Attention-deficit/hyperactivity disorder is characterized by a delay in cortical maturation

P. Shaw†*, K. Eckstrand†, W. Sharp†, J. Blumenthal†, J. P. Lerch§, D. Greenstein†, L. Clasen†, A. Evans§, J. Giedd†, and J. L. Rapoport†

†Child Psychiatry Branch, National Institute of Mental Health, Room 3N202, Building 10, Center Drive, Bethesda, MD 20892; and §Montreal Neurological Institute, McGill University, Montreal, QC, Canada H3A 2T5

Edited by Leslie G. Ungerleider, National Institutes of Health, Bethesda, MD, and approved October 5, 2007 (received for review August 17, 2007)

There is controversy over the nature of the disturbance in brain development that underpins attention-deficit/hyperactivity disorder (ADHD). In particular, it is unclear whether the disorder results from a delay in brain maturation or whether it represents a complete deviation from the template of typical development. Using computational neuroanatomic techniques, we estimated cortical thickness at >40,000 cerebral points from 824 magnetic resonance scans acquired prospectively on 223 children with ADHD and 223 typically developing controls. With this sample size, we could define the growth trajectory of each cortical point, delineating a phase of childhood increase followed by adolescent decrease in cortical thickness (a quadratic growth model). From these trajectories, the age of attaining peak cortical thickness was derived and used as an index of cortical maturation. We found maturation to progress in a similar manner regionally in both children with and without ADHD, with primary sensory areas attaining peak cortical thickness before polymodal, high-order association areas. However, there was a marked delay in ADHD in attaining peak thickness throughout most of the cerebrum: the median age by which 50% of the cortical points attained peak thickness for this group was 10.5 years (SE 0.01), which was significantly later than the median age of 7.5 years (SE 0.02) for typically developing controls (log rank 10.5 years (SE 0.01), which was significantly later than the median of the cortical points attained peak thickness for this group was 10.5 years (SE 0.01), which was significantly later than the median age of 7.5 years (SE 0.02) for typically developing controls (log rank test $\chi^2(12) = 5,609, P < 1.0 \times 10^{-20}$). The delay was most prominent in prefrontal regions important for control of cognitive processes including attention and motor planning. Neuroanatomic documentation of a delay in regional cortical maturation in ADHD has not been previously reported.

A

Attention-deficit/hyperactivity disorder (ADHD) is the most common neurodevelopmental disorder of childhood affecting between 3% and 5% of school-aged children (1). Since its earliest descriptions, there has been debate as to whether the disorder is a consequence partly of delay in brain maturation or as a complete deviation from the template of typical development (2). Several studies find that brain activity at rest and in response to cognitive probes is similar between children with ADHD and their slightly younger but typically developing peers, evidence congruent with a maturational lag in cortical development (3–5). However, others report a qualitatively distinct neurophysiology, with a unique architecture of the electroencephalogram and some highly anomalous findings in functional imaging studies, more in keeping with ADHD as a deviation from typical development (6–10).

In a previous longitudinal study, we found parallel trajectories of gray lobar volume change in children with ADHD and typically developing controls, but more focal changes in cortical maturation occurring at a sublobar level would not be detected by this lobar measure (11). We thus aimed to define the trajectory of cortical development using a measure of cortical thickness that affords exquisite spatial resolution. Cortical thickness was chosen as a metric that both captures the columnar architecture of the cortex and is sensitive to developmental change in typically developing and clinical populations (12–15).

