ABSTRACT

BACKGROUND: The COVID-19 pandemic has caused significant stress and disruption for young people, likely leading to alterations in their mental health and neurodevelopment. In this context, it is not clear whether youths who lived through the pandemic and its shutdowns are comparable psychobiologically to their age- and sex-matched peers assessed before the pandemic. This question is particularly important for researchers who are analyzing longitudinal data that span the pandemic.

METHODS: We compared carefully matched youths assessed before the pandemic (n = 81) and after the pandemic-related shutdowns ended (n = 82).

RESULTS: We found that youths assessed after the pandemic shutdowns had more severe internalizing mental health problems, reduced cortical thickness, larger hippocampal and amygdala volume, and more advanced brain age.

CONCLUSIONS: The COVID-19 pandemic not only appears to have led to poorer mental health and accelerated brain aging in adolescents, but it also poses significant challenges to researchers analyzing data from longitudinal studies of normative development that were interrupted by the pandemic.

https://doi.org/10.1016/j.bpsgos.2022.11.002
larger longitudinal study assessing the effects of early life stress on psychobiology across puberty (N = 214) (12–14).

Exclusion criteria were postpubertal status, nonfluency in English, inability to undergo magnetic resonance imaging (MRI), and history of neurological disorder or major medical illness. Participants were invited to return for follow-up assessments approximately every 2 years; however, the approximately 1-year-long COVID-19 pandemic shutdown beginning in March 2020 interrupted participants’ in-person assessments for the youngest pre-COVID participants were too young to be matched to the peri-COVID group (and we were compensated for their time. All study procedures were approved by the Stanford University Institutional Review Board.

From this larger cohort, we constructed 2 matched subsamples using data collected either before the pandemic (from November 2016 to November 2019; pre-COVID group, n = 81) or during the pandemic but following the end of the Bay Area shutdown (from October 2020 to March 2022; peri-COVID group, n = 82). We conducted these subsamples to maximize group sizes and to match the 2 groups on sex, age, pubertal status, race/ethnicity, parental education, annual household income, and severity of early life stress based on panel ratings of participants’ responses to interview (12,13).

Specifically, we attempted to match the peri-COVID participants with pre-COVID participants with respect to age and sex as closely as possible at the group level. Not all peri-COVID participants could be matched to pre-COVID participants given their older age, and not all pre-COVID participants were needed to be matched to the smaller peri-COVID group (and the youngest pre-COVID participants were too young to be matched to the peri-COVID participants). For the mental health symptoms sample of 81 pre-COVID and 82 peri-COVID participants (see below), we excluded from analyses 50 pre-COVID and 12 peri-COVID participants who could not be appropriately matched. For the neuroimaging sample, we were able to match by age and sex 64 of the 104 participants who underwent scanning in the peri-COVID period with 64 pre-COVID participants.

Mental Health Symptoms

Participants self-reported their depressive symptoms using the 10-item version of the Children’s Depression Inventory (16). This widely used reliable measure (17) has been shown to have convergent validity with clinician ratings of depression symptoms and diagnosis (18). We assessed anxiety symptoms using total score of the Social Anxiety and Physical Symptoms subscales of the Multidimensional Anxiety Scale for Children (MASC) (19). The full MASC assesses a wide range of anxiety symptoms, including symptoms that are not as relevant for the age range of our participants (e.g., separation anxiety). For this reason and to reduce participant burden, we administered only the Social Anxiety and Physical Symptoms subscales of the MASC for this study; therefore, the MASC total score in this study reflects the sum of these 2 subscales. Finally, we assessed internalizing and externalizing symptoms using the validated subscales of the Youth Self-Report version of the Child Behavior Checklist (20).

