Archival Report

Biological Psychiatry

Neuroanatomical Diversity of Corpus Callosum and Brain Volume in Autism: Meta-analysis, Analysis of the Autism Brain Imaging Data Exchange Project, and Simulation

Aline Lefebvre, Anita Beggiato, Thomas Bourgeron, and Roberto Toro

ABSTRACT

BACKGROUND: Patients with autism have been often reported to have a smaller corpus callosum (CC) than control subjects.

METHODS: We conducted a meta-analysis of the literature, analyzed the CC in 694 subjects of the Autism Brain Imaging Data Exchange project, and performed computer simulations to study the effect of different analysis strategies.

RESULTS: Our meta-analysis suggested a group difference in CC size; however, the studies were heavily underpowered (20% power to detect Cohen's $d = .3$). In contrast, we did not observe significant differences in the Autism Brain Imaging Data Exchange cohort, despite having achieved 99% power. However, we observed that CC scaled nonlinearly with brain volume (BV): large brains had a proportionally smaller CC. Our simulations showed that because of this nonlinearity, CC normalization could not control for eventual BV differences, but using BV as a covariate in a linear model would. We also observed a weaker correlation of IQ and BV in cases compared with control subjects. Our simulations showed that matching populations by IQ could then induce artifactual BV differences.

CONCLUSIONS: The lack of statistical power in the previous literature prevents us from establishing the reality of the claims of a smaller CC in autism, and our own analyses did not find any. However, the nonlinear relationship between CC and BV and the different correlation between BV and IQ in cases and control subjects may induce artifactual differences. Overall, our results highlight the necessity for open data sharing to provide a more solid ground for the discovery of neuroimaging biomarkers within the context of the wide human neuroanatomical diversity.

Keywords: Autism, Brain volume, Computational neuroanatomy, Corpus callosum, Meta-analysis, Statistical power

http://dx.doi.org/10.1016/j.biopsych.2015.02.010

Autism spectrum disorders (ASD) are pervasive developmental disorders with qualitative impairments in social interaction and communication, along with restricted, repetitive, and stereotyped patterns of behavior. Several cognitive studies have suggested that difficulty integrating multiple sources of stimulation may be a common characteristic of ASD, which has led, for example, to the influential weak central coherence hypothesis (1). The neural basis of these difficulties has been hypothesized to be an imbalance between local and distant connections: local overconnectivity and long-distance underconnectivity (2,3). The connectivity hypothesis has been a major subject of study and discussion in ASD research (4,5).

The corpus callosum (CC)—the largest commissure connecting the left and right hemispheres—appeared then as a natural candidate to look for evidence of connectivity abnormalities. The CC exists exclusively within eutherian mammals (kangaroos and other marsupials lack a CC) and has been suggested to play an important role in the evolution of lateralization. The number of callosal axons is disproportionally smaller in large-brain mammals (like humans) compared with small-brain mammals (like mice). A smaller number of callosal fibers, and their increased length, could hinder the formation of interhemispheric synchronous neuronal populations, thus facilitating local recruitment and leading to functional lateralization (6–8). The CC has been, as a consequence, one of the most studied white matter tracts in ASD. Numerous reports have indeed described significantly smaller CC among patients compared with control subjects, and a series of studies have suggested a higher incidence of ASD within cases of CC agenesis or callosotomy patients (9–11).

However, many of these analyses have relied on small cohorts (~30 patients, ~30 control subjects), without statistical power to find even effect sizes as large as .5 standard deviations between groups. Despite the lack of power, studies often report statistically significant differences. A solution to the methodological problems associated with small cohorts has been recently proposed by the Autism Brain Imaging Data Exchange project (Abide) (12). Large cohorts are difficult to gather and analyze by any single research group. Abide provides open access to

SEE COMMENTARY ARTICLE ON PAGE 81

behavioral and neuroimaging data for almost 600 patients and 600 age-, sex-, and IQ- matched control subjects from an international consortium of 17 research groups. Abide provides the research community with the statistical power necessary to detect even small differences in case/control designs, and to use more sophisticated analysis strategies, providing a wider perspective on neuroanatomical diversity.

Here, we first present a review of studies of CC size differences in ASD. We observed a general lack of statistical power (only 20% power to detect two-sided differences of .3 standard deviations at .05 level of significance), which contrasted with the frequent report of significant findings (9 out of 17 studies). Next, we present our analysis of the diversity of the CC in Abide and differences related to scanning sites, age, sex, brain size polymorphism, and diagnostic group. Even though previous studies have reported diagnostic group differences as large as .3 to .7 standard deviations and despite having analyzed a number of subjects comparable with the sum of all previously studied, we did not find significant differences between patients with ASD and control subjects. Finally, we discuss possible ways in which analysis strategies, such as the normalization of CC size by total brain volume (BV) or the matching of subjects by IQ scores, could lead to artifactual differences in BV and CC size between patients and control subjects.