Most of the 446 children in the current study had repeated neuroanatomic imaging—112 (25%) had two scans, 88 (20%) had three scans, and 30 (7%) had four or more scans, performed at a mean interval between scans of 2.8 years. Such longitudinal data can be combined with cross-sectional data by using mixed-model regression to model developmental change, with the longitudinal data being particularly informative. For cortical thickness data, the simplest trajectory that can be fitted to describe its change over time is a straight line. More complex growth models include distinct phases of increase and decrease in cortical thickness: A quadratic model has two such phases (typically an initial increase that reaches a peak before declining) and a cubic model has three. Derived properties of these developmental curves are frequently used as developmental indices, such as the age at which points of inflection in the curve are attained (16, 17). When considering cortical change, the age at which peak cortical thickness is reached—the point where increase gives way to decrease in cortical thickness—emerges as a particularly useful index. Note that the ability to detect a quadratic or cubic growth model is a prerequisite for defining the age of peak cortical thickness; it cannot be determined from a linear model.

We thus compared the age of attaining peak cortical thickness in children with and without ADHD to determine whether the disorder is characterized by a delay in cerebral cortical maturation.

Results

The temporal sequence of cortical maturation, reflected by the age of reaching peak cortical thickness at cortical points where a quadratic model was appropriate, was similar in both groups [see supporting information (SI) Movies 1 and 2 and Fig. 1]. In the frontal cortex, the superior, precentral, and polar regions reached an early peak in cortical thickness, followed by a centrifugal wave moving toward the middle prefrontal cortex. In the temporal cortex, posterior portions of the middle and superior temporal cortex matured before more anterior temporal regions. In the occipital cortex, for both the typically developing and ADHD subjects, there were early peaks with little developmental change in the age period covered. Direct comparison of cortical change in the parietal regions was complicated because the groups differed in the regions where a quadratic model was appropriate.

However, although the overall pattern of development was similar, there were pronounced diagnostic differences in timing. Where a peak age could be determined, the ADHD group

Author contributions: P.S. and J.L.R. designed research; P.S., W.S., J.B., L.C., J.G., and J.L.R. performed research; K.E., J.B., J.P.L., and A.E. contributed new reagents/analytic tools; P.S., K.E., and D.G. analyzed data; and P.S. and J.L.R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

To whom correspondence should be addressed. E-mail: shawp@mail.nih.gov.

This article contains supporting information online at www.pnas.org/cgi/content/full/0707741104DC1.

© 2007 by The National Academy of Sciences of the USA
generally reached this milestone later than the typically developing controls; see Fig. 2. Kaplan–Meier curves showed that the median age by which 50% of the cortical points had attained peak thickness for the ADHD group was 10.5 years (SE 0.01), which was significantly later than the median age of 7.5 years (SE 0.02) for the typically developing controls (log-rank test $\chi^2 = 5,609, P < 1.0 \times 10^{-20}$); Fig. 3. Differences were most prominent in the middle prefrontal cortex, where the ADHD group reached their peak thickness 5 years after the typically developing controls, and to a lesser extent in the superior prefrontal and medial prefrontal cortex (with the ADHD group peaking $\approx$2 years later). Kaplan–Meier curves for the prefrontal region demonstrated that, although both groups had a similar rates of attaining cortical thickness, this was delayed in the ADHD group with a median age of 10.4 years (SE 0.02), compared with typically developing control median age of 7.5 years (SE 0.02) ($\chi^2 = 9,599, P < 1.0 \times 10^{-20}$). Posteriorly, delay was present bilaterally in the middle and superior temporal cortex, extending to the middle occipital gyri, with the ADHD group having a peak age of 10.6 years (SE 0.04) and the typically developing controls peaking at 6.8 years (SE 0.08) ($\chi^2 = 303, P < 1.0 \times 10^{-20}$).

The ADHD group had an earlier peak thickness predominately in the primary motor cortex, with a median age by which 50% of points within this region peaked at 7 years (SE = 0.16) compared with 7.4 years (SE 0.12) for the typically developing controls (log-rank test $\chi^2 = 10, P < 0.001$); Fig. 4. The pattern of results held when the degree of motion artifact was entered into the regression equation (see SI Figs. 5 and 6).