Neuroimaging

A subset of the participants (matched n = 64 per group) completed a T1-weighted MRI scan at the Center for Cognitive and Neurobiological Imaging at Stanford University. All participants in the pre-COVID group completed their scans using a 3T Discovery MR750 (GE Medical Systems). As of March 16, 2020, the Discovery MR750 was upgraded to an ultra-high performance system. Thus, all peri-COVID participants were scanned on the upgraded scanner. Participants in both groups were scanned using a 32-channel head coil (Nova Medical, Inc.). Prior work suggests that FreeSurfer-based cortical thickness and subcortical measures are highly reliable across scanner upgrades (21–23). For example, Han et al. (22) did not find evidence that scanner upgrades introduce bias for cortical thickness measures, and Brown et al. (23) found that hippocampal measures are reliable across scanners. In addition, we conducted analyses with our own data to assess potential differences in T1-weighted image quality related to scanner upgrade. Specifically, in a subset of 31 participants with imaging data before and after the scanner upgrade, we tested within-participant changes in gray–white matter contrast-to-noise ratio using FreeSurfer’s mri_cnr quality metric command. We did not find significant differences in contrast-to-noise ratio from pre- to post-upgrade in either the left (t\(_{30} = 0.81, p = .425\)) or the right (t\(_{30} = 0.66, p = .513\)) hemisphere. Thus, the scanner upgrade does not appear to have introduced a systematic bias in image quality. Whole-brain T1-weighted images were collected for all participants using the following spoiled gradient echo pulse sequence: 186 sagittal slices, repetition time/echo time/inversion time = 6.24/2.34/450 ms, flip angle = 12°, voxel size = 0.9 mm \(\times\) 0.9 mm \(\times\) 0.9 mm, scan duration = 315 seconds. The spoiled gradient echo sequence was repeated up to 2 additional times if the first acquisition did not yield clear images. For each participant with multiple acquisitions, the single spoiled gradient echo image with the clearest structural boundaries (i.e., that was free from motion or other artifacts) was used for further analysis.

Segmentation of Cortical and Subcortical Regions

We used FreeSurfer version 6.0 (https://surfer.nmr.mgh.harvard.edu) recon-all function to automatically skull strip and segment cortical and subcortical volumes from the T1-weighted structural images (24), which has been shown to have acceptable scan-rescan reliability (21) and comparable accuracy to manual labeling techniques (24–26). We implemented structural image processing protocols established by ENIGMA (https://enigma.ini.usc.edu/protocols/imaging-protocols/) to extract and perform quality assurance checks on the cortical thickness and subcortical volume estimates from the FreeSurfer outputs. Using FreeView image viewer, all cortical and subcortical outputs were visually inspected to quality check for processing and segmentation errors. As previously described (27,28), we converted gray matter volumes from each hemisphere into z scores; volumes with z scores greater than 2.5 or less than –2.5 were visually examined again for accuracy, and any segmentations that failed any of these steps were removed from final analyses. We focused on mean cortical thickness (average of

Biological Psychiatry: Global Open Science  October 2023; 3:912–918  www.sobp.org/GOS

913
cortical thickness values across individual regions as defined by the Desikan-Killiany atlas (23) and unstandardized residuals of subcortical volumes regressed on total intracranial volume.

### Brain Age Gap Estimates

Based on cortical and subcortical features, we computed BrainAGE values for male and female participants using sex-specific machine learning–based models developed by the ENIGMA-Brain Age working group (30). These models use data from 14 subcortical gray matter volumes, 2 lateral ventricles, 68 cortical thickness measures, 68 surface area measures, and total intracranial volume to predict chronological age (i.e., predicted brain age). We computed brain age gap estimates by subtracting chronological age from predicted brain age. Given that BrainAGE values are often overestimated in younger individuals and underestimated in older individuals, Le et al. (31) proposed adjusting for chronological age in analyses of BrainAGE. Therefore, we regressed gap estimates onto chronological age and used the unstandardized residuals as the BrainAGE outcome variable in our statistical analyses.

### Statistical Approach

All statistical analyses were conducted using R version 4.0.2 (32). To examine whether adolescents who experienced the pandemic differed from their pre-pandemic peers, we conducted between-group tests on measures of internalizing and externalizing symptoms, cortical thickness, and subcortical volume (regions of interest were bilateral amygdala, hippocampus, and nucleus accumbens). For analyses of mental health problems, we first conducted a one-way multiple analysis of variance test to examine whether there were group differences in overall mental health scores across measures. We conducted follow-up independent sample t tests to examine whether the pre-COVID and peri-COVID groups differed in specific aspects of mental health as assessed by different measures. We repeated these steps for analyses of brain metrics. Given our expectations based on recent work suggesting that mental health problems have increased during the pandemic (5), we used one-tailed hypothesis tests for follow-up analyses of mental health outcomes. We used two-tailed hypothesis tests for follow-up analyses of brain outcomes.

