

RESEARCH ARTICLE

WILEY

Adolescent to young adult longitudinal development of subcortical volumes in two European sites with four waves

Lea L. Backhausen^{1,2}  | Juliane H. Fröhner¹ | Hervé Lemaître^{3,4} |
 Eric Artiges^{5,6} | Marie-Laure Palillère Martinot^{5,7} | Megan M. Herting⁸  |
 Fabio Sticca⁹ | Tobias Banaschewski¹⁰ | Gareth J. Barker¹¹ |
 Arun L. W. Bokde¹² | Sylvane Desrivières¹³ | Herta Flor^{14,15} | Antoine Grigis³ |
 Hugh Garavan¹⁶ | Penny Gowland¹⁷ | Andreas Heinz¹⁸ | Rüdiger Brühl¹⁹ |
 Frauke Nees^{10,14,20} | Dimitri Papadopoulos-Orfanos³ | Luise Poustka²¹ |
 Sarah Hohmann^{10,22} | Lauren Robinson²³ | Henrik Walter¹⁸ |
 Jeanne Winterer^{18,24} | Robert Whelan²⁵ | Gunter Schumann^{13,26,27} |
 Jean-Luc Martinot⁵ | Michael N. Smolka¹  | Nora C. Vetter^{1,2,28}  | the IMAGEN Consortium

Correspondence

Lea L. Backhausen, Medical Faculty and University Hospital Carl Gustav Carus, TUD Dresden University of Technology, Fetscherstraße 74, 01307 Dresden, Germany. Email: lea_luise.backhausen@tu-dresden.de

Funding information

European Union-funded FP6 Integrated Project IMAGEN, Grant/Award Number: LSHM-CT-2007-037286; Horizon 2020 funded ERC Advanced Grant “STRATIFY”, Grant/Award Number: 695313; Human Brain Project, Grant/Award Numbers: 945539, 785907; Medical Research Council Grant “c-VEDA”, Grant/Award Number: MR/N000390/1; National Institute of Health (NIH), Grant/Award Number: R01DA049238; Bundesministerium für Bildung und Forschung, Grant/Award Numbers: 01GL1745B, 01EE1406B, 01EE1406A, 01EV0711, 01GS08152; the Deutsche Forschungsgemeinschaft, Grant/Award Numbers: NE 1383/14-1, 290210763, 402170461, 178833530, 186318919; the Medical Research Foundation and Medical Research Council, Grant/Award Numbers: MR/S020306/1, MR/R00465X/1; National

Abstract

Adolescent subcortical structural brain development might underlie psychopathological symptoms, which often emerge in adolescence. At the same time, sex differences exist in psychopathology, which might be mirrored in underlying sex differences in structural development. However, previous studies showed inconsistencies in subcortical trajectories and potential sex differences. Therefore, we aimed to investigate the subcortical structural trajectories and their sex differences across adolescence using for the first time a single cohort design, the same quality control procedure, software, and a general additive mixed modeling approach. We investigated two large European sites from ages 14 to 24 with 503 participants and 1408 total scans from France and Germany as part of the IMAGEN project including four waves of data acquisition. We found significantly larger volumes in males versus females in both sites and across all seven subcortical regions. Sex differences in age-related trajectories were observed across all regions in both sites. Our findings provide further evidence of sex differences in longitudinal adolescent brain development of subcortical regions and thus might eventually support the relationship of underlying brain development and different adolescent psychopathology in boys and girls.

Jean-Luc Martinot, and Michael N. Smolka contributed equally.

For affiliations refer to page 13

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *Human Brain Mapping* published by Wiley Periodicals LLC.

Institutes of Health (NIH) funded ENIGMA, Grant/Award Numbers: 1R56AG058854-01, 5U54EB020403-05; ANR, Grant/Award Numbers: AAPG2019, ANR-12-SAMA-0004; Eranet Neuron, Grant/Award Numbers: ANR-18-NEUR00002-01, AF12-NEUR0008-01; Fondation de France, Grant/Award Numbers: 2012-00033703, 00081242; Fondation pour la Recherche Médicale, Grant/Award Numbers: DPP20151033945, DPA20140629802; Mission Interministérielle de Lutte-contre-les-Drogues-et-les-Conduites-Addictives (MILDECA); Assistance-Publique-Hôpitaux-de-Paris and INSERM; Paris Sud University; Fondation de l'Avenir, Grant/Award Number: AP-RM-17-013; Fédération pour la Recherche sur le Cerveau; National Institutes of Health, Science Foundation Ireland, Grant/Award Number: 16/ERC/3797; NIH Consortium grant, Grant/Award Number: U54 EB020403

KEYWORDS

adolescence, longitudinal, neurodevelopment, sex differences, structural magnetic resonance imaging (sMRI), subcortical

Practitioner Points

- 503 participants from two large European samples similar in sociodemographics underwent up to four structural MRI scans from ages 14 to 24.
- Age trajectories in subcortical volumes differed between sexes in all regions of interest.
- Several regions differed between sites, possibly due to different imaging sequences.

1 | INTRODUCTION

Adolescence is a core incidence phase for the development of mental disorders (Paus et al., 2008). Mental disorders occur in around 20% of adolescents with important sex differences, that is, higher rates of internalizing disorders in girls and externalizing disorders in boys (Kessler et al., 2005; Zahn-Waxler et al., 2008). These sex differences might be mirrored in underlying sex differences in subcortical structural development (Gogtay & Thompson, 2010; Shaw et al., 2010) since alterations in subcortical structures have been associated with psychopathological disorders (Hoogman et al., 2017; Noordermeer et al., 2017). Hence, characterizing subcortical structural development and sex-related differences during adolescence is vital to better understand adolescent psychopathology.

Cross-sectional (Boedhoe et al., 2020; Brain Development Cooperative Group, 2012; Østby et al., 2009) and longitudinal studies (Dennison et al., 2013; Goddings et al., 2014; Herting et al., 2018; Narvacan et al., 2017; Tamnes et al., 2013; Wierenga et al., 2014) began to investigate the typical development of subcortical structures across adolescence, with some studies also analyzing sex effects (Godings et al., 2014; Herting et al., 2018; Narvacan et al., 2017; Wierenga et al., 2014; Wierenga, Bos, et al., 2018). These studies demonstrate differential trajectories for some subcortical structures. We review here the trajectories from mid-adolescence to young adulthood, specifically the age span of 14 to 24, which will also be covered by our study. Basal ganglia (mainly caudate nucleus, putamen, globus pallidus, and nucleus accumbens) showed volume reductions in cross-sectional studies (Brain Development Cooperative Group, 2012; Østby et al., 2009), including a recent mega-analysis (Dima et al., 2021). Longitudinal studies found similar results across adolescence for some basal ganglia structures (Dennison et al., 2013; Goddings et al., 2014; Narvacan et al., 2017; Raznahan et al., 2014; Tamnes et al., 2013; Wierenga et al., 2014; Wierenga, Sexton, et al., 2018; for methodical details of the longitudinal studies, including our sample, please see Table S1). However, divergent findings exist. The caudate nucleus remained stable across adolescence in

three studies (for both sexes: Narvacan et al., 2017; Wierenga, Bos, et al., 2018); for males only (Raznahan et al., 2014). For the globus pallidus three studies found a volume increase (Dennison et al., 2013; Wierenga et al., 2014; Wierenga, Bos, et al., 2018). A multi-site study found a descriptive, albeit non-significant increase (Herting et al., 2018). For the nucleus accumbens, a stable trajectory was found (for both sexes; Dennison et al., 2013 or males only; Herting et al., 2018) and the putamen revealed a non-linear increase (Wierenga, Bos, et al., 2018). For the thalamus, previous results converged in indicating volume decreases (Herting et al., 2018 for males only; Raznahan et al., 2014; Tamnes et al., 2013; Wierenga, Bos, et al., 2018) while Narvacan et al. (2017) only showed decreases in their cross-sectional and no changes in their longitudinal cohort. Diverging findings also exist regarding the trajectories of the amygdala and hippocampus across adolescence (Koolschijn & Crone, 2013; Østby et al., 2009; Sowell et al., 2002; Wierenga et al., 2014). While some cross-sectional (Durston et al., 2001; Giedd et al., 1996) and longitudinal studies (Dennison et al., 2013; Herting et al., 2018) found an increase, other longitudinal studies showed decreases (Tamnes et al., 2013) or little to no change (Dennison et al., 2013; Wierenga, Bos, et al., 2018). Wierenga et al. (2014) reported an increase for both the amygdala and hippocampus with a peak at about age 17/18 followed by a slight decrease.

1.1 | Sex differences

Regarding sex differences, cross-sectional and longitudinal studies consistently demonstrate larger subcortical volumes in boys versus girls (Dennison et al., 2013; Goddings et al., 2014; Herting et al., 2018; Narvacan et al., 2017; Raznahan et al., 2014; Tamnes et al., 2013; Wierenga et al., 2014; Wierenga, Bos, et al., 2018), ranging from 9% for the putamen to 15% for the amygdala (Narvacan et al., 2017).

Age-related trajectories of subcortical structures also exhibit sex differences, but previous studies have been rather inconsistent. Overall differences in trajectories between sexes were found in some

studies (Dennison et al., 2013; Goddings et al., 2014; Herting et al., 2018), for example, later peak volumes in males (Lenroot et al., 2007; Raznahan et al., 2014). The thalamus showed a decline either for females (Dennison et al., 2013) or males (Herting et al., 2018). A slight decrease for the amygdala and hippocampus was found in girls at about 15 years versus a continued increase for boys (Goddings et al., 2014). Differing nonlinear trajectories across sexes for the amygdala and hippocampus have been identified (Herting et al., 2018), however some studies also reported no sex differences in these trajectories (Narvacan et al., 2017; Wierenga et al., 2014; Wierenga, Bos, et al., 2018). Importantly, few studies tested for sex-differences in trajectories (Herting et al., 2018; Wierenga, Bos, et al., 2018) but only descriptively compared shapes of trajectories. In addition, the ability to model complex trajectories was limited by the fact that most studies had very few participants with more than two waves (see Table S1). Moreover, some studies included only a few subcortical regions (Goddings et al., 2014; Raznahan et al., 2014), which limits inferences about the trajectories of all subcortical structures.