**Discussion**

Cortical development in children with ADHD lagged behind that of typically developing children by several years. However, the ordered sequence of regional development, with primary sensory and motor areas attaining their peak cortical thickness before high-order association areas, was similar in both groups, suggesting that ADHD is characterized by delay rather than deviance in cortical maturation. This contrasts with other neurodevelopmental disorders such as autism in which there appears to be a dramatic shift of brain growth curves to the right along the age axis, resulting in peak brain volumes being reached at a
much earlier age—the opposite of the pattern we note in ADHD (18, 19).

The cortical maturation delay in ADHD was most prominent in the lateral prefrontal cortex, the region with the most consistent reports of structural anomalies in the disorder (11, 20), particularly within the superior and dorsolateral prefrontal regions (21–23). The prefrontal cortex supports a host of cognitive functions, such as the ability to suppress inappropriate responses and thoughts (24, 25), the executive "control" of attention (26), evaluation of reward contingencies (27, 28), higher-order motor control (5), and working memory (29). Deficits in these cognitive functions have all been implicated in the pathogenesis of ADHD (30), and prefrontal cortical hypoactivation in children with ADHD during performance of many of these tasks is a relatively consistent finding (10).

Delay was also found in the temporal cortex, most prominently in the posterior portions of the middle/superior temporal gyrus bilaterally, relatively circumscribed on the left, and with more posterior extension on the right. Structural change in the temporal lobes is a common finding in studies of ADHD, from the level of the entire lobe (11) through more focal gray matter density and cortical thickness anomalies (31, 32) and may have metabolic (9, 33), functional (10, 34–36), and electrophysiological correlates (37, 38). A unifying feature of the frontal and temporal regions with greatest maturational delay is the involvement of heteromodal cortex (39). These are interconnected cortical regions that integrate information from lower-order sensory areas giving higher-order percepts that guide the control of attention and action. Structural anomalies of this system have been implicated in the pathogenesis of ADHD (31).

By contrast, the primary motor cortex was the only cortical area in which the ADHD group showed slightly earlier maturation. It is possible that the combination of early maturation of the primary motor cortex with late maturation of higher-order motor control regions may reflect or even drive the excessive and poorly controlled motor activity cardinal to the syndrome.

Reaching peak cortical thickness at a younger age also means the typically developing children enter earlier the phase of cortical thinning that dominates adolescence (40, 41). Because of the limited age range, we were not able to define the age at which the adolescent phase of cortical thinning levels off, transitioning into stable adult cortical dimensions. We predict that the age of reaching this essentially static adult phase would also be later in the subjects with ADHD.

To our knowledge, neuroanatomic evidence supportive of the theory of delay in cortical maturation in ADHD has not been previously reported. The use of a cortical measure that affords exquisite spatiotemporal resolution allows us to demonstrate considerable variability in timing of cortical maturation within each lobe not detectable by our earlier lobar volumetric analyses (11). Additionally, we are able to localize the greatest maturational delay to prefrontal cortical regions implicated in the pathogenesis of ADHD.

In other work on a subsample of subjects with clinical outcome data from this cohort, we were only able to detect linear patterns of change in cortical thickness (and thus could not define the age of peak cortical thickness) and found generally parallel trajectories with the exception of a region in the right parietal cortex (12). By including additional subjects, we are able to detect higher-order effects of age and thus map out diagnostic regional differences in the age of attaining peak cortical thickness (12). Because we lacked clinical outcome data on the majority of the ADHD subjects in the current study, we were unable to examine the possibility that good or poor clinical outcome is linked to differences in the timing of key developmental markers, such as the age of peak cortical thickness.

Returning to the central finding, the generally older age of attaining peak cortical thickness in ADHD presumably represents a temporal shift in the balance between the cellular

**Fig. 3.** Kaplan–Meier curves illustrating the proportion of cortical points that had attained peak thickness at each age for all cerebral cortical points (Left) and the prefrontal cortex (Right). The median age by which 50% of cortical points had attained their peak differed significantly between the groups (all P < 1.0 × 10⁻²⁰).