### Table 1. Participant Characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mental Health Subsample&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Brain Subsample&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-COVID, n = 81</td>
<td>Peri-COVID, n = 82</td>
</tr>
<tr>
<td>Sex, Female</td>
<td>n = 82</td>
<td>n = 82</td>
</tr>
<tr>
<td>Age, Years</td>
<td>15.87 (1.14)</td>
<td>16.17 (0.93)</td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian/Asian American</td>
<td>13 (16%)</td>
<td>11 (13%)</td>
</tr>
<tr>
<td>Biracial</td>
<td>9 (11%)</td>
<td>21 (26%)</td>
</tr>
<tr>
<td>Black/African American</td>
<td>6 (7%)</td>
<td>7 (9%)</td>
</tr>
<tr>
<td>Hispanic/Latinx</td>
<td>5 (6%)</td>
<td>7 (9%)</td>
</tr>
<tr>
<td>Other&lt;sup&gt;c&lt;/sup&gt;</td>
<td>5 (6%)</td>
<td>3 (4%)</td>
</tr>
<tr>
<td>White</td>
<td>43 (53%)</td>
<td>33 (40%)</td>
</tr>
<tr>
<td>Income-to-Needs Ratio</td>
<td>1.37 (0.53)</td>
<td>1.27 (0.54)</td>
</tr>
<tr>
<td>Early Life Stress</td>
<td>6.60 (4.89)</td>
<td>6.31 (4.97)</td>
</tr>
<tr>
<td>Parental Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No GED/high school diploma</td>
<td>0 (0%)</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>GED/high school diploma</td>
<td>0 (0%)</td>
<td>4 (5%)</td>
</tr>
<tr>
<td>Some college</td>
<td>10 (12%)</td>
<td>15 (18%)</td>
</tr>
<tr>
<td>2-Year college degree</td>
<td>5 (6%)</td>
<td>7 (9%)</td>
</tr>
<tr>
<td>4-Year college degree</td>
<td>29 (36%)</td>
<td>25 (30%)</td>
</tr>
<tr>
<td>Master’s degree</td>
<td>30 (37%)</td>
<td>20 (24%)</td>
</tr>
<tr>
<td>Professional degree</td>
<td>1 (1%)</td>
<td>5 (6%)</td>
</tr>
<tr>
<td>Doctorate</td>
<td>4 (5%)</td>
<td>2 (2%)</td>
</tr>
<tr>
<td>Not reported</td>
<td>2 (2%)</td>
<td>3 (4%)</td>
</tr>
<tr>
<td>COVID-19 Impact</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individual diagnosis</td>
<td>N/A</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>Household diagnosis</td>
<td>N/A</td>
<td>3 (4%)</td>
</tr>
<tr>
<td>Financial strain</td>
<td>N/A</td>
<td>13 (16%)</td>
</tr>
<tr>
<td>Job loss</td>
<td>N/A</td>
<td>7 (9%)</td>
</tr>
</tbody>
</table>

Values are mean (SD) or n (%).
GED, General Educational Development; N/A, not applicable.
<sup>a</sup>Subsample of participants who completed the measures of mental health.
<sup>b</sup>Subsample of participants who also successfully completed the neuroimaging protocol.
<sup>c</sup>Primarily participants who identified as multiracial.
COVID-19, Mental Health, and Brain Maturation

RESULTS

Participant Characteristics

The demographic and clinical characteristics of the pre- and peri-COVID subgroups of participants are presented in Table 1. Participants’ parents reported on their annual household income, from which we computed an income-to-needs ratio by dividing the midpoint of their reported income bin by the low-income value for Santa Clara County. Importantly, this calculation considers the number of people in the home and the time period in which the study occurred [https://www.huduser.gov/portal/datasets/il/il2017/2017summary.odn; (33)]. Attesting to the success of our careful matching procedure, there were no significant group differences in participant characteristics between the pre-COVID and peri-COVID subgroups for either the mental health or the brain samples (all individual ps > .06).

