormala ZL, Petty RE. Mediation analysis in social psychology: Current practices and new recommendations. Social and Personality Psychology Compass 2011; 5(6): 359-71.

23. Koponen S, Taiminen T, Kurki T, Portin R, Isoniemi H, Himanen L, *et al.* MRI findings and Axis I and II psychiatric disorders after traumatic brain injury: a 30-year retrospective follow-up study. Psychiatry research 2006; 146(3): 263-70.

24. Kumar A, Bilker W, Jin Z, Udupa J. Atrophy and high intensity lesions: complementary neurobiological mechanisms in late-life major depression. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology 2000; 22(3): 264-74.

25. Rosenberg DR, Macmaster FP, Mirza Y, Smith JM, Easter PC, Banerjee SP, *et al.* Reduced anterior cingulate glutamate in pediatric major depression: a magnetic resonance spectroscopy study. Biological psychiatry 2005; 58(9): 700-4.

26. Rosso IM, Cintron CM, Steingard RJ, Renshaw PF, Young AD, Yurgelun-Todd DA. Amygdala and hippocampus volumes in pediatric major depression. Biological psychiatry 2005; 57(1): 21-6.

27. Bora E, Fornito A, Yucel M, Pantelis C. The effects of gender on grey matter abnormalities in major psychoses: a comparative voxelwise meta-analysis of schizophrenia and bipolar disorder. Psychological medicine 2012; 42(2): 295-307.

28. Eklund A, Nichols TE, Knutsson H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. Proceedings of the National Academy of Sciences of the United States of America 2016; 113(28): 7900-5.

29. Salmond CH, Ashburner J, Vargha-Khadem F, Connelly A, Gadian DG, Friston KJ. Distributional assumptions in voxel-based morphometry. NeuroImage 2002; 17(2): 1027-30.

30. Grieve SM, Korgaonkar MS, Koslow SH, Gordon E, Williams LM. Widespread reductions in gray matter volume in depression. NeuroImage Clinical 2013; 3: 332-9.

31. Papmeyer M, Giles S, Sussmann JE, Kielty S, Stewart T, Lawrie SM, *et al.* Cortical Thickness in Individuals at High Familial Risk of Mood Disorders as They Develop Major Depressive Disorder. Biological psychiatry 2015; 78(1): 58-66.

32. Tu PC, Chen LF, Hsieh JC, Bai YM, Li CT, Su TP. Regional cortical thinning in patients with major depressive disorder: a surface-based morphometry study. Psychiatry research 2012; 202(3): 206-13.

33. van Tol MJ, Li M, Metzger CD, Hailla N, Horn DI, Li W, *et al.* Local cortical thinning links to resting-state disconnectivity in major depressive disorder. Psychological medicine 2014; 44(10): 2053-65.

34. Kempton MJ, Salvador Z, Munafo MR, Geddes JR, Simmons A, Frangou S, *et al.* Structural neuroimaging studies in major depressive disorder. Meta-analysis and comparison with bipolar disorder. Arch Gen Psychiatry 2011; 68(7): 675-90.

35. Ajilore O, Narr K, Rosenthal J, Pham D, Hamilton L, Watari K *et al.* Regional cortical gray matter thickness differences associated with type 2 diabetes and major depression. Psychiatry research 2010; 184(2): 63-70.

36. Lebedev AV, Beyer MK, Fritze F, Westman E, Ballard C, Aarsland D. Cortical changes associated with depression and antidepressant use in Alzheimer and Lewy body dementia: an MRI surface-based morphometric study. The American journal of geriatric psychiatry : official journal of the American Association for Geriatric Psychiatry 2014; 22(1): 4-13 e1.

37. Weissman MM, Brown AS, Talati A. Translational epidemiology in psychiatry: linking population to clinical and basic sciences. Arch Gen Psychiatry 2011; 68(6): 600-8.

38. Fava M, Rosenbaum JF, McCarthy M, Pava J, Steingard R, Bless E. Anger attacks in depressed outpatients and their response to fluoxetine. Psychopharmacology bulletin 1991; 27(3): 275-9.

39. Papmeyer M, Sussmann JE, Stewart T, Giles S, Centola JG, Zannias V, *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 research 2016; 248: 119-25.

40. Fava M, Anderson K, Rosenbaum JF. "Anger attacks": possible variants of panic and major depressive disorders. Am J Psychiatry 1990; 147(7): 867-70.