**Fig. 4.** Regions where the ADHD group had early cortical maturation, as indicated by a younger age of attaining peak cortical thickness.
processes that result in an initial increase and later decrease in cortical thickness. The exact nature of these processes in typically developing children is yet to be determined. Extrapolating from animal studies, the increase in cortical thickness may be driven by mechanisms such as dendritic spine growth and elaboration of supporting glia and vasculature (42, 43). Cortical thinning in adolescence may reflect intracortical myelination and the use-dependent selective elimination of synapses that may help create and sculpt neural circuits, including those supporting cognitive abilities (44–46). Turning to ADHD, animal models are mostly based on perturbations in monoaminergic neurotransmission arising in response to either early insults (e.g., induced transient hyperthyroidism, or neonatal 6-OHDA lesions) or anomalies of neurotransmitters (such as the 160-bp insertion in exon 3 of the dopamine transport gene in the spontaneously hypertensive rat) (47, 48). How such changes might influence the dynamics of cortical development remains unclear but would be an important area for future research.

What etiological factors might underpin this delay? Trophic effects of treatment with psychostimulants in the ADHD group are possible but unlikely, given our previous reports of no effect of psychostimulants on gray matter volume (11). Because our studies have been observational, however, any conclusions about stimulants are tentative. Our overall results cannot be attributed to group differences in intelligence and gender, which, although they affect cortical development (14, 41, 49, 50), were strictly controlled in our design. Genetic factors will certainly play a role, with a perturbation in the developmental sequence of the activation and deactivation of genes that sculpt cortical architecture. In this context, neurotrophins, essential for the proliferation, differentiation, and survival of neuronal and nonneuronal cells, emerge as promising candidates, and, indeed, polymorphisms within the brain-derived neurotrophic factor and nerve growth-factor 3 genes have already been tentatively linked with ADHD (51, 52).

Trajectories of brain development built on longitudinal and cross-sectional neuroanatomic data sets are providing rich insights into ADHD. Not only do they inform key debates that have existed since the earliest descriptions of the disorder (2), but they may also guide the future search for factors that delay, rather than derail, cortical development.

Table 1. Demographic and clinical details of the subjects

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>ADHD, N = 223</th>
<th>Controls, N = 223</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at initial scan, yr, mean (SD)*</td>
<td>10.2 (3.2)</td>
<td>10.6 (3.5)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male:female</td>
<td>141:82</td>
<td>141:82</td>
</tr>
<tr>
<td>Estimated IQ,' mean (SD)</td>
<td>109 (15)</td>
<td>111 (13)</td>
</tr>
<tr>
<td>Comorbid diagnoses</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oppositional defiant disorder, no. (%)</td>
<td>77 (35)</td>
<td>NA</td>
</tr>
<tr>
<td>Conduct disorder, no. (%)</td>
<td>15 (7)</td>
<td>NA</td>
</tr>
<tr>
<td>Learning disorder, no. (%)</td>
<td>19 (9)</td>
<td>NA</td>
</tr>
<tr>
<td>Mood disorder, no. (%)</td>
<td>8 (4)</td>
<td>NA</td>
</tr>
<tr>
<td>Anxiety disorder, no. (%)</td>
<td>13 (6)</td>
<td>NA</td>
</tr>
<tr>
<td>Tic (NOS), no. (%)</td>
<td>14 (6)</td>
<td>NA</td>
</tr>
<tr>
<td>Clinical details</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinical Global Assessment Scale, mean (SD)</td>
<td>48 (7)</td>
<td>NA</td>
</tr>
<tr>
<td>CBCL Attention Problems T score, mean (SD)</td>
<td>71 (8)</td>
<td>NA</td>
</tr>
<tr>
<td>TRF Attention Problems T score, mean (SD)</td>
<td>66 (10)</td>
<td>NA</td>
</tr>
<tr>
<td>Prior stimulant treatment, no. (%)</td>
<td>108 (66)</td>
<td>NA</td>
</tr>
</tbody>
</table>

NA, not applicable; NOS, not otherwise specified; CBCL, Child Behavior Checklist; TRF, Teacher Report Form.