1.2 | Methodological challenges

In summary, inconsistent results regarding subcortical trajectories and their potential sex differences exist. This might be related to differences in sample characteristics, study design, and analysis regarding sampling type (mostly accelerated longitudinal designs), age and sex, sociodemography, inter-scan interval, number of total scans, image acquisition, quality control, segmentation software, and analysis models. Mostly, parametric linear mixed effects models or hierarchical regression were used instead of the more flexible generalized additive mixed models (GAMMs), which describe the best relationship between predictor and outcome variables of interest without a priori knowledge of the inherent form of the data. Therefore, GAMMs allow for different nonlinear shapes of age trajectories for categories of a variable such as sex. For more details on methodical differences, please see supplemental material S1.

These methodical variations continue to restrict our ability to draw conclusions about subcortical structural trajectories in mid-adolescence to young adulthood, primarily due to limited studies with more than two waves and a sufficient amount of participants in this age span. Hence, investigating sex differences in subcortical trajectories during adolescence remains crucial. Our study aimed to examine these trajectories, covering the thalamus, globus pallidus, caudate nucleus, putamen, nucleus accumbens, hippocampus, and amygdala as regions of interest (ROIs) and their sex differences using a consistent quality control procedure, segmentation software, that is, FreeSurfer 6.0.0, and a GAMM approach. Contrary to most previous accelerated studies, our design was a single cohort approach across two sites (Dresden, Germany [$N = 248$]; Paris, France [$N = 255$]) where all participants started at the same age and more than half were followed across the entire age-range of interest. We included four waves of data acquisition at the ages of $M = 14.47$ years (range 13.33–15.72), $M = 16.48$ years (range 15.65–17.88), $M = 19.24$ years (range 17.26–22.54), and

$M = 22.35$ years (range 20.1–24.76) across sites, with a total of 1408 scans. In examining two similar sites, we can extend previous research that showed sex differences in trajectories dependent on specific sites (Herting et al., 2018) and thus explore age-constant and time-varying site differences closely focused on the mid-adolescent to young adult age span with more waves and scans within this particular age-range enabling more precise estimation of trajectories.

2 | MATERIALS AND METHODS

Access to the IMAGEN dataset is available with an accepted proposal from the IMAGEN executive committee (<https://imagen-europe.com/resources/imagen-project-proposal/>). All code is posted to Open Science Framework and can be assessed at <https://osf.io/5tfh4/>.

2.1 | Participants and study design

The study is part of the larger European Imaging Genetics project (IMAGEN) project assessing magnetic resonance imaging (MRI) data from typically developing adolescents at the age of 14, 18, and 22. For a detailed description of recruitment and assessment procedures in the IMAGEN study, please refer to Schumann et al. (2010). At the sites Dresden and Paris MRI data was assessed additionally at age 16, resulting in four waves (see Table 1 for sample characteristics and Figure 1 for age and sex distributions per site), therefore these two sites were chosen. Participants provided written informed consent and assent (in case of participants <18 years also their legal guardians). The study had been approved by the local ethics committees (Technische Universität Dresden and University of Paris) and was performed in accordance with the Declaration of Helsinki. Exclusion criteria included existing bipolar disorder, schizophrenia, and major neuro-developmental disorders such as autism, as well as a premature birth, head trauma, and history of several neurological or medical disorders. In this manuscript, female and male is defined as “assigned female at birth” and “assigned male at birth,” respectively. For Dresden, 260 participants took part in the overall IMAGEN study of whom 256 were scanned for at least one wave and 248 participants were included in the analyses after quality control. For Paris, 274 participants took part in the study of whom 258 were scanned for at least one wave and 255 participants were included in the analyses after quality control (see below and Table S2 for details). In total, the present study included 503 participants (250 females) and 1408 scans covering the age range of 13.33–24.76 years, including 447 scans at wave 1 (Dresden: 201, Paris: 246), 336 scans at wave 2 (Dresden: 207, Paris: 129), 372 scans at wave 3 (Dresden: 172, Paris: 200), and 253 scans at wave 4 (Dresden: 130, Paris: 123). Distribution of females to males did not significantly differ across waves (wave 1: 226/221, wave 2: 169/167, wave 3: 184/188, wave 4: 139/114) which was confirmed by a chi-square test, $\chi^2(3) = 2.03$, $p = .567$. Of the 503 participants included in the study 94% were White (97% Dresden, 90% Paris) and the majority stemmed from rather well-educated households (as a proxy for socioeconomic status), with

TABLE 1 Sample characteristics. Sample characteristics including demographics of typically developing adolescents from two sites of the IMAGEN study using single cohort designs.

	Dresden, Germany			Paris, France			Group comparison (Dresden all/Paris all)		
	All	Female	Male	All	Female	Male	t	df	p
Total scans across all waves	710	352	358	698	366	332			
Age (years)									
Wave 1, M (SD)	14.54 (.32)	14.58 (.32)	14.49 (.33)	14.42 (.5)	14.43 (.5)	14.41 (.5)			
Wave 2, M (SD)	16.6 (.42)	16.64 (.4)	16.56 (.42)	16.80 (.62)	16.84 (.59)	16.76 (.64)			
Wave 3, M (SD)	18.68 (.57)	18.72 (.56)	18.64 (.58)	19.73 (.73)	19.76 (.74)	19.69 (.73)			
Wave 4, M (SD)	22.1 (.74)	22.04 (.62)	22.12 (.85)	22.64 (.55)	22.62 (.52)	22.67 (.59)			
N									
Total	248	121	127	255	129	126			
Participants with 1 scan	35	14	21	41	17	24			
Participants with 2 scans	47	26	21	51	25	28			
Participants with 3 scans	83	38	45	95	49	44			
Participants with 4 scans	83	43	40	68	38	30			
Demographics—Data from the first wave is reported^a									
Pubertal status, ^b M (SD)	3.65 (.66)	4.03 (.41)	3.28 (.64)	3.55 (.78)	3.98 (.42)	3.1 (.82)	1.45	498	.148
Height in cm, M (SD)	169.42 (7.21)	166.55 (6.45)	172.13 (6.85)	165.57 (8.03)	162.67 (5.54)	168.6 (9.07)	5.6	492	***
Weight in kg, M (SD)	59.17 (11.74)	56.82 (10.23)	61.39 (12.64)	55.3 (9.26)	53.61 (7.9)	57.05 (10.25)	4.08	492	***
BMI, M (SD)	20.53 (3.29)	20.42 (3.03)	20.62 (3.51)	20.11 (2.61)	20.23 (2.6)	19.98 (2.65)	1.56	492	.119
IQ, ^c M (SD)	112.95 (10.72)	113.05 (9.93)	112.86 (11.48)	109.35 (9.43)	108.5 (8.95)	110.25 (9.87)	3.94	487	***

^aNote that data for pubertal status, height, weight, BMI, and IQ is reported for all participant in the final data set independent of a high-quality scan at the first wave. BMI, Body mass index (BMI = weight in kg/height in m²).

^bPubertal status ranged from 1 for prepubertal to 5 for postpubertal status, measured with the Pubertal Development Scale (PDS; Petersen et al., 1988).

^cGeneral cognitive ability estimated with the subtests Similarities, Block Design, Vocabulary, and Matrices from the Wechsler Intelligence Scale For Children (WISC-IV, German adaptation; Petermann & Petermann, 2007).

Note: * $p < .05$; ** $p < .01$; *** $p < .001$.

around 60% of the parents having obtained a university or college (university of applied sciences) degree. For more information about the parental education assessment in the IMAGEN study see tab. 1 in Schumann et al. (2010).

2.2 | Structural imaging

2.2.1 | Image acquisition

We acquired T1-weighted anatomical scans using the respective local Siemens (Erlangen, Germany) 3 T whole-body MR tomograph in Dresden and Paris. For both sites a standardized quality control assessment for the scanner was regularly performed (phantom scan

and in-vivo scan, see Schumann et al. (2010)). High-resolution T1-weighted images were collected using a magnetization prepared rapid acquisition gradient-echo (MPRAGE) sequence across all waves (Dresden: repetition time [TR] = 1900 ms, echo time [TE] = 2.26 ms, inversion time [TI] = 900 ms, voxel size = 0.5 × 0.5 × 1.0 mm, flip angle = 9°; matrix size = 256 × 256 mm; 176/slices; [Paris: TR = 2300 ms, TE = 2.93 ms, TI = 900 ms, voxel size = 1.1 × 1.1 × 1.1 mm, flip angle = 9°; matrix size = 256 × 256 mm; 160 slices]). The sequence in Dresden varied slightly with a higher resolution. This is due to differences in imaging data assessment waves at the each site of the IMAGEN project. Only the sites Dresden and Paris assessed imaging data at age 16 with Dresden using the local MPRAGE sequence only. All images were examined by a clinical neuroradiologist for structural abnormalities.

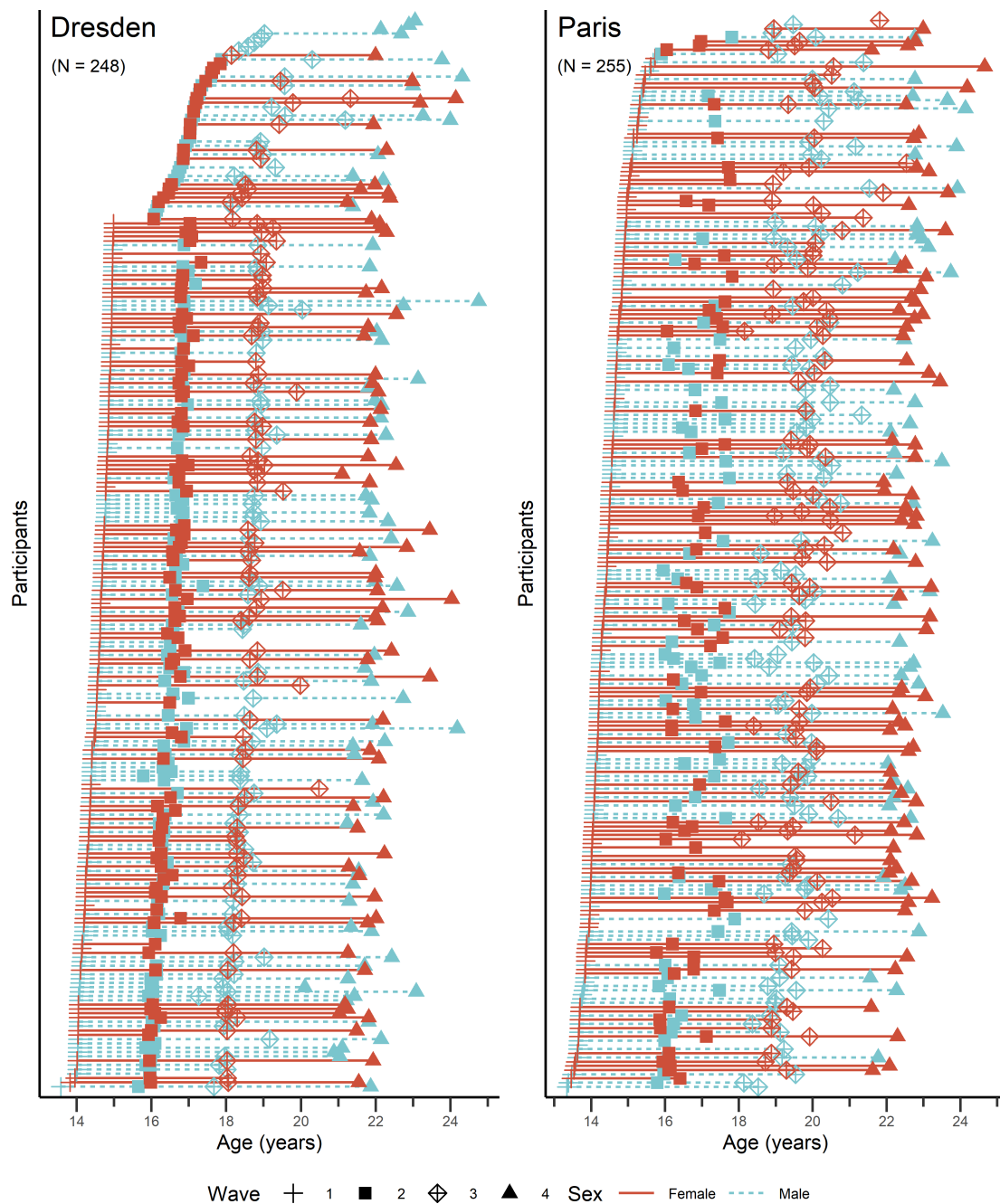


FIGURE 1 Age and sex distributions for each site.

2.2.2 | Image processing

Images (also single time points) were processed using the longitudinal pipeline of FreeSurfer 6.0.0 (<http://surfer.nmr.mgh.harvard.edu>; Fischl et al., 2002; Reuter et al., 2012). This pipeline creates an unbiased within-subject template space and image using robust, inverse consistent registration (Reuter et al., 2010). Several processing steps, such as skull stripping, Talairach transforms, atlas registration, as well as spherical surface maps and parcellations are then initialized with common information from the within-subject template. This process

significantly improves robustness with respect to noise, intensity scaling and outliers, significantly increasing statistical power (Reuter et al., 2012) and scan-rescan reliability as compared to conventional cross-sectional analysis pipelines (Jovicich et al., 2013). We further used the recon-all-flag-3 T which alters FreeSurfer's internal N3 bias field correction parameters (Sled et al., 1998) making it more appropriate for 3 T MRI (Zheng et al., 2009).

As previous studies demonstrated no evidence of developmental difference between hemispheres in our subcortical ROIs (Brain Development Cooperative Group, 2012; Østby et al., 2009), we report

averaged volumes over both hemispheres similarly to previous studies (Herting et al., 2018; Wierenga, Bos, et al., 2018). Correlations between hemispheres across sites at each wave were high for each ROI (r 's > .7) except for the nucleus accumbens (r 's = .46 to .79) and the amygdala (r 's = .6 to .79; see Table S3 for all correlations). Following the methodology of previous studies (e.g., Herting et al., 2018; Wierenga, Bos, et al., 2018) we used absolute volume values and did not correct ROI volumes for intra-cranial volume.

2.2.3 | Quality control

Before statistical analyses, trained operators performed the following multi-step quality control procedure: For pre-processing quality control all raw images were visually inspected for motion and technical artifacts according to our lab's three-category (fail, check, pass) rating scheme checking for image sharpness, blurred regions, susceptibility, and ringing artifacts and whether all subcortical structures as well as gray and white matter could be well-differentiated (for more details, see Backhausen et al., 2016). We retained "check" and "pass" images and excluded "fail" images from longitudinal pipeline processing to ensure high quality of the within-subject template. In post-processing quality control trained operators inspected longitudinally processed images with questionable raw image quality (rated "check") for accuracy of subcortical segmentation. Subcortical structures were excluded when incorrectly segmented (see Table S2). No manual adjustments of subcortical volumes were made.

2.2.4 | Statistical analyses

For analysis of GAMMs we used R version 3.6.1 with the mgcv package version 1.8-33 (Wood, 2004, 2006) and the packages itsadug version 2.3 and ggplot2 version 3.3.2 for visualization of smooth curves. We used a shrinkage version of penalized cubic regression splines as smooth terms in all models. The k parameters, which specify the number of basis functions to build the curves and thus influence potential over-fitting, were set to 4 after k.check (mgcv package) analyses. We closely followed the approach of Herting et al. (2018) and built GAMM models to examine age-constant and time-varying sex and site differences in trajectories of subcortical volumes. We explored differences in trajectories by examining significance of difference smooths between the trajectories of males versus females in the Dresden and Paris site. GAMM models were implemented to (1) test an age-constant sex difference (main effect of sex) as well as a sex difference in the age trajectory (age*sex interaction), while controlling for site at the level of main effect and site*age interaction and (2) test an age-constant site difference (main effect of site) as well as a site difference in the age trajectory (age*site interaction), while controlling for sex at the level of main effect and sex*age interaction. All models included an individual-level random effect intercept per participant. Sex and site were coded as factors (male = 0, female = 1; Dresden = 0, Paris = 1). To directly test sex and site differences, previous GAMM

TABLE 2 GAMMS examining sex and age effects across both sites on subcortical regions of interest.

Thalamus				
Main effect	β	SE	t	p
Sex (F vs. M)	-734.91	53.75	-13.67	***
Trajectory				
	edf	Ref.df	F	p
S (age)	2.85	3	24.44	***
S (age, F vs. M)	2.29	3	11.44	***
Globus pallidus				
Main effect	β	SE	t	p
Sex (F vs. M)	-184.97	17.39	-10.64	***
Trajectory				
	edf	Ref.df	F	p
S (age)	0.81	3	0.45	**
S (age, F vs. M)	2.46	3	56.84	***
Caudate nucleus				
Main effect	β	SE	t	p
Sex (F vs. M)	-302.60	40.35	-7.50	***
Trajectory				
	edf	Ref.df	F	p
S (age)	1.88	3	1.78	***
S (age, F vs. M)	2.62	3	40.02	***
Putamen				
Main effect	β	SE	t	p
Sex (F vs. M)	-595.96	45.89	-12.99	***
Trajectory				
	edf	Ref.df	F	p
S (age)	1.76	3	1.53	***
S (age, F vs. M)	2.53	3	20.27	***
Nucleus accumbens				
Main effect	β	SE	t	p
Sex (F vs. M)	-63.17	9.0	-7.02	***
Trajectory				
	edf	Ref.df	F	p
S (age)	0.00	3	0.00	0.62
S (age, F vs. M)	1.80	3	11.52	***
Hippocampus				
Main effect	β	SE	t	p
Sex (F vs. M)	-360.67	30.39	-11.87	***
Trajectory				
	edf	Ref.df	F	p
S (age)	2.31	3	1.88	0.55
S (age, F vs. M)	2.33	3	19.17	***
Amygdala				
Main effect	β	SE	t	p
Sex (F vs. M)	-238.48	15.28	-15.61	***
Trajectory				
	edf	Ref.df	F	p
S(age)	0.00	3	0.00	0.76
S(age, F vs. M)	2.27	3	23.22	***

Note: GAMM estimates for age, sex, and age*sex across both sites. Sex was coded as a factor (male = 0, female = 1), allowing for each term to reflect the following: Sex (F vs. M) reflected the main difference in females as compared to males; S(age) reflected the global smooth trajectory of age; S(age, F vs. M) reflected the difference in trajectory of females compared males. Site and site*age were included as covariates in all models. Estimate, standard error (SE), t -value, and associated p -values for each main difference. Smooth function (S(age)/edf), degrees of freedom (Ref.df), F -statistic and associated p -values for each smooth term. * p < .05; ** p < .01; *** p < .001.

models were updated in coding sex/site and sex*age/site*age as contrasting ordered factors. For GAMM estimates of these smooth terms see Table 2.

To test whether the inclusion of age*sex interaction terms significantly improve model fit and explain variance, we analyzed three models following procedures by Pedersen et al. (2019), Wierenga, Bos, et al. (2018), and Wierenga et al. (2019) for each subcortical structure as a smooth function of age of the individual i at wave j , a random intercept per participant effect (u_i) and error ($error_{ij}$), while controlling for site (in main effect and site*age interaction) in each model. We compared these models using Akaike Information Criterion (AIC; Akaike, 1973) and Bayesian Information Criterion (BIC; Schwarz, 1978). The model that had the lowest AIC and BIC scores, and significantly differed from more parsimonious models was selected as the best fit model.

Model 1: age only model including only a global smooth term for age

$$y_{ij} = \beta_0 + \beta_1(\text{site}_i) + s_1(\text{age}_{ij}) + s_2(\text{age}_{ij})\text{site}_j + u_i + \text{error}_{ij}.$$

Model 2: age plus main effect of sex model including a global smooth term for age and a main effect of sex

$$y_{ij} = \beta_0 + \beta_1(\text{site}_i) + \beta_2(\text{sex}_i) + s_1(\text{age}_{ij}) + s_2(\text{age}_{ij})\text{site}_j + u_i + \text{error}_{ij}.$$

Model 3: age plus sex*age interaction including a global smooth term for age, a main effect of sex, and group-level smooth terms for males and females with differing complexity

$$y_{ij} = \beta_0 + \beta_1(\text{sample}_i) + \beta_2(\text{sex}_i) + s_1(\text{age}_{ij}) + s_2(\text{age}_{ij})\text{site}_j + s_3(\text{age}_{ij})\text{sex}_j + u_i + \text{error}_{ij}.$$

Here, s_1 , s_2 , and s_3 denote the arbitrary smooth functions (shrinkage version of cubic regression splines). Further, β_0 stands for the random intercepts per participant, β_1 for parameter estimates of site, and β_2 for parameter estimates of sex.

3 | RESULTS

For all seven subcortical structures, models with the interaction age*sex (Model 3) had lower AIC and BIC values than the more parsimonious models, and were thus deemed best fitting models (Table S4). Moreover, for all seven ROIs we found significant sex differences in age trajectories when examining difference smooths between females and males. Hence, we present results for GAMM models including age, sex, as well as an age*sex interaction to investigate age-constant sex differences, as well as a sex difference in age trajectories (age*sex interaction), while controlling for site (main effect and age*site interaction). Notably, trajectories for each sex follow the global age trajectory (i.e., across all participants of both sexes) as they are penalized by the model to not differ too much from it (Pedersen et al., 2019). To better understand these differences, we examined age trajectories in GAMMs for each sex separately (independent from each other and the global age trajectory) while controlling for site (see Table 3 and Figure 2).

TABLE 3 GAMM estimates for age across both sites in females and males separately.

Females	edf	Ref.df	F	p	Males				
					edf	Ref.df	F	p	
Thalamus									
S (age)	2.77	3	54.82	***	S (age)	2.58	3	5.18	***
Globus pallidus									
S (age)	2.06	3	58.88	***	S (age)	2.70	3	80.86	***
Caudate nucleus									
S (age)	2.70	3	197.81	***	S (age)	2.75	3	177.71	***
Putamen									
S (age)	2.58	3	127.74	***	S (age)	2.68	3	61.43	***
Nucleus accumbens									
S (age)	1.75	3	10.61	***	S (age)	0.00	3	0.00	0.78
Hippocampus									
S (age)	2.45	3	4.22	**	S (age)	1.76	3	1.15	***
Amygdala									
S (age)	0.00	3	0.00	0.74	S (age)	2.70	3	18.82	***

Note: Smooth function (S (age)/edf), degrees of freedom (Ref.df), F-statistic, and associated p-values for each term. * $p < .05$; ** $p < .01$; *** $p < .001$.

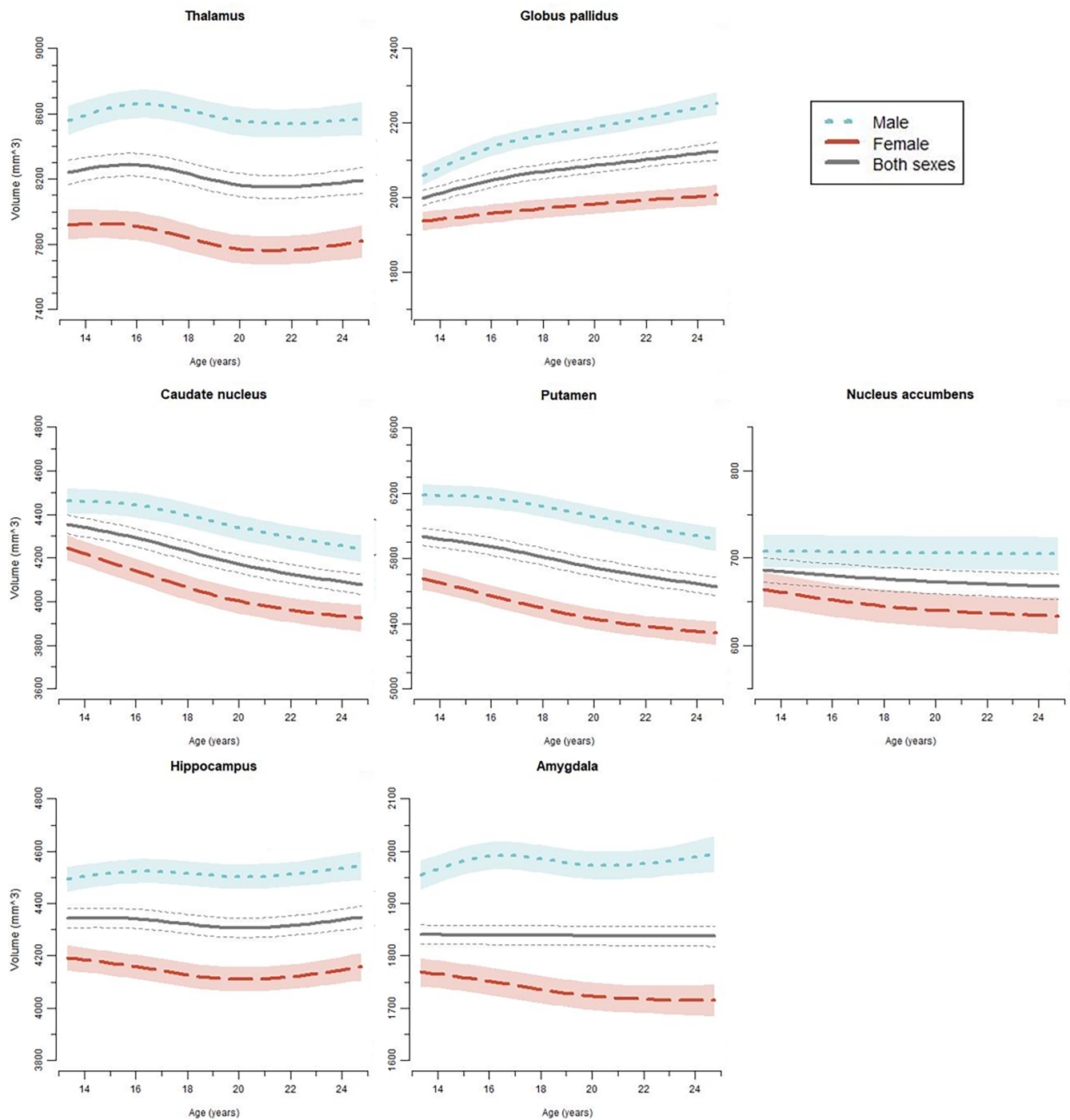


FIGURE 2 Age trajectories for subcortical ROIs across the two sites Dresden and Paris. GAMM fitting with 95% confidence intervals plotted separately for females, males, and both sexes across Dresden and Paris.

3.1 | Overall age trajectories; that is, not differentiating sex and site

Estimation of smooth terms including all participants indicated overall volume decrease in caudate nucleus and putamen and overall volume increase in globus pallidus. Further, we found a nonlinear pattern of change in the thalamus and no significant volume change in nucleus accumbens, hippocampus, and amygdala.

3.2 | Sex differences

3.2.1 | Sex differences in raw volumes

To investigate volume differences between the sexes for all data sets and in each wave separately, we calculated percentage differences between males and females from raw values, not from GAMM modeling. For all structures, males had larger volumes as

TABLE 4 Percentage volume differences of subcortical brain structures between sexes, ordered by largest to smallest in the overall difference.

Structure	Difference in volume (male - female %)									
	All (N = xx)		TP1 (N = xx)		TP2 (N = xx)		TP3 (N = xx)		TP4 (N = xx)	
	M	CI	M	CI	M	CI	M	CI	M	CI
Amygdala	13.66	11.86–15.45	12.31	9.26–15.35	12.38	8.85–15.9	15.41	11.74–19.07	15.16	10.74–19.58
Putamen	10.81	9.18–12.43	9.43	6.72–12.	10.61	7.3–13.91	12.48	9.12–15.83	10.66	6.58–14.47
Thalamus	9.83	8.27–11.39	8.62	6.01–11.23	9.70	6.53–12.86	10.17	7.1–13.24	11.63	7.68–15.58
Pallidum	9.42	7.89–10.95	7.24	4.83–9.64	8.98	5.91–12.05	11.04	7.86–14.22	11.75	7.78–15.73
Nucleus accumbens	9.22	7.71–10.73	7.9	5.46–10.48	8.77	5.75–11.79	11.90	8.61–15.19	7.88	4.56–11.2
Hippocampus	8.75	7.27–10.23	7.70	5.22–10.19	8.20	5.25–11.15	9.59	6.59–12.59	10.22	6.49–13.95
Caudate nucleus	7.38	6.02–8.75	5.56	3.43–7.68	7.46	4.65–10.28	8.73	5.86–11.6	8.17	4.79–11.54

Note: For *n* of each structure per wave please refer to Table S2. CI = lower and upper bound of normal approximation to the binomial calculation with 95% confidence level.

compared to females across the entire age range with overall differences ranging from 8.7% for the hippocampus to 13.7% for the amygdala (see Table 4). Percentage differences increased with age in thalamus, globus pallidus, caudate nucleus, hippocampus, and amygdala, which is partly mirrored by increasingly diverging trajectories of females and males over time (see Table 4; Figure 2). All data points for the subcortical ROIs are shown in Figure S1.

3.2.2 | Sex differences in age trajectories

Thalamus

Males showed a volume increase until about age 17, while females remained rather stable until about age 16. After this, both sexes followed a similar nonlinear trajectory starting with a decrease and then a slight increase.

Globus pallidus

While females showed a nonlinear increase across the entire age range, the curve for males increased steeper until age 16 and flattened afterwards.

Caudate nucleus

For females, volume decreased until about age 22, then decrease flattened. Males showed a flatter decrease until about age 17, followed by a nonlinear decrease.

Putamen

For females, volume decreased until about age 20, then decrease flattened. Males showed a flatter decrease until about age 16, followed by a nonlinear decrease.

Nucleus accumbens

We did not find volume change in males, but a significant volume decrease in females which differed from the global age trajectory.

Hippocampus

While both groups showed significant change over time we did not find obvious decrease or increase over time. Males and females differed significantly from each other showing nonlinear patterns with females showing a volume decrease until age 20, then increasing again to about their initial volume level. Males increased slightly until age 17, followed by a slight increase.

Amygdala

We found no volume change in females but a volume increase in males until about age 16 in males with a following non-linear trajectory. Males also differed significantly from the global trajectory while females did not.

When examining sex differences in each site separately, significant differences between males' and females' trajectories were seen for all subcortical structures in Paris and for all subcortical structures except the nucleus accumbens in Dresden (see Figure 3).

3.3 | Site differences

3.3.1 | Age-constant site differences in volumes

Overall, we found significantly smaller volumes in the Paris site for thalamus, globus pallidus, and hippocampus. Significantly larger volumes in the Paris site were seen for putamen, nucleus accumbens, and amygdala, while the caudate nucleus volume did not differ between the sites.

3.3.2 | Site differences in age trajectories

We followed a similar logic for examining the site differences as for the sex differences. For all subcortical structures we found significant differences in age trajectories between the Dresden and Paris site. Examining the sites separately revealed significant volume change in thalamus, globus pallidus, caudate nucleus, and putamen for both

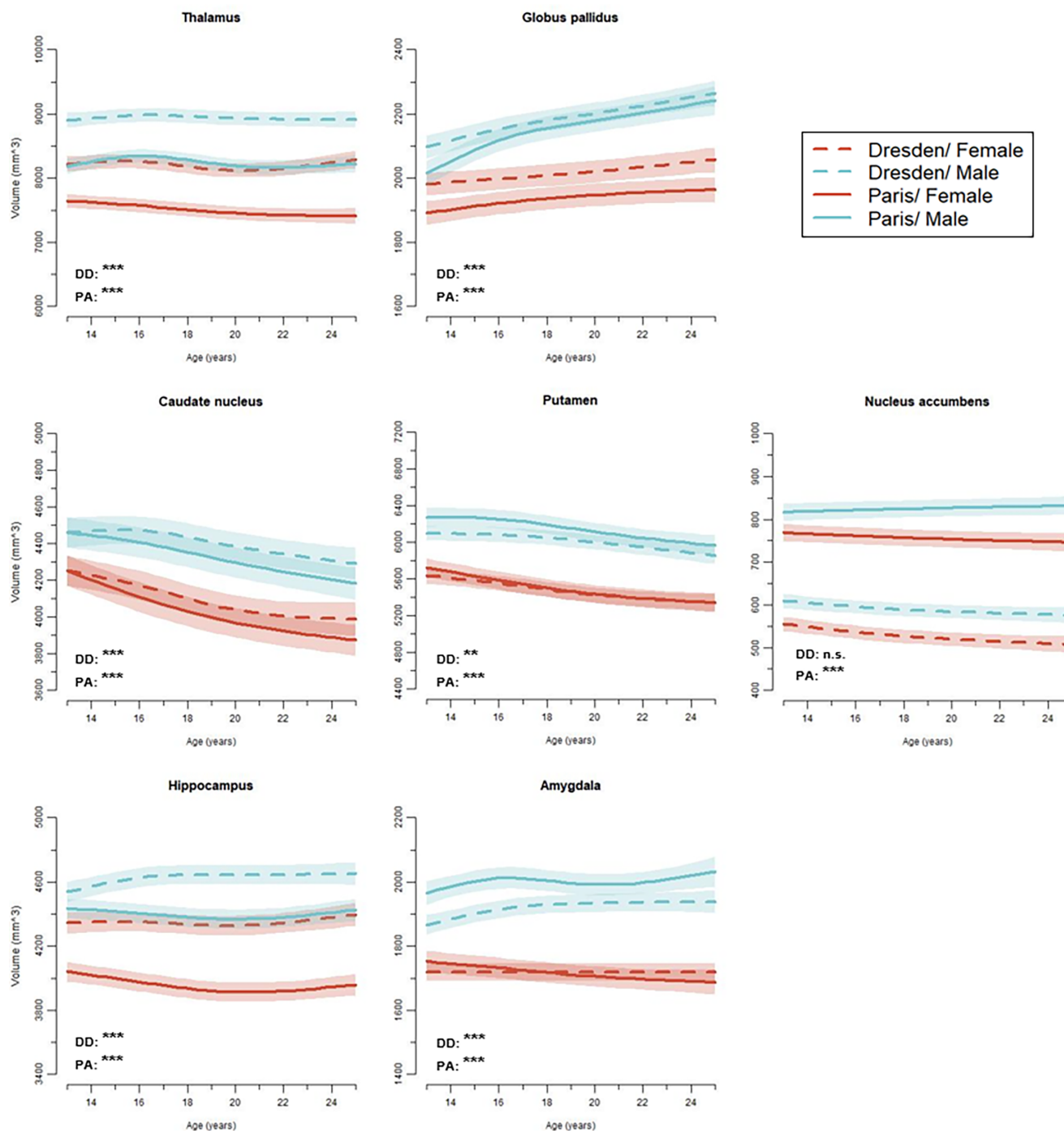


FIGURE 3 Site differences of developmental trajectories for the subcortical structures. GAMM fitting with 95% confidence intervals plotted separately for females and males in Dresden and Paris. p -Values represent sex differences per site. DD, Dresden; PA, Paris; n.s., not significant; ** $p < .01$; *** $p < .001$.

sites, significant change only for Dresden in the nucleus accumbens, and no change in both sites for the amygdala.

4 | DISCUSSION

This study investigated age trajectories and their sex differences across ages 14 to 24 of the volumes of seven subcortical structures in

two large samples from Germany and France as part of the IMAGEN project including four waves of data acquisition. This is the first study with a single cohort design to focus on the mid-adolescent to young adult age span providing a continuous data set with 1408 total scans in 503 participants with 36% of participants having three, and 30% of participants having four waves of data acquisition. Our data indicate larger volumes in males compared to females and different age trajectories between the sexes across all subcortical regions. Although the

current sample was partly included in the meta-analysis of Dima et al. (2021) we here present all four waves from age 14 to 24 processed in the longitudinal stream with FreeSurfer 6.0 (available only for the Dresden and Paris site) versus only the first wave in Dima et al. (2021) was processed using the cross-sectional stream with FreeSurfer 5.3 (available for all IMAGEN sites).

4.1 | Overall age trajectories across both sexes

We found a decrease in the thalamus from age 16, which was in the range of other studies (15–19 years; Herting et al., 2018; Raznahan et al., 2014; Wierenga, Bos, et al., 2018). Our data show a stabilization thereafter. The slight increase from about age 21 is unexpected and might stem from the few scans at later ages (40 scans after age 23). We found an overall volume increase in the globus pallidus. This is in line with four other studies including a multi-site study (Dennison et al., 2013; Herting et al., 2018; Wierenga et al., 2014; Wierenga, Bos, et al., 2018). Wierenga et al. (2014) had very few scans after age 17, presumably leading to a peak with a following decrease. In contrast, Raznahan et al. (2014), found a decrease. Overall, we found volume decreases in putamen and caudate. This is in line with cross-sectional studies (Brain Development Cooperative Group, 2012; Østby et al., 2009) and a recent mega-analysis (Dima et al., 2021) as well as longitudinal studies (Dennison et al., 2013; Goddings et al., 2014; Tamnes et al., 2013; Wierenga et al., 2014). In contrast, other findings suggest no volume change for the caudate nucleus (Narvacan et al., 2017; Wierenga, Bos, et al., 2018) or an increase for the putamen (Wierenga, Bos, et al., 2018). We demonstrated no volume change in the nucleus accumbens. This is in line with (Dennison et al., 2013), while Wierenga, Bos, et al. (2018) found a decrease. Further, we found a decrease in the hippocampus followed by a slight increase from about age 20. The decreasing pattern is in line with Tamnes et al. (2013) but others found an increase (Dennison et al., 2013; Herting et al., 2018). We demonstrated no volume change in the amygdala which is in line with some previous studies (Dennison et al., 2013; Wierenga, Bos, et al., 2018) but divergent to others (Goddings et al., 2014; Østby et al., 2009; Wierenga et al., 2014). However, findings seem to be driven by sex differences (see e.g., Goddings et al., 2014 and next section on sex differences).

4.2 | Sex differences

4.2.1 | Sex differences in raw volumes

We found larger volumes across subcortical structures for males versus females without allometric scaling. This is in line with previous studies (Herting et al., 2018; Wierenga et al., 2014; Wierenga, Bos, et al., 2018; Wierenga, Sexton, et al., 2018) and the volume differences of 7% to 14% are in a similar range as found by Narvacan et al. (2017).

4.2.2 | Sex differences in age trajectories

While many previous studies only had few data sets per age group our single cohort study should have the computational strength to precisely model age trajectories during adolescence and detect sex differences. We found significant sex differences in age trajectories for all subcortical ROIs across the two sites. This is in line with Herting et al. (2018) and Raznahan et al. (2014) who also found sex differences in age trajectories for all their subcortical ROIs (Raznahan et al., 2014: striatum, globus pallidus, thalamus). We also found sex differences for each site separately, except for the nucleus accumbens in Dresden. This differs from Herting et al. (2018), who did not find sex differences for all regions when examining their three sites separately. The consistent sex differences for each site might be due to the more homogenous samples and more time points, that is, power in our study. No sex differences in age trajectories were found in three studies (Narvacan et al., 2017; Wierenga et al., 2014; Wierenga, Sexton, et al., 2018). This is most likely an issue of power (studies covering about 1/2, 1/6, or 1/10 of our total scans) or larger age-ranges (from about primary school until mid-to late twenties). Interestingly, Dennison et al. (2013) had similarly low power (around 1/11 of our scans) in a longitudinal design with participants in the adolescent age range and still found sex differences in three subcortical regions. Only Herting et al. (2018) and Raznahan et al. (2014) had similarly powered samples but covered larger age-ranges. Single ROIs will be discussed below.

Overall, descriptively, we observed volume changes for males across several structures until age 16/ 17 while females showed different trajectories in this time span (i.e., stabilization/decrease). Presumably, females reached the peaks and turning points earlier. While the thalamus has been shown to change for females or males only (Dennison et al., 2013; Herting et al., 2018), we found a significant change in both sexes. In line with Dennison et al. (2013) and Herting et al. (2018) both sexes first decreased from about age 16 (females) or 17 (males) on. Differing from our study, Dennison et al. (2013) did not cover the age beyond 17. Raznahan et al. (2014) found a rather stable trajectory in males with a slight decrease in females after about age 14. In contrast to our study, Raznahan et al. (2014) had an accelerated design and assessed fewer participants in adolescence while covering earlier ages. Our slight increase from about age 21 is not in line with previous studies and might stem from the few scans after age 23. For the globus pallidus, females showed a nonlinear increase across the entire age range, while the curve for males increased steeper until age 16 and flattened afterwards. Herting et al. (2018) only found a descriptive albeit non-significant increase for each sex. In contrast, Raznahan et al. (2014) demonstrated a similar slight decrease between the sexes across the adolescent age-range, while Dennison et al. (2013) found an increase in the globus pallidus and no sex differences. Taken together, sex differences in volume age trajectories for the globus pallidus were not convergingly found in the previous literature (also including Narvacan et al., 2017; Wierenga et al., 2014; Wierenga, Sexton, et al., 2018) and our trajectories show an increase in both sexes with just a steeper increase in males. Since caudate

nucleus and putamen form the striatum and show similar trajectories in our data, we discuss them together. We found a decrease for both structures: For females, volume decreased until about age 22 for the caudate nucleus, respectively age 20 for the putamen 20, then decrease flattened. Males showed a flatter decrease until about age 17 (putamen 16), followed by a nonlinear decrease. This is in line with Dennison et al. (2013) who found a decrease for both sexes in both structures. No decrease was shown by Herting et al. (2018) for males, but a significant decrease for females in both structures. Raznahan et al. (2014) investigated both structures together as “the striatum” and found an earlier peak for females at age 12 versus 14 for males with a slight decrease thereafter.

In line with Herting et al. (2018) we found a stable trajectory for the nucleus accumbens in males and a decrease in females while other studies (Dennison et al., 2013; Narvacan et al., 2017; Wierenga et al., 2014; Wierenga, Sexton, et al., 2018) did not find sex differences for this region. Taken together, our study together with Herting et al. (2018) is the only one so far to indicate sex differences in the nucleus accumbens. We discuss amygdala and hippocampus trajectories together since they have similar trajectories. We can only descriptively compare our data with Herting et al. (2018) since they investigated also younger age groups from 8 years on and therefore sex differences in trajectories from age 14 to 24 were not statistically tested. We found no volume change in females but a volume increase in males until about age 16 with a following non-linear trajectory. Males also differed significantly from the global trajectory while females did not. For the amygdala, in line with Herting et al. (2018) who reported stable volumes from age 14 on for females and a slight increase in males, we did not find significant change in females but a nonlinear increase in males. Like Herting et al. (2018), we found significant nonlinear change in the hippocampus for both males and females while our data indicated slight increase from age 20 which was not shown by Herting et al. (2018). Other studies (Dennison et al., 2013; Narvacan et al., 2017; Wierenga et al., 2014; Wierenga, Bos, et al., 2018) did not find sex differences for these regions.

Notably, a recent study investigated sex and puberty effects on structural maturation subcortical structures in a subsample (scans at age 14 and 16) from Dresden and Paris using a voxel-based morphometry and ROI approach (Bézivin-Frere et al., 2020). The authors reported significant differences in volume trajectories with age and puberty in the amygdala-hippocampal complex, caudate, putamen, and thalamus. Together with another study pointing to independent and interactive influences of puberty and age on subcortical development (Goddings et al., 2014) the authors call attention to the importance of pubertal markers when looking at brain development during adolescence, as these processes are likely influenced by differences in sex hormones between females and males. As we only had information on pubertal status for waves one and two and assume pubertal maturation complete afterwards, we did not include pubertal effects in our analyses. Still, differences in subcortical brain development with age and puberty, next to cognitive and social-emotional differences, may add insight into the development of sexual dimorphism in mental

disorders during adolescence and later in life (Zahn-Waxler et al., 2008). One mechanism might be the influence of amygdala and hippocampus volume change which differently related to the development of self-reported positive characteristics (e.g., how generous, affectionate, and caring a person is) during adolescence between females and males (Bézivin-Frere et al., 2020).

Reviewing results across all subcortical structures, there is still need for further studies. So far, our study is only the seventh study with three studies spanning only low numbers of 60–147 participants and including only two waves or few participants with three time points (Dennison et al., 2013; Narvacan et al., 2017; Wierenga et al., 2014). All except one study had accelerated designs, so there is a need for larger studies with single cohort designs.

4.3 | Site effects

Similar to Herting et al. (2018) who also used consistent quality control, longitudinal preprocessing and statistical analyses, we observed site differences and interactions of site by sex depending on the subcortical ROIs. Overall differences in volume between Dresden and Paris were especially visible in thalamus, hippocampus, and nucleus accumbens, and Herting et al. (2018) point to site differences in the thalamus, globus pallidus, caudate, and hippocampus. Interestingly, while Herting et al. (2018) analyzed more heterogeneous sites (differences in number of total scans, age spans, and sex distributions), our two European samples were vastly comparable on sample characteristics such as number of total scans, inter-scan interval, age range, sex distribution, and pubertal and socioeconomic status. It is hence unlikely that differences in sample characteristics lead to overall volume differences in Dresden versus Paris. These systematic effects may rather stem from differences in the imaging sequence used on the two sites (higher resolution in Dresden) which may have locally affected the FreeSurfer segmentation algorithm. Further, a scanner update was completed during wave three in Paris (Siemens Trio to Prisma) which may have introduced additional bias in results as Medawar et al. (2021) and Plitman et al. (2021) reported very good ICC but also percent volume differences in several cortical and subcortical regions when comparing Siemens Scanners (Trio versus Prisma/Verio versus Skyra). Since many recent longitudinal studies on brain development included several samples with differences in imaging sequence and scanner type (Dennison et al., 2013; Mills et al., 2021; Tamnes et al., 2017; Wierenga, Sexton, et al., 2018), results should be interpreted with caution.

4.4 | Limitations

All findings of the present study must be considered in light of the following limitations. Some subcortical structures might suffer from low reliability, that is, the nucleus accumbens (Mills et al., 2021) putamen or globus pallidus (Wonderlick et al., 2009). The globus pallidus is less

distinct from its surrounding white matter as compared to other subcortical regions such as the thalamus or caudate nucleus (Fischl et al., 2002; Wonderlick et al., 2009). Further, the putamen seems to be systematically overestimated in Freesurfer 5.3, which has been improved in the Freesurfer 6.0 (see release notes). This becomes obvious when, for example, comparing putamen volumes at around age 14 in the samples of Herting et al. (2018) of around 6500 mm³ for female and 7000 mm³ for male with our values (5600 mm³ for female 6200 mm³ for male).

Concerning GAMM statistics, when global age smooth term is included additionally to group-level smooth term (as done here following Herting et al., 2018) the group-level smooth terms are penalized to not deviate too much from the global smooth terms. This inherently prevents female and male age trajectories from differing too much when fitted in these types of GAMM compared to when fitted separately. Further, concurvity (nonparametric equivalent of multicollinearity) is present when fitting models with separate global and group-level smooth terms as the global term could be approximated by a combination of female and male group-level smooth terms. This might lead to estimation issues in the model and inflated type 1 errors.

Moreover, current results are naturally dependent on power effects: when classes of sexes/sites are modeled separately the number of data sets is divided in half and significance of effects might be changed (see discussion in Herting et al., 2018). Results further depend on differences in site composition over time as in our study more females than males returned in later waves which might have affected estimation of sex trajectories. Our data also provide more power for the first two waves. Thus, we can be more confident in describing trajectories during mid to late adolescence (until about age 17) while conclusions become less reliable after this time due to attrition of the site (especially after age 23).

While the strength of some previous studies was to show the development in-and-out of adolescence, we focused on the mid-adolescent to young adult age span and provide 1408 total scans within this age span. However, we cannot draw any conclusions about a potential turning point of trajectories before age 14. Additionally, in line with most other studies, the young participants were ethnically homogeneous (i.e., mostly White participants) and stemmed from a well-educated background which restricts generalizability. Future studies could try to recruit participants with a more diverse ethnical and socioeconomic background.

5 | CONCLUSION

Our detailed descriptions of regional subcortical trajectories provide further evidence of sex differences in longitudinal adolescent brain development from mid-adolescence to young adulthood. Additionally, we illuminate pitfalls in large multi-site longitudinal brain imaging studies in an attempt to improve replicability of typical subcortical trajectories over the lifespan.

AUTHOR CONTRIBUTIONS

Tobias Banaschewski, Gareth J. Barker, Herta Flor, Hugh Garavan, Andreas Heinz, Gunter Schumann, Jean-Luc Martinot, and Michael N. Smolka acquired the funding and designed the study. Juliane H. Fröhner, Nora C. Vetter, and all the other members of the IMAGEN consortium participated in data collection and data management. Lea L. Backhausen did the analysis. Nora C. Vetter and Lea L. Backhausen drafted the manuscript. Nora C. Vetter, Megan M. Herting, and Fabio Sticca assisted with data analysis and interpretation of findings. All authors critically reviewed the content and approved the final version for publication.

AFFILIATIONS

¹Department of Psychiatry and Psychotherapy, TUD Dresden University of Technology, Dresden, Germany

²Department of Child and Adolescent Psychiatry, Medical Faculty and University Hospital Carl Gustav Carus, TUD Dresden University of Technology, Dresden, Germany

³NeuroSpin, CEA, Université Paris-Saclay, Gif-sur-Yvette, France

⁴Institut des Maladies Neurodégénératives, UMR 5293, CNRS, CEA, Université de Bordeaux, Bordeaux, France

⁵Institut National de la Santé et de la Recherche Médicale, INSERM U1299 "Trajectoires Développementales en Psychiatrie", Université Paris-Saclay, Ecole Normale supérieure Paris-Saclay, CNRS, Centre Borelli, Gif-sur-Yvette, France

⁶Department of Psychiatry, Lab-D-Psy, EPS Barthélémy Durand, Etampes, France

⁷AP-HP, Sorbonne Université, Department of Child and Adolescent Psychiatry, Pitié-Salpêtrière Hospital, Paris, France

⁸Departments of Population and Public Health Sciences and Pediatrics, University of Southern California, Los Angeles, California, USA

⁹Institute for Educational Support for Behaviour, Social-Emotional, and Psychomotor Development, University of Teacher Education in Special Needs, Zurich, Switzerland

¹⁰Department of Child and Adolescent Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

¹¹Department of Neuroimaging, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK

¹²Discipline of Psychiatry, School of Medicine and Trinity College Institute of Neuroscience, Trinity College Dublin, Dublin, Ireland

¹³Centre for Population Neuroscience and Precision Medicine (PONS), Institute of Psychiatry, Psychology & Neuroscience, SGDP Centre, King's College London, London, UK

¹⁴Institute of Cognitive and Clinical Neuroscience, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

¹⁵Department of Psychology, School of Social Sciences, University of Mannheim, Mannheim, Germany

¹⁶Departments of Psychiatry and Psychology, University of Vermont, Burlington, Vermont, USA

¹⁷Sir Peter Mansfield Imaging Centre School of Physics and Astronomy, University of Nottingham, University Park, Nottingham, UK

¹⁸Department of Psychiatry and Neurosciences, Charité—Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany

¹⁹Physikalisch-Technische Bundesanstalt (PTB), Braunschweig, Germany

²⁰Institute of Medical Psychology and Medical Sociology, University Medical Center Schleswig Holstein, Kiel University, Kiel, Germany

²¹Department of Child and Adolescent Psychiatry and Psychotherapy, University Medical Centre Göttingen, Göttingen, Germany

²²Department of Child and Adolescent Psychiatry, Psychotherapy and Psychosomatics, University Medical Center Hamburg Eppendorf, Hamburg, Germany

²³Department of Psychological Medicine, Section for Eating Disorders, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

²⁴Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany

²⁵School of Psychology and Global Brain Health Institute, Trinity College Dublin, Dublin, Ireland

²⁶PONS Research Group, Dept of Psychiatry and Psychotherapy, Campus Charite Mitte, Humboldt University, Berlin and Leibniz Institute for Neurobiology, Magdeburg, Germany

²⁷Institute for Science and Technology of Brain-Inspired Intelligence (ISTBI), Fudan University, Shanghai, China

²⁸Department of Psychology, MSB Medical School Berlin, Berlin, Germany

ACKNOWLEDGMENTS

This work received support from the following sources: the European Union-funded FP6 Integrated Project IMAGEN (Reinforcement-related behaviour in normal brain function and psychopathology) (LSHM-CT-2007-037286), the Horizon 2020 funded ERC Advanced Grant “STRATIFY” (Brain network based stratification of reinforcement-related disorders) (695313), Human Brain Project (HBP SGA 2, 785907, and HBP SGA 3, 945539), the Medical Research Council Grant “c-VEDA” (Consortium on Vulnerability to Externalizing Disorders and Addictions) (MR/N000390/1), the National Institute of Health (NIH) (R01DA049238, A decentralized macro and micro gene-by-environment interaction analysis of substance use behavior and its brain biomarkers), the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London, the Bundesministerium für Bildung und Forschung (BMBF grants 01GS08152; 01EV0711; Forschungsnetz AERIAL 01EE1406A, 01EE1406B; Forschungsnetz IMAC-Mind 01GL1745B), the Deutsche Forschungsgemeinschaft (DFG project numbers 186318919 (FOR 1617), 178833530 (SFB 940), 402170461 (TRR 265), 290210763 (VE 892/2-1), NE 1383/14-1); Faculty of Medicine at the Technische Universität Dresden, MeDDrive Grant; the Medical Research Foundation and Medical

Research Council (grants MR/R00465X/1 and MR/S020306/1), the National Institutes of Health (NIH) funded ENIGMA (grants 5U54EB020403-05 and 1R56AG058854-01). Further support was provided by grants from: the ANR (ANR-12-SAMA-0004, AAPG2019—GeBra), the Eranet Neuron (AF12-NEUR0008-01—WM2NA; and ANR-18-NEUR00002-01—ADoRe), the Fondation de France (00081242, 2012-00033703), the Fondation pour la Recherche Médicale (FRM; DPA20140629802; DPP20151033945), the Mission Interministérielle de Lutte-contre-les-Drogues-et-les-Conduites-Addictives (MILDECA), the Assistance-Publique-Hôpitaux-de-Paris and INSERM (interface grant), Paris Sud University IDEX 2012, the Fondation de l'Avenir (grant AP-RM-17-013), the Fédération pour la Recherche sur le Cerveau (FRC Neurodon 2015); the National Institutes of Health, Science Foundation Ireland (16/ERC/D/3797), U.S.A. (Axon, Testosterone and Mental Health during Adolescence; RO1 MH085772-01A1), and by NIH Consortium grant U54 EB020403, supported by a cross-NIH alliance that funds Big Data to Knowledge Centres of Excellence. Resources of The Center for Information Services and High Performance Computing (ZIH) at TU Dresden were used for fast data processing. We thank Jonas Granzow for help with data processing, quality control, and some figures and tables. We further wish to thank Annabell Baake, Leonie Epple, Carolin Fritzsche, and Isabell Theilig for assistance with data management and quality control. Lastly, we thank all participants and their families for their enduring commitment to the study over the years. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

Dr Banaschewski served in an advisory or consultancy role for ADHS digital, Infectopharm, Lundbeck, Medice, Neurim Pharmaceuticals, Oberberg GmbH, Roche, and Takeda. He received conference support or speaker's fee by Medice and Takeda. He received royalties from Hogrefe, Kohlhammer, CIP Medien, Oxford University Press. The present work is unrelated to the above grants and relationships. Dr Barker has received honoraria from General Electric Healthcare for teaching on scanner programming courses. Dr Poustka served in an advisory or consultancy role for Roche and Viforpharm and received speaker's fee by Shire. She received royalties from Hogrefe, Kohlhammer, and Schattauer. The present work is unrelated to the above grants and relationships. The other authors report no biomedical financial interests or potential conflicts of interest.

DATA AVAILABILITY STATEMENT

Access to the IMAGEN dataset is available with an accepted proposal from the IMAGEN executive committee (<https://imagen-europe.com/resources/imagen-project-proposal/>). All code is openly available in Open Science Framework at <https://osf.io/5tfh4/>.

ORCID

Lea L. Backhausen  <https://orcid.org/0000-0001-6482-1316>

Megan M. Herting  <https://orcid.org/0000-0002-0840-4582>

Michael N. Smolka  <https://orcid.org/0000-0001-5398-5569>

Nora C. Vetter  <https://orcid.org/0000-0003-1477-5395>

REFERENCES

- Akaike, H. (1973). *Information theory and extension of the maximum likelihood principle* (pp. 267–281).
- Backhausen, L. L., Herting, M. M., Buse, J., Roessner, V., Smolka, M. N., & Vetter, N. C. (2016). Quality control of structural MRI images applied using freesurfer—A hands-on workflow to rate motion artifacts. *Frontiers in Neuroscience*, 10, 558. <https://doi.org/10.3389/fnins.2016.00558>
- Bézivin-Frere, P., Vetter, N. C., Artiges, E., Filippi, I., Miranda, R., Vulser, H., Paillère-Martinot, M.-L., Ziesch, V., Conrod, P., Cattrell, A., Walter, H., Gallinat, J., Bromberg, U., Jurk, S., Menningen, E., Frouin, V., Papadopoulos Orfanos, D., Stringaris, A., Penttilä, J., ... Lemaître, H. (2020). Sex effects on structural maturation of the limbic system and outcomes on emotional regulation during adolescence. *NeuroImage*, 210, 116441. <https://doi.org/10.1016/j.neuroimage.2019.116441>
- Boedhoe, P. S. W., van Rooij, D., Hoogman, M., Twisk, J. W. R., Schmaal, L., Abe, Y., Alonso, P., Ameis, S. H., Anik, A., Anticevic, A., Arango, C., Arnold, P. D., Asherson, P., Assogna, F., Auzias, G., Banaschewski, T., Baranov, A., Batistuzzo, M. C., Baumeister, S., ... van den Heuvel, O. A. (2020). Subcortical brain volume, regional cortical thickness, and cortical surface area across disorders: findings from the ENIGMA ADHD, ASD, and OCD working groups. *American Journal of Psychiatry*, 177(9), 834–843. <https://doi.org/10.1176/appi.ajp.2020.19030331>
- Brain Development Cooperative Group. (2012). Total and regional brain volumes in a population-based normative sample from 4 to 18 years: The NIH MRI Study of Normal Brain Development. *Cerebral Cortex* (New York, N.Y.: 1991), 22(1), 1–12. <https://doi.org/10.1093/cercor/bhr018>
- Dennison, M., Whittle, S., Yücel, M., Vijayakumar, N., Kline, A., Simmons, J., & Allen, N. B. (2013). Mapping subcortical brain maturation during adolescence: Evidence of hemisphere- and sex-specific longitudinal changes. *Developmental Science*, 16(5), 772–791. <https://doi.org/10.1111/desc.12057>
- Dima, D., Modabbernia, A., Papachristou, E., Doucet, G. E., Agartz, I., Aghajani, M., Akudjedu, T. N., Albajes-Eizaguirre, A., Alnæs, D., Alpert, K. I., Andersson, M., Andreassen, N. C., Andreassen, O. A., Asherson, P., Banaschewski, T., Bargallo, N., Baumeister, S., Baur-Streubel, R., Bertolino, A., ... Frangou, S. (2021). Subcortical volumes across the lifespan: Data from 18,605 healthy individuals aged 3–90 years. *Human Brain Mapping*, 43, 452–469. <https://doi.org/10.1002/hbm.25320>
- Durston, S., Hulshoff Pol, H. E., Casey, B. J., Giedd, J. N., Buitelaar, J. K., & Van Engeland, H. (2001). Anatomical MRI of the Developing Human Brain: What Have We Learned? *Journal of the American Academy of Child & Adolescent Psychiatry*, 40(9), 1012–1020. <https://doi.org/10.1097/00004583-200109000-00009>
- Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., & Dale, A. M. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3), 341–355. [https://doi.org/10.1016/S0896-6273\(02\)00569-X](https://doi.org/10.1016/S0896-6273(02)00569-X)
- Giedd, J. N., Snell, J. W., Lange, N., Rajapakse, J. C., Casey, B. J., Kozuch, P. L., Vaituzis, A. C., Vauss, Y. C., Hamburger, S. D., Kaysen, D., & Rapoport, J. L. (1996). Quantitative magnetic resonance imaging of human brain development: Ages 4–18. *Cerebral Cortex*, 6(4), 551–559. <https://doi.org/10.1093/cercor/6.4.551>
- Goddings, A.-L., Mills, K. L., Clasen, L. S., Giedd, J. N., Viner, R. M., & Blakemore, S.-J. (2014). The influence of puberty on subcortical brain development. *NeuroImage*, 88, 242–251. <https://doi.org/10.1016/j.neuroimage.2013.09.073>
- Gogtay, N., & Thompson, P. M. (2010). Mapping gray matter development: Implications for typical development and vulnerability to psychopathology. *Brain and Cognition*, 72(1), 6–15. <https://doi.org/10.1016/j.bandc.2009.08.009>
- Herting, M. M., Johnson, C., Mills, K. L., Vijayakumar, N., Dennison, M., Liu, C., Goddings, A.-L., Dahl, R. E., Sowell, E. R., Whittle, S., Allen, N. B., & Tamnes, C. K. (2018). Development of subcortical volumes across adolescence in males and females: A multisample study of longitudinal changes. *NeuroImage*, 172, 194–205. <https://doi.org/10.1016/j.neuroimage.2018.01.020>
- Hoogman, M., Bralten, J., Hibar, D. P., Mennes, M., Zwiers, M. P., Schweren, L. S. J., van Hulzen, K. J. E., Medland, S. E., Shumskaya, E., Jahanshad, N., de Zeeuw, P., Szekely, E., Sudre, G., Wolfers, T., Onnink, A. M. H., Dammers, J. T., Mostert, J. C., Vives-Gilbert, Y., Kohls, G., ... Franke, B. (2017). Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: A cross-sectional mega-analysis. *The Lancet Psychiatry*, 4(4), 310–319. [https://doi.org/10.1016/S2215-0366\(17\)30049-4](https://doi.org/10.1016/S2215-0366(17)30049-4)
- Jovicich, J., Mariztoni, M., Sala-Llonch, R., Bosch, B., Bartrés-Faz, D., Arnold, J., Benninghoff, J., Wiltfang, J., Roccatagliata, L., Nobili, F., Hensch, T., Tränkner, A., Schönknecht, P., Leroy, M., Lopes, R., Bordet, R., Chanoine, V., Ranjeva, J.-P., Didic, M., ... Frisoni, G. B. (2013). Brain morphometry reproducibility in multi-center 3T MRI studies: A comparison of cross-sectional and longitudinal segmentations. *NeuroImage*, 83, 472–484. <https://doi.org/10.1016/j.neuroimage.2013.05.007>
- Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry*, 62(6), 593–602. <https://doi.org/10.1001/archpsyc.62.6.593>
- Koolschijn, P. C. M. P., & Crone, E. A. (2013). Sex differences and structural brain maturation from childhood to early adulthood. *Developmental Cognitive Neuroscience*, 5, 106–118. <https://doi.org/10.1016/j.dcn.2013.02.003>
- Lenroot, R. K., Gogtay, N., Greenstein, D. K., Wells, E. M., Wallace, G. L., Clasen, L. S., Blumenthal, J. D., Lerch, J., Zijdenbos, A. P., Evans, A. C., Thompson, P. M., & Giedd, J. N. (2007). Sexual dimorphism of brain developmental trajectories during childhood and adolescence. *NeuroImage*, 36(4), 1065–1073. <https://doi.org/10.1016/j.neuroimage.2007.03.053>
- Medawar, E., Thieleking, R., Manuilova, I., Paerisch, M., Villringer, A., Witte, A. V., & Beyer, F. (2021). Estimating the effect of a scanner upgrade on measures of grey matter structure for longitudinal designs. *PLoS One*, 16(10), e0239021. <https://doi.org/10.1371/journal.pone.0239021>
- Mills, K. L., Siegmund, K. D., Tamnes, C. K., Ferschmann, L., Wierenga, L. M., Bos, M. G. N., Luna, B., Li, C., & Herting, M. M. (2021). Inter-individual variability in structural brain development from late childhood to young adulthood. *NeuroImage*, 242, 118450. <https://doi.org/10.1016/j.neuroimage.2021.118450>
- Narvacan, K., Treit, S., Camicioli, R., Martin, W., & Beaulieu, C. (2017). Evolution of deep gray matter volume across the human lifespan. *Human Brain Mapping*, 38(8), 3771–3790. <https://doi.org/10.1002/hbm.23604>
- Noordermeer, S. D., Luman, M., Greven, C. U., Veroude, K., Faraone, S. V., Hartman, C. A., Hoekstra, P. J., Franke, B., Buitelaar, J. K., Heslenfeld, D. J., & Oosterlaan, J. (2017). Structural brain abnormalities of attention-deficit/hyperactivity disorder with oppositional defiant disorder. *Biological Psychiatry*, 82(9), 642–650. <https://doi.org/10.1016/j.biopsych.2017.07.008>
- Østby, Y., Tamnes, C. K., Fjell, A. M., Westlye, L. T., Due-Tønnessen, P., & Walhovd, K. B. (2009). Heterogeneity in subcortical brain development: A structural magnetic resonance imaging study of brain maturation from 8 to 30 years. *Journal of Neuroscience*, 29(38), 11772–11782. <https://doi.org/10.1523/JNEUROSCI.1242-09.2009>

- Paus, T., Keshavan, M., & Giedd, J. N. (2008). Why do many psychiatric disorders emerge during adolescence? *Nature Reviews Neuroscience*, 9(12), 947–957. <https://doi.org/10.1038/nrn2513>
- Pedersen, E. J., Miller, D. L., Simpson, G. L., & Ross, N. (2019). Hierarchical generalized additive models in ecology: An introduction with mgcv. *PeerJ*, 7, e6876. <https://doi.org/10.7717/peerj.6876>
- Petermann, F., & Petermann, U. (2007). *Hamburg wechsler intelligenztest für kinder–IV (HAWIK-IV)*. Huber.
- Petersen, A. C., Crockett, L., Richards, M., & Boxer, A. (1988). A self-report measure of pubertal status: Reliability, validity, and initial norms. *Journal of Youth and Adolescence*, 17(2), 117–133. <https://doi.org/10.1007/BF01537962>
- Plitman, E., Bussy, A., Valiquette, V., Salaciak, A., Patel, R., Cupo, L., Bèland, M.-L., Tullo, S., Tardif, C. L., Rajah, M. N., Near, J., Devenyi, G. A., & Chakravarty, M. M. (2021). The impact of the Siemens Tim Trio to Prisma upgrade and the addition of volumetric navigators on cortical thickness, structure volume, and 1H-MRS indices: An MRI reliability study with implications for longitudinal study designs. *NeuroImage*, 238, 118172. <https://doi.org/10.1016/j.neuroimage.2021.118172>
- Raznahan, A., Shaw, P. W., Lerch, J. P., Clasen, L. S., Greenstein, D., Berman, R., Pipitone, J., Chakravarty, M. M., & Giedd, J. N. (2014). Longitudinal four-dimensional mapping of subcortical anatomy in human development. *Proceedings of the National Academy of Sciences*, 111(4), 1592–1597. <https://doi.org/10.1073/pnas.1316911111>
- Reuter, M., Rosas, H. D., & Fischl, B. (2010). Highly accurate inverse consistent registration: A robust approach. *NeuroImage*, 53(4), 1181–1196. <https://doi.org/10.1016/j.neuroimage.2010.07.020>
- Reuter, M., Schmansky, N. J., Rosas, H. D., & Fischl, B. (2012). Within-subject template estimation for unbiased longitudinal image analysis. *NeuroImage*, 61(4), 1402–1418. <https://doi.org/10.1016/j.neuroimage.2012.02.084>
- Schumann, G., Loth, E., Banaschewski, T., Barbot, A., Barker, G., Büchel, C., Conrod, P. J., Dalley, J. W., Flor, H., Gallinat, J., Garavan, H., Heinz, A., Itterman, B., Lathrop, M., Mallik, C., Mann, K., Martinot, J.-L., Paus, T., Poline, J.-B., ... Struve, M. (2010). The IMAGEN study: Reinforcement-related behaviour in normal brain function and psychopathology. *Molecular Psychiatry*, 15(12), 1128–1139. <https://doi.org/10.1038/mp.2010.4>
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461–464.
- Shaw, P., Gogtay, N., & Rapoport, J. (2010). Childhood psychiatric disorders as anomalies in neurodevelopmental trajectories. *Human Brain Mapping*, 31(6), 917–925. <https://doi.org/10.1002/hbm.21028>
- Sled, J. G., Zijdenbos, A. P., & Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Transactions on Medical Imaging*, 17(1), 87–97. <https://doi.org/10.1109/42.668698>
- Sowell, E. R., Trauner, D. A., Gamst, A., & Jernigan, T. L. (2002). Development of cortical and subcortical brain structures in childhood and adolescence: A structural MRI study. *Developmental Medicine and Child Neurology*, 44(1), 4–16. <https://doi.org/10.1111/j.1469-8749.2002.tb00253.x>
- Tamnes, C. K., Herting, M. M., Goddings, A.-L., Meuwese, R., Blakemore, S.-J., Dahl, R. E., Güroğlu, B., Raznahan, A., Sowell, E. R., Crone, E. A., & Mills, K. L. (2017). Development of the cerebral cortex across adolescence: A multisample study of inter-related longitudinal changes in cortical volume, surface area, and thickness. *Journal of Neuroscience*, 37(12), 3402–3412. <https://doi.org/10.1523/JNEUROSCI.3302-16.2017>
- Tamnes, C. K., Walhovd, K. B., Dale, A. M., Østby, Y., Grydeland, H., Richardson, G., Westlye, L. T., Roddey, J. C., Hagler, D. J., Duenes, P., Holland, D., & Fjell, A. M. (2013). Brain development and aging: Overlapping and unique patterns of change. *NeuroImage*, 68, 63–74. <https://doi.org/10.1016/j.neuroimage.2012.11.039>
- Wierenga, L. M., Bos, M. G. N., Schreuders, E., vd Kamp, F., Peper, J. S., Tamnes, C. K., & Crone, E. A. (2018). Unraveling age, puberty and testosterone effects on subcortical brain development across adolescence. *Psychoneuroendocrinology*, 91, 105–114. <https://doi.org/10.1016/j.psyneuen.2018.02.034>
- Wierenga, L. M., Bos, M. G. N., van Rossenberg, F., & Crone, E. A. (2019). Sex effects on development of brain structure and executive functions: Greater variance than mean effects. *Journal of Cognitive Neuroscience*, 31(5), 730–753. https://doi.org/10.1162/jocn_a_01375
- Wierenga, L. M., Langen, M., Ambrosino, S., van Dijk, S., Oranje, B., & Durston, S. (2014). Typical development of basal ganglia, hippocampus, amygdala and cerebellum from age 7 to 24. *NeuroImage*, 96, 67–72. <https://doi.org/10.1016/j.neuroimage.2014.03.072>
- Wierenga, L. M., Sexton, J. A., Laake, P., Giedd, J. N., Tamnes, C. K., & Pediatric Imaging, Neurocognition, and Genetics Study. (2018). A key characteristic of sex differences in the developing brain: Greater variability in brain structure of boys than girls. *Cerebral Cortex (New York, N.Y.: 1991)*, 28(8), 2741–2751. <https://doi.org/10.1093/cercor/bhx154>
- Wonderlick, J. S., Ziegler, D. A., Hosseini-Varnamkhasi, P., Locascio, J. J., Bakkour, A., van der Kouwe, A., Triantafyllou, C., Corkin, S., & Dickerson, B. C. (2009). Reliability of MRI-derived cortical and subcortical morphometric measures: Effects of pulse sequence, voxel geometry, and parallel imaging. *NeuroImage*, 44(4), 1324–1333. <https://doi.org/10.1016/j.neuroimage.2008.10.037>
- Wood, S. N. (2004). Stable and efficient multiple smoothing parameter estimation for generalized additive models. *Journal of the American Statistical Association*, 99(467), 673–686. <https://doi.org/10.1198/016214504000000980>
- Wood, S. N. (2006). *Generalized additive models: An introduction with R*. Chapman and Hall/CRC. <https://doi.org/10.1201/9781420010404>
- Zahn-Waxler, C., Shirtcliff, E. A., & Marceau, K. (2008). Disorders of childhood and adolescence: Gender and psychopathology. *Annual