41. Yang Y, Joshi SH, Jahanshad N, Thompson PM, Baker LA. Neural correlates of proactive and reactive aggression in adolescent twins. Aggressive behavior 2016.

Table 1. Demographic and clinical characteristics of 82 participants with scans of W5 and W6

		High risk (n = 43)	Low risk (n = 39)	Test statistic	P-value
Age, mean, years	*W5*	31.7 ± 12.8	23.1 ± 12.8	t = 3.1	0.003
	W6	39.4 ± 12.7	31.2 ± 12.8	t = 2.9	0.005
Age, by generation, years					
Second generation (G2)	*W5*	38.7 ± 8.1	39.7 ± 5.6	t = 0.5	0.63
	W6	46.2 ± 8.1	47.8 ± 5.9	t = 0.7	0.48
Third generation (G3)	*W5*	15.8 ± 5.0	14.7 ± 3.9	t = 0.7	0.50
	W6	23.8 ± 4.9	23.0 ± 4.3	t = 0.5	0.62
Generation					
Second generation (G2)		30 (69.8%)	13 (33.3%)	$X^2 = 10.9$	0.001
Third generation (G3)		13 (30.2%)	26 (66.7%)		
Gender					
Male		20 (46.5%)	17 (43.6%)	$X^2 = 0.07$	0.79
Female		23 (53.5%)	22 (56.4%)		
Depressive symptoms[1]					
Adult	*W5*	4 ± 4.9	3.2 ± 4.2	t = 0.6	0.55
	W6	3.6 ± 5.1	1.4 ± 4.5	t = 2.0	0.05
Child	*W5*	20.2 ± 9.0	21.0 ± 7.1	t = 0.2	0.82
	W6	21.3 ± 4.5	18.0 ± 16.7	t = 0.3	0.77
Anxiety symptoms[2]					
Adult	*W5*	3.7 ± 4.1	2.5 ± 3.5	t = 1.1	0.28
	W6	2.6 ± 3.4	1.0 ± 3.0	t = 2.2	0.04
Child	*W5*	8.7 ± 8.3	5.8 ± 6.7	t = 0.9	0.37
	W6	2.3 ± 3.2	3.0 ± 5.2	t = 0.3	0.86
Current/lifetime depressive disorder		24 (55.8%)	12 (30.8%)	$X^2 = 5.2$	0.02
Current/lifetime anxiety disorder		28 (65.1%)	19 (48.7%)	$X^2 = 2.2$	0.13

Values are mean ± SD.

[1]Depressive symptoms were determined by Hamilton Depression Rating Scale and the Children's Depression Inventory for adults and children, respectively.

[2]Anxiety symptoms were determined by the Hamilton Anxiety Rating Scale and the Revised Children's Manifest Anxiety Scale for adults and children, respectively.

Symptoms scores were z-transformed for adults and children independently before they were included as covariates in the statistical models.

Table 2. Cortical ROIs with significant thinning at either wave of assessment

Cortical Region	**W5**		**W6**	
	Size (mm^2)	Beta mean (mm)	Size (mm^2)	Beta mean (mm)
Right superior frontal	2983.5	-0.158	2985.1	-0.112
Left superior frontal	3835.0	-0.143	4023.9	-0.105
Left caudal middle frontal	1200.4	-0.141	1020.2	-0.102
Right caudal middle frontal	1411.2	-0.137	1291.0	-0.098
Left inferior parietal	850.4	-0.093	772.2	-0.096

The size of thinning regions and their corresponding average beta values are reported here for statistically significant ones.

Figure Titles and Legends

Figure 1. Maps of group differences in cortical thickness of 82 common subjects (43 high-risk, 39 low-risk) at W5 and W6.