Mental Health

Group differences on the mental health measures are presented in Figure 1. A one-way multivariate analysis of variance indicated that the pre-COVID (n = 81) and peri-COVID (n = 82) groups differed significantly in their self-reported mental health difficulties ($F_{2,158} = 2.67, p = .034$). Follow-up t tests showed that the peri-COVID group reported more severe symptoms of anxiety ($t_{161} = 3.15, p < .001$; Cohen’s $d = 0.49$), depression ($t_{161} = 1.92, p = .029; d = 0.30$), and internalizing problems ($t_{161} = 1.77, p = .039; d = 0.28$); the 2 groups did not differ in externalizing problems ($t_{161} = 1.25, p = .108$).

Brain Metrics

Group differences in cortical thickness, subcortical volumes, and BrainAGE are shown in Figure 2. A one-way multivariate analysis of variance conducted on all the brain metrics yielded a significant difference between the pre-COVID (n = 61–64) and peri-COVID (n = 63–64) groups ($F_{5,161} = 7.13, p < .001$). Follow-up tests indicated that the peri-COVID group had reduced bilateral cortical thickness ($t_{122} = 3.67, p < .001; d = 0.66$) and, controlling for intracranial volume, larger bilateral hippocampal volume ($t_{125} = 3.56, p < .001; d = 0.63$) and bilateral amygdala volume ($t_{125} = 2.01, p = .047; d = 0.36$); the 2 groups did not differ in bilateral nucleus accumbens volume ($t_{125} = 0.68, p = .248; d = 0.12$). Finally, despite the fact that the 2 groups were matched on age and other relevant demographic characteristics, adolescents in the peri-COVID group had an older BrainAGE than their peers who were assessed before the pandemic ($t_{125} = 2.31, p = .022; d = 0.41$).

Interval Between COVID-19 Shutdown and Peri-COVID Assessments

Finally, given the possibility that participants’ mental health difficulties and their brain metrics increased with the duration of the pandemic, we examined our clinical functioning and brain metrics as a function of time since the Bay Area shelter-in-place orders were initiated (March 17, 2020). The peri-COVID participants completed measures of clinical functioning between January 10, 2021, and September 30, 2021, and MRI scans between October 13, 2020, and March 22, 2022. Within the peri-COVID group, we examined associations between the number of days from the start of shelter-in-place orders to the dates that participants completed measures of psychopathology (mean [SD] = 346.49 [131.70] days; range, 133–720 days). There were no significant associations between this interval and participants’ scores on the measures of depression ($r_{62} = 0.01, p = .901$), anxiety ($r_{62} = 0.06, p = .544$), internalizing symptoms ($r_{62} = 0.07, p = .506$), or externalizing symptoms ($r_{62} = 0.00, p = .980$). We repeated these analyses for the brain metrics (mean [SD] interval = 379.00 [119.24] days; range, 210–735 days). Again, there were no significant associations between the interval and residuals of amygdala volume ($r_{62} = 0.01, p = .935$), hippocampal volume ($r_{62} = 0.15, p = .245$), nucleus accumbens volume ($r_{62} = 0.05, p = .681$), mean cortical thickness ($r_{61} = 0.04, p = .303$), or residuals of BrainAGE ($r_{62} = -0.09, p = .459$).
DISCUSSION

In addition to replicating prior findings that the pandemic has adversely affected the mental health of young people (5), we found that adolescents assessed during the pandemic have neuroanatomical features that are more typical of individuals who are older or who experienced significant adversity in childhood. Compared with carefully matched peers assessed before the pandemic, adolescents assessed during the pandemic showed signs of advanced cortical thinning and had larger bilateral hippocampal and amygdala volumes. Given that volume in these structures typically increases over adolescence (34), these neural alterations may reflect accelerated brain maturation in the context of the pandemic. Indeed, adolescents assessed during the pandemic also had larger positive brain age gap estimates, indicative of older-appearing brains.

It appears, therefore, that the pandemic not only has adversely affected mental health of adolescents, but also has accelerated their brain maturation. These findings have critical implications for researchers who are conducting longitudinal studies that were interrupted due to pandemic-related shutdowns. In our own longitudinal study, we had been assessing a sample of approximately 200 adolescents at each of 4 time points, at 2-year intervals, to examine the effects of early adversity on trajectories of neurodevelopment and clinical symptoms. At the time of the shutdown, we were two thirds of the way through the third assessment, when our participants were 13 to 17 years of age. We had originally planned to simply use participants’ age in analyzing trajectories from our 4 time points of data. Although some participants would have had a longer interval than others between assessments that bracketed the shutdown, we would control statistically for those differences. It is important to recognize that this analytic approach assumes that, for example, 16-year-olds who were assessed after the shutdown ended are equivalent in their clinical functioning and neurodevelopment to 16-year-olds who were assessed before the pandemic and would simply be grouped together. Our results suggest that this assumption is not correct. Rather, the pandemic appears to have altered adolescent mental health and neurodevelopment, at least in the short term, which will present a challenge for researchers in analyzing longitudinal data from studies of normative development that were interrupted by the pandemic.

In order to not confound age-related changes in brain maturation with experiences and consequences of the COVID-19 pandemic, some researchers, including our group, have used a dummy-coded variable to control statistically for whether participants were assessed before or during the pandemic [e.g. (35)]. Nevertheless, restrictions around

Figure 2. Raw data are plotted for visualization. Significance levels are based on group differences in subcortical volumes (in mm) adjusted for intracranial volume, cortical thickness, and brain age gap estimation (BrainAGE) adjusted for chronological age. *p < .05, ***p < .001. n.s., not significant.
COVID-19 are constantly changing; therefore, additional measures may need to be used as covariates, including the interval between shelter-in-place orders and time of assessment as well as the nature and severity of the individual’s stress and experience during the pandemic (e.g., COVID-19 infection, upheaval in living situation, financial strain). We should note that our sample is of relatively high socioeconomic status and represents the racial/ethnic composition of the San Francisco Bay Area. Researchers have reported that sample composition influences age-related effects on brain structure (36) and, more specifically, that the psychosocial and health consequences of the pandemic have been more severe among individuals from socially marginalized groups [e.g., lower socioeconomic status (37–39)]. Therefore, it is important that investigators examine the effects of the COVID-19 pandemic on psychopathology and brain metrics in more diverse samples of adolescents that are representative of the broader population.

Another critical task for future research is to determine whether these alterations are temporary effects of the pandemic or stable changes that will characterize the current generation of youths. If these changes are found to be enduring, accounting for and interpreting data acquired during this period will require additional attention and consideration. For example, as more researchers publish data concerning normative developmental trajectories of MRI-derived anatomical features [e.g., (40)], it will be possible to compare COVID-19–impacted neurodevelopmental trajectories with normative trajectories and, indeed, to compute COVID–19–adjusted metrics of brain maturation. Regardless, however, we emphasize that it is important that we continue to follow and assess individuals who were recruited and assessed before the pandemic; this type of research offers the strongest possibility for us to examine the effects of a major stressor experienced on a global scale.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by the National Institutes of Health (Grant Nos. R37MH101495 [to IHG] and K01MH117442 [to TCH]).

The National Institutes of Health played no role in the study design, data collection, analysis, or preparation, review, or approval of the manuscript. IHG designed the study and helped to write the manuscript; JGM, LRB, and TCH helped to collect the data and write the manuscript. We thank the youths and their families for participating in this research. The authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Department of Psychology, Stanford University, Stanford, California (IHG, JGM, LRB, SMC, LAC, JMG); and Department of Psychiatry and Weill Institute for Neurosciences, University of California San Francisco, San Francisco, California (TCH). Address correspondence to Ian H. Gotlib, Ph.D., at ian.gotlib@stanford.edu.

Received Aug 8, 2022; revised Nov 5, 2022; accepted Nov 7, 2022. Supplementary material cited in this article is available online at https://doi.org/10.1016/j.bpsgos.2022.11.002.

REFERENCES