METHODS AND MATERIALS

Meta-analysis

We included all studies from the recent review by Frazier and Hardan (13) and searched PubMed (http://www.ncbi.nlm.nih.gov/pubmed/) for additional studies reporting differences in CC between patients with ASD and control subjects: (autism OR PDD OR "pervasive developmental disorder") AND "corpus callosum". Table 1 describes the cohort in the articles meta-analyzed (see Figure S1 and Supplemental Materials in Supplement 1 for further details on the inclusion procedure).

Analysis of Abide

We used FreeSurfer v5.1 (http://surfer.nmr.mgh.harvard.edu/) to process the 1102 subjects in Abide (http://fcon_1000.projects.nitrc.org/indi/abide) with T1-weighted magnetic resonance imaging data available. We developed an open online tool to visually control the accuracy of the segmentations (http://siphonophore.org/corpuscallosum). Based on this quality control, we excluded 380 subjects, 331 of them because of a mislabeled fornix that was included in the middle segment of the CC (Figure 1; Figure S2 in Supplement 1). Not excluding the subjects with a mislabeled fornix, however, did not change our results. We only included subjects from 7.5 to 40 years old. The final sample consisted of 694 subjects: 328 patients (290 male subjects, 38 female subjects) and 366 control subjects (304 male subjects, 62 female subjects). Full IQ (FIQ) was available for 672 subjects and verbal IQ (VIQ) and performance IQ (PIQ) for 538 subjects. Table 2 describes the groups of subjects retained from each scanning site. Statistical analyses were performed using JMP Pro 10.0.2 (http://www.jmp.com), R (http://www.r-project.org/), iPython (http://ipython.org/), and G*Power (http://gpower.hhu.de/) (see Supplemental Materials in Supplement 1 for further details on the analysis).

Simulations

Simulations were written in Python to analyze the effect on case/control comparisons produced by 1) normalization by

Table 1. Population in the Studies Included in the Meta-analysis

Reference	Age_{ASD} ± SD	Age_{Ctrl} ± SD	IQ Level (LF/HF)	N_{ASD} (F)	N_{Ctrl} (F)	Matching Strategy		
						Relative	GLM	IQ
Gaffney *et al.* 1987 (36)	11 ± 5	12 ± 5	LF/HF	13 (3)	35 (14)	No	No	No
Egaas *et al.* 1995 (37)	16 ± 10	16 ± 10	LF/HF	51 (6)	51 (6)	No	No	No
Piven *et al.* 1997 (38)	18 ± 5	20 ± 4	LF/HF (PIQ > 70)	35 (6)	36 (16)	No	Yes	No
Manes *et al.* 1999 (39)	14 ± 7	12 ± 5	LF	27 (5)	17 (6)	Yes	No	Yes
Elia *et al.* 2000 (40)	11 ± 4	11 ± 3	LF	22 (0)	11 (0)	No	No	No
Rice *et al.* 2005 (41)	12 ± 4	13 ± 4	HF/LF	12 (0)	8 (0)	No	Yes	No
Vidal *et al.* 2006 (42)	10 ± 3	11 ± 3	HF	24 (0)	26 (0)	Yes	Yes	Yes
Boger-Mediggo *et al.* 2006 (43)	4 ± 3	4 ± .5	–	45 (7)	26 (8)	Yes	Yes	No
Alexander 2007 (44)	16 ± 7	16 ± 6	HF (PIQ)	43 (–)	34 (–)	No	Yes	Yes
Just *et al.* 2007 (3)	27 ± 12	25 ± 10	HF	18 (1)	18 (3)	Yes	No	Yes
Hardan *et al.* 2009 (45)	11 ± 1	11 ± 1	HF	22 (0)	23 (0)	Yes	No	Yes
Freitag *et al.* 2009 (46)	18 ± 4	19 ± 1	HF	15 (2)	15 (2)	No	Yes	Yes
Keary 2009 (47)	20 ± 10	19 ± 9	HF	32 (2)	34 (2)	No	Yes	Yes
Anderson *et al.* 2011 (48)	22 ± 7	21 ± 7	HF/LF	53 (0)	39 (0)	No	No	Yes
Hong *et al.* 2011 (49)	9 ± 2	10 ± 2	HF	18 (0)	16 (0)	No	Yes	Yes
Frazier *et al.* 2012 (50)	11 (8–12)[a]	11 (7–13)[a]	HF	23 (0)	23 (0)	No	Yes	No
Prigge *et al.* 2013 (51)	14 ± 8	15 ± 7	(HF/HF PIQ > 70)	68 (0)	47 (0)	No	Yes	Yes

Matching strategy: Relative: differences in BV were accounted by dividing CC by BV; GLM: differences in BV were accounted by including BV as a covariate in a general linear model; IQ: differences in IQ were accounted for by matching groups by intelligence scores.

ASD, autism spectrum disorders; BV, total brain volume; CC, corpus callosum; Ctrl, control subjects; F, females; GLM, general linear model; HF, high functioning; LF, low functioning; PIQ, performance IQ.

[a]Age range.