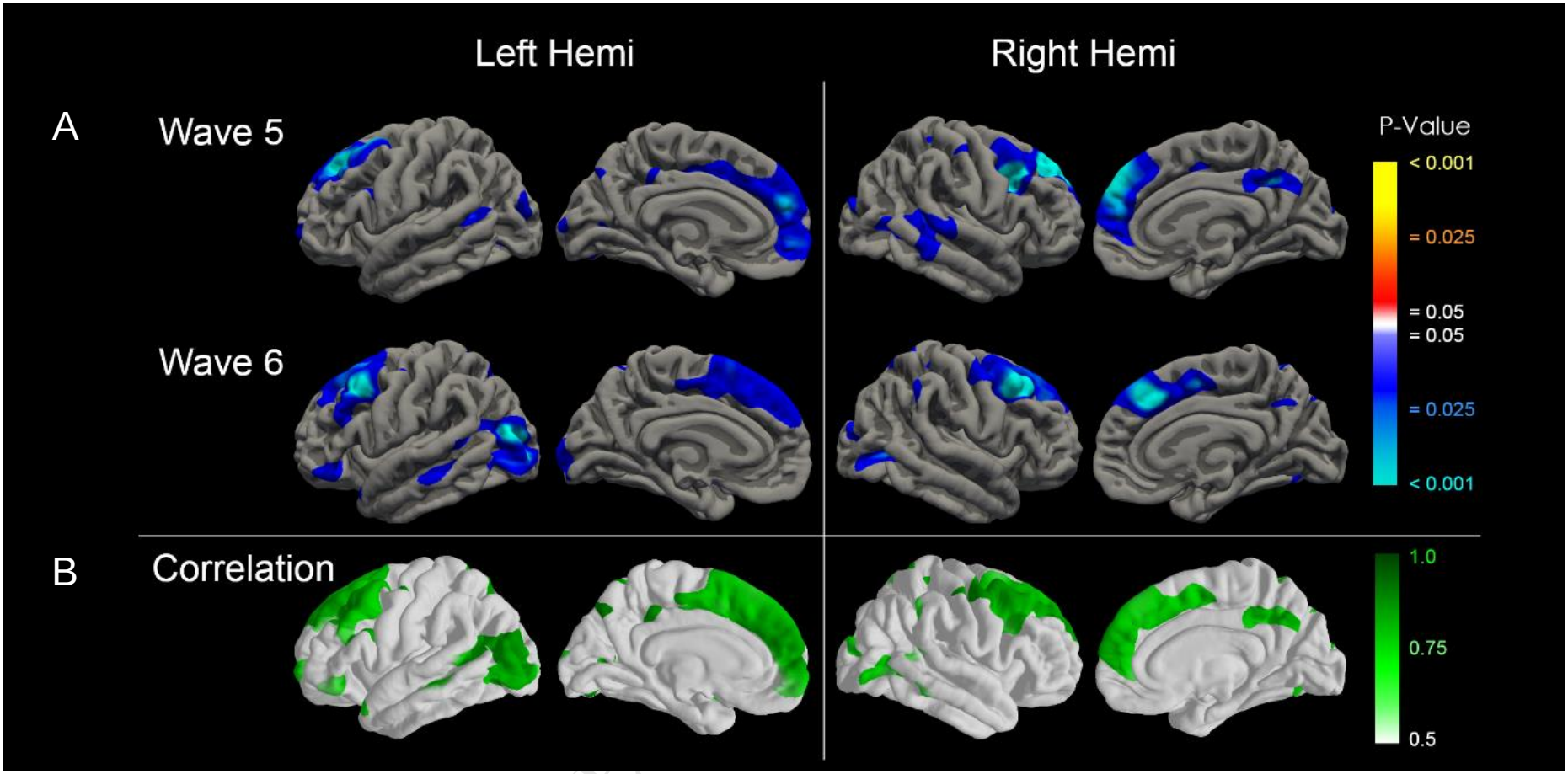
Panel A: At each point on the cerebral surface, the statistical significance (probability values) of differences in cortical thickness across groups (high vs. low risk) in participants from generations G2 and G3 are color coded. Warm colors (yellow, orange, and red) represent significantly thicker cortices in the high-risk group; cooler colors (blue and cyan) represent thinner cortices in that group. The color bar indicates the color-coding of p-values for testing of statistical significance at each point on the surface of the brain. The statistical models controlled for the age, and the sex of all participants.