*P (ADHD vs controls), t (444) = −1.4, P = 0.16.

†P (ADHD vs controls), t (426) = −1.5, P = 0.14

Methods

Subjects. The clinical group comprised 223 children and adolescents with Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV)-defined ADHD. Diagnosis was based on the Parent Diagnostic Interview for Children and Adolescents (53), Conner’s Teacher Rating Scales (54), and the Teacher Report Form; see Table 1. Exclusion criteria were IQ under 80 and evidence of medical or neurological disorders. Two hundred five (92%) had combined-type ADHD at baseline, 16 (6%) had inattentive subtype, and 5 (2%) had hyperactive/impulsive subtype. One hundred fifty-four unrelated singletons and 25 sets of affected singleton siblings (with 53 individuals) and 16 twin-births (one child per twin pair) were included. Typically developing controls were recruited, and each subject completed the Childhood Behavior Checklist as a screening tool and then underwent a structured diagnostic interview by a child psychiatrist to rule out any psychiatric or neurological diagnoses (55). The typically developing participants in this study were matched to the ADHD group on gender, age, and intelligence as measured by age-appropriate version of the Wechsler Intelligence Scales. There were 169 singletons, 17 sets of unaffected siblings (with 38 individuals), and 16 twin births (one child per twin pair). The institutional review board of the National Institutes of Health approved the research protocol, and written informed consent and assent to participate in the study were obtained from parents and children, respectively.

The total number of subjects scanned at each age is given in Table 2, which also shows the numbers of subjects undergoing repeated scanning and the mean age at each wave of scan acquisition. The mean interscan interval was 2.9 years (SD 1.5) for the ADHD group and 2.8 years (SD 1.4) for the typically developing controls [t (317) = 1.3, P = 0.2].

Neuroimaging. All children had neuroanatomic magnetic resonance imaging on the same 1.5-T General Electric Signa scanner throughout the study. Imaging parameters were echo time of 5 ms, repetition time of 24 ms, flip angle of 45°, acquisition matrix of 256 × 192, number of excitations equaling 1, and 24-cm field of view. Head placement was standardized as described (56). The same 1.5-T General Electric Signa scanner was used throughout the study. The native MRI scans were registered into standardized stereotaxic space by using a linear transformation and corrected for...
preserves cortical topological features and represents considerably less cortex than the equivalent volumetric Gaussian blurring kernel (60). All scans were rated for degree of motion artifact (none, mild, moderate, or severe), as detailed in SI Text and ref. 61. Scans with moderate or severe motion artifact were excluded from further analyses; scans with mild motion artifact were included.

Statistical Analyses. First, we determined developmental trajectories, using mixed model regression analysis that allows the inclusion of multiple measurements per person, missing data, and irregular intervals between measurements, thereby increasing statistical power (62). A random effect for each individual was nested within a random effect for each family, thus accounting for both within-person and within-family dependence. Our classification of developmental trajectories was based on a step-down model selection procedure: At each cortical point, we modeled cortical thickness by using a mixed-effects polynomial regression model, testing for cubic, quadratic, and linear age effects. If the cubic age effect was not significant at $P < 0.05$, it was removed, and we stepped down to the quadratic model and so on. In this way, we were able to classify the development of each cortical point as being best explained by a cubic, quadratic, or linear function of age. A quadratic model proved appropriate for much of the cortex, in which 4th cortical thickness of the $i$th individual in the $j$th family was modeled as Thickness$_{ijk} = \text{intercept} + d_{ij} + \beta_1(\text{age} - \text{mean age}) + \beta_2(\text{age} - \text{mean age})^2 + e_{ijk}$, where $d_{ij}$ are nested random effects modeling within-person and within family dependence, the intercept and $\beta$ terms are fixed effects, and $e_{ijk}$ represents the residual error. Specifically, for both the ADHD and typically developing controls, a quadratic model was appropriate throughout most of the lateral prefrontal and medial prefrontal cortex, the superior and middle temporal cortex, superior and middle occipital cortex, and angular and supramarginal gyri. The ADHD group showed a linear fit in the superior parietal lobules and postcentral gyri, unlike the typically developing controls, for whom a quadratic model held. The analyses were repeated, entering the degree of motion artifact into the regression equation.