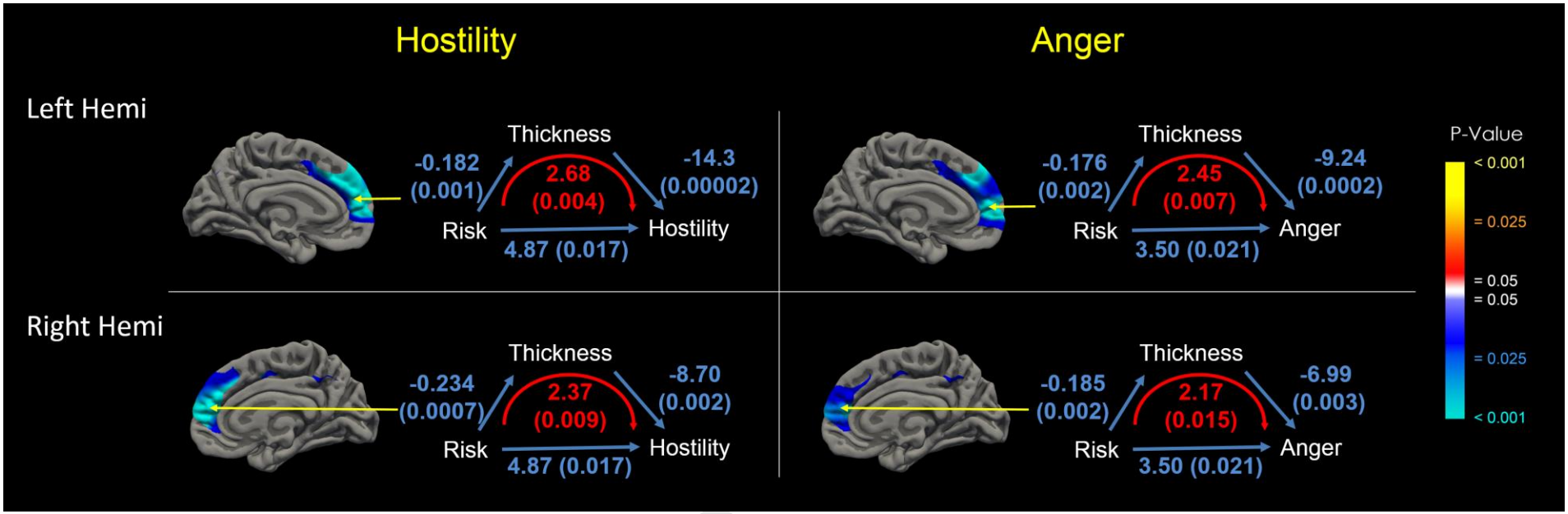
Panel B: Plot of Pearson correlation coefficients of cortical thickness between W5 and W6 in conjunct thinning regions.

Figure 2. Correlation and mediation of cortical thickness with measures of aggression hostility

Brain images show correlation of cortical thickness at the first scan (i.e., w5) with measures of hostility (left panel) and anger (right). Regions shown in warm colors (yellow to red) represent those where thicker cortex was associated with higher hostility/anger scores (there were no regions meeting these criteria). Regions shown in cool colors (blue and cyan) represent those where thinner cortex was associated with higher hostility/anger. The color bar on the right indicates the color-coding of p-values for testing of statistical significance at each point on the surface of the brain.

The arrow diagrams illustrate the pathways tested in the mediation models, with the blue pathways indicating the three direct paths, and the red pathway, the indirect path. Each path lists the beta values (adjusted for age at first scan, age difference between scans, and sex), with corresponding p-values in parentheses. The red path indicates the indirect path, that is, the portion of the association between familial risk and anger/hostility outcomes that is mediated (explained) by the cortical thinning in the respective region.


Left Hemi
Right Hemi
A
Wave 5
Wave 6
P-Value
< 0.001
= 0.025
= 0.05
= 0.05
= 0.025
< 0.001
B
Correlation
1.0
0.75
0.5


Hostility
Anger
Left Hemi
Right Hemi
Thickness
-0.182 (0.001)
2.68 (0.004)
-14.3 (0.00002)
Risk
Hostility
4.87 (0.017)
-0.176 (0.002)
2.45 (0.007)
-9.24 (0.0002)
Anger
3.50 (0.021)
-0.234 (0.0007)
2.37 (0.009)
-8.70 (0.002)
-0.185 (0.002)
2.17 (0.015)
-6.99 (0.003)
P-Value
< 0.001
= 0.025
= 0.05
= 0.05
= 0.025
< 0.001

Stability of Cortical Thinning in Persons at Increased Familial Risk for Major Depression Across Eight Years

Supplemental Information

Primary analyses presented in Table 1 and Figure 1 document stability of cortical thinning among families at high (compared to low) familial risk for depression across two time points approximately 8 years apart. However, there were a number of changes that co-occurred in the same eight year window, raising the possibility that these changes could confound group differences. These include changes in scanner, differential changes in developmental phase (puberty) and overall age across participants, and clinical changes.

We thus performed a series of sensitivity analyses to address the potential impact these factors could have had on the main finding.

Scanner change

We used a 1.5 Tesla Siemens scanner at W5 and a 3 Tesla GE scanner at W6. Differing field strength of the scanner led to different contrast and image quality at the two waves (**Figure S1**) which could result in differences in cortical thickness measurements across waves. Moreover, the voxel size of a W5 scan was about 70% bigger than that of a W6 scan. We compared the thickness changes between W5 and W6 across the pial surface for high- and low-risk groups. Though different contrast resulted from the two different scanners, together with different spatial resolution at W5 and W6 led to significantly different cortical thickness measurements between W5 and W6. Measurements at W6, in general, generated thicker cortex than W5, except at the frontal portion of the left hemisphere where cortical thickness at W6 was thinner. The differences, though, were consistent across high- and low-risk groups and thus we did not find any group difference in cortical thickness changes between the two waves. This implies that even though there were significant differences in cortical thickness measurements between W5 and W6, the effects were similar across high- and low-risk groups and would most likely cancel out when we

compared cortical thickness between risk groups within W5 or W6. Thus the thinning effects in high-risk group were stable and consistent across W5 and W6, and were not affected much by the scanner change.

Developmental phase

Differences in the thinning effects at W5 and W6 could be caused by subjects who went through puberty between the two different waves. To assess the effects of those subjects, we removed 9 subjects who went from under 12 years old at W5 to over 12 at W6 and reran the analyses at both waves (**Figure S2**) (73 subjects, 40 high-risk, 33 low-risk). Furthermore, we removed 13 subjects who were less than 13 at W5 or less than 18 at W6 (69 subjects, 39 high-risk, 30 low-risk, not shown). The results from both analyses were similar or identical to the results that included those who went through puberty between W5 and W6 (Fig. 1).

Age differences

As there were overall age differences by group (Table 1), we investigated its effect on the cortical thinning results by plotting the *age x risk status* interaction term on the surface of the brain. There was no significance of the interaction term across both hemispheres in either W5 or W6, suggesting that differences in age in high versus low- risk groups were unlikely to bias the observed cortical thinning at W5 or W6. Furthermore, we compared age-matched 66 subjects of high- and low-risk subgroups and found the similar thinning pattern (**Figure S3**) as shown in the main figure.

Clinical status change

There were no new onsets of MDD in the 82 common subjects between the two scan waves, though there were 8 recurrences. To test if those clinical status changes confounded the detected cortical thinning, we excluded those offspring from the analyses and generated the cortical thinning results in the 74 remaining subjects (37 high-risk, 37 low-risk, **Figure S4**). The effects were largely similar to the ones that included those 8 recurrences (Fig. 1). Statistically adjusting for offspring's prior history of depression and anxiety

(**Figure S5**), and current depressive and anxiety symptoms (not shown) also did not significantly change the findings. We compared cortical thickness of subjects with lifetime history of MDD and subjects without MDD and found no significant thinning in the regions where thinning in the high-risk group was considered an MDD trait, i.e., left superior frontal cortex, left caudal middle frontal, left inferior parietal lobule, right superior frontal cortex.

Finally, although no offspring developed a new diagnosis of depression, six participants developed a diagnosis of an anxiety disorder, and one of a substance use disorder, between the two scans. The findings remained unchanged when we conducted sensitivity analyses removing these individuals from the analyses.

Scanner Change

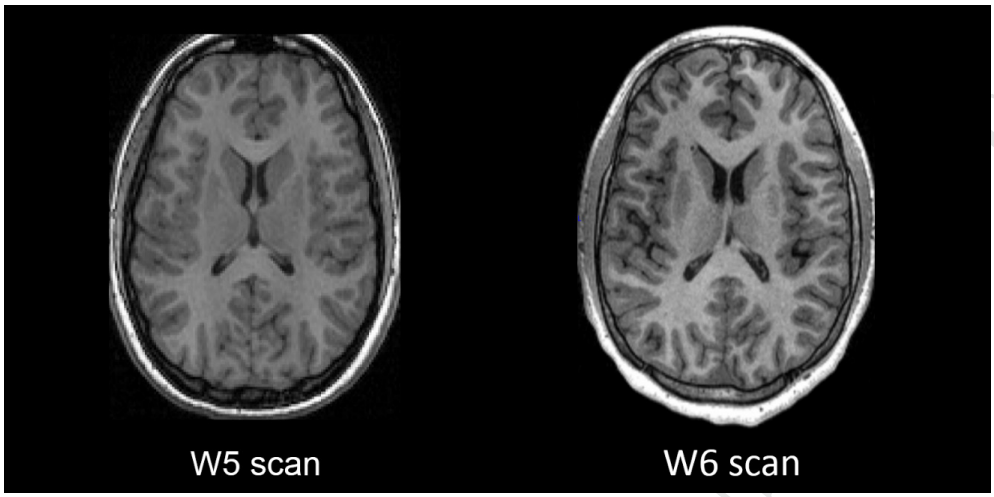


Figure S1: Scan quality at W5 and W6. W6 data had better scan quality due to the higher field strength of the scanner.

Adjusting for the effects of development

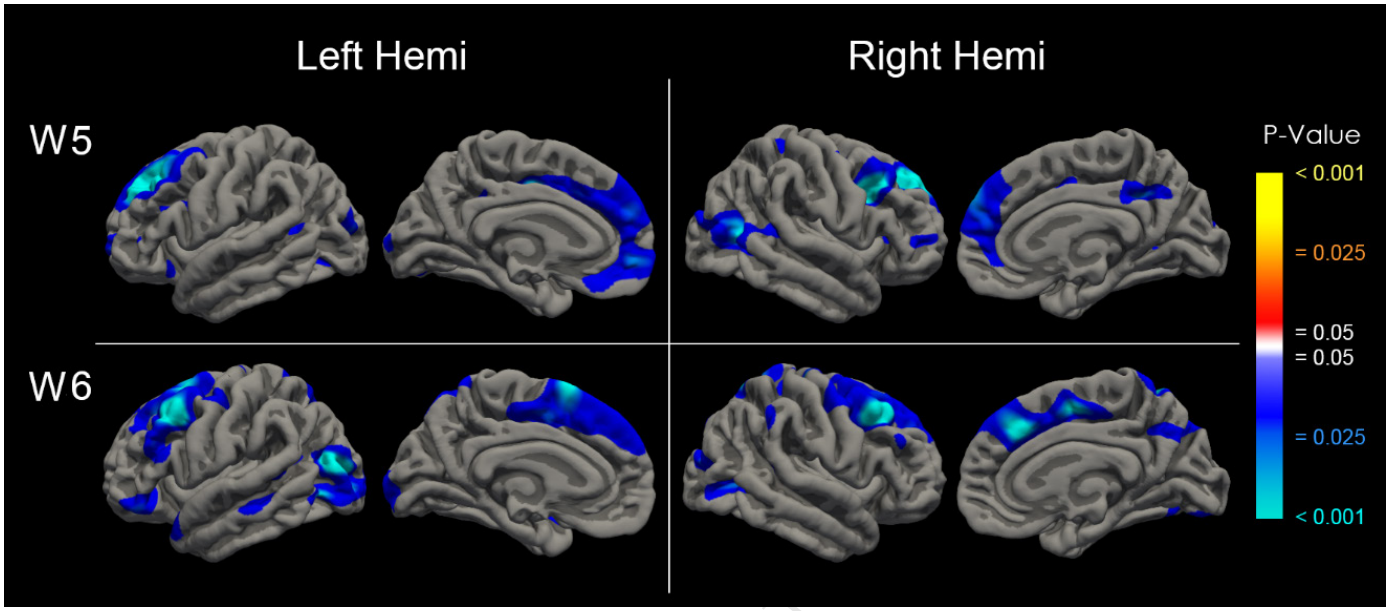


Figure S2: Maps of group differences in cortical thickness of 73 common subjects (40 high-risk, 33 low-risk, excluding 9 subjects who went from under 12 years at W5 to over 12 at W6) at W5 and W6. At each point on the cerebral surface, the statistical significance (probability values) of differences in cortical thickness across groups (high vs. low risk) in participants from generations G2 and G3 are color coded. Warm colors (yellow, orange, and red) represent significantly thicker cortices in the high-risk group; cooler colors (blue and cyan) represent thinner cortices in that group. The color bar indicates the color-coding of p-values for testing of statistical significance at each point on the surface of the brain. The statistical models controlled for the age, and the sex of all participants.

Adjusting for the effects of age

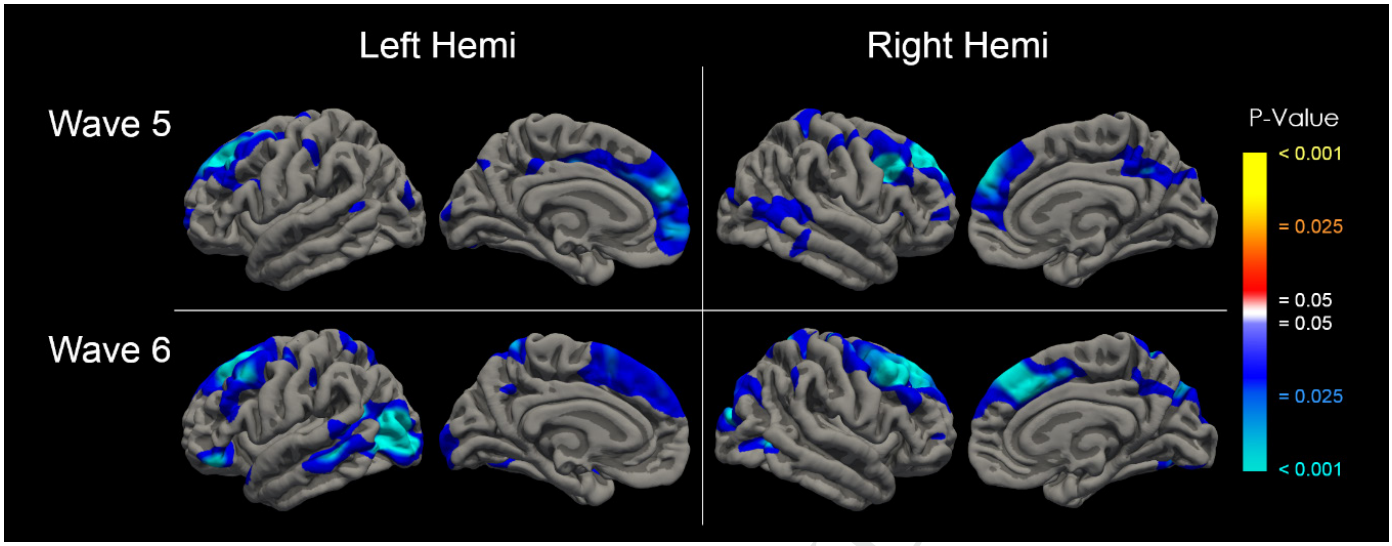


Figure S3: Maps of group differences in cortical thickness of 66 common subjects (33 high-risk, 33 low-risk, age matched between risk groups) at W5 and W6. At each point on the cerebral surface, the statistical significance (probability values) of differences in cortical thickness across groups (high vs. low risk) in participants from generations G2 and G3 are color coded. Warm colors (yellow, orange, and red) represent significantly thicker cortices in the high-risk group; cooler colors (blue and cyan) represent thinner cortices in that group. The color bar indicates the color-coding of p-values for testing of statistical significance at each point on the surface of the brain. The statistical models controlled for the age, and the sex of all participants.

Adjusting for the effects of clinical depression

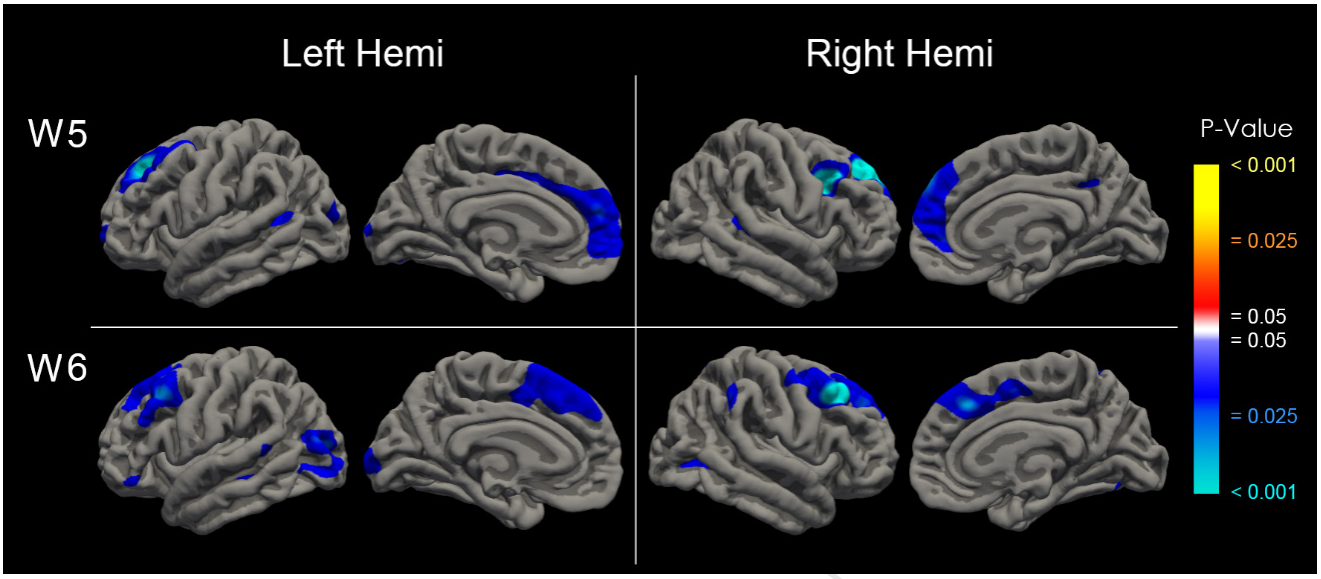


Figure S4. Maps of group differences in cortical thickness of 74 common subjects (37 high-risk, 37 low-risk, excluded 8 subjects who had recurrence of MDD between W5 and W6) at W5 and W6. At each point on the cerebral surface, the statistical significance (probability values) of differences in cortical thickness across groups (high vs. low risk) in participants from generations G2 and G3 are color coded. Warm colors (yellow, orange, and red) represent significantly thicker cortices in the high-risk group; cooler colors (blue and cyan) represent thinner cortices in that group. The color bar indicates the color-coding of p-values for testing of statistical significance at each point on the surface of the brain. The statistical models controlled for the age, and the sex of all participants.

Adjusting for the effects of clinical symptoms

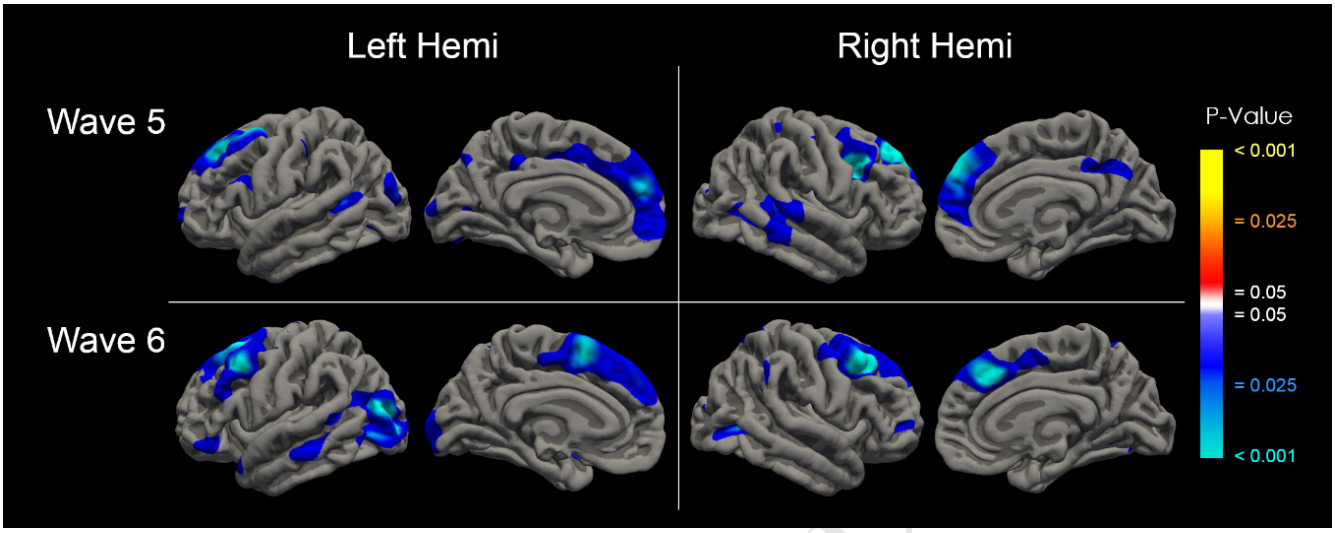


Figure S5. Maps of group differences in cortical thickness of 82 common subjects (43 high-risk, 39 low-risk) at W5 and W6, controlled for history of depression and anxiety. At each point on the cerebral surface, the statistical significance (probability values) of differences in cortical thickness across groups (high vs. low risk) in participants from generations G2 and G3 are color coded. Warm colors (yellow, orange, and red) represent significantly thicker cortices in the high-risk group; cooler colors (blue and cyan) represent thinner cortices in that group. The color bar indicates the color-coding of p-values for testing of statistical significance at each point on the surface of the brain. The statistical models also controlled for the age and the sex of all participants.