Next, the age of reaching peak cortical thickness for each group was calculated in these regions from the first-order derivatives of the fitted curves and illustrated through dynamic time-lapse sequences (‘movies’). Kaplan–Meier curves were constructed showing the proportion of cortical points that had reached peak cortical thickness throughout the age range covered. The significance of the group difference in the median age by which half of the cortical points had attained their peak thickness was calculated by using the log-rank (Mantel–Cox) test. Brain maps show the regions where the ADHD group attained peak thickness at either an earlier or later age.

We thank F. X. Castellanos for initiating the study and for advice and support and the children and their families who participated in the study. This work was supported by the Intramural Research Program of the National Institutes of Health. The sponsor of the study had no role in study design, data interpretation, or writing of the report.

Table 2. Scan acquisition details

<table>
<thead>
<tr>
<th>Age Range</th>
<th>ADHD</th>
<th>Typically developing controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 and under</td>
<td>13</td>
<td>24</td>
</tr>
<tr>
<td>6</td>
<td>18</td>
<td>24</td>
</tr>
<tr>
<td>7</td>
<td>37</td>
<td>28</td>
</tr>
<tr>
<td>8</td>
<td>44</td>
<td>24</td>
</tr>
<tr>
<td>9</td>
<td>45</td>
<td>31</td>
</tr>
<tr>
<td>10</td>
<td>40</td>
<td>39</td>
</tr>
<tr>
<td>11</td>
<td>33</td>
<td>38</td>
</tr>
<tr>
<td>12</td>
<td>33</td>
<td>38</td>
</tr>
<tr>
<td>13</td>
<td>24</td>
<td>33</td>
</tr>
<tr>
<td>14</td>
<td>29</td>
<td>30</td>
</tr>
<tr>
<td>15</td>
<td>22</td>
<td>30</td>
</tr>
<tr>
<td>16</td>
<td>17</td>
<td>25</td>
</tr>
<tr>
<td>17</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>18</td>
<td>13</td>
<td>17</td>
</tr>
<tr>
<td>19</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>20 and over</td>
<td>13</td>
<td>13</td>
</tr>
</tbody>
</table>

No. of subjects at each wave of scan acquisition†

- **Time 1**: 223
- **Time 2**: 111
- **Time 3**: 59
- **Time 4**: 15

Mean age at each scan, yr (SD)

- **Time 1**: 10.2 (3.2)
- **Time 2**: 12.8 (3.6)
- **Time 3**: 15.3 (3.7)
- **Time 4**: 17.9 (3.6)

*Absolute number of scans obtained at each age for each group.
†Number of subjects in each group who had scans at each wave of acquisition and the mean age (and SD) of each wave.

Supplementary Figure 1: Regions where the ADHD group had delayed cortical maturation, as indicated by an older age of attaining peak cortical thickness. The regression equation included the degree of motion artifact. As be seen through comparison with Figure 2, there was little change in the pattern of cortical delay.
Supplementary Figure 2: Regions where the ADHD group had earlier cortical maturation, as indicated by an older age of attaining peak cortical thickness. The regression equation included the degree of motion artifact. As seen through comparison with Figure 4, there was little change in the pattern of earlier cortical maturation in the ADHD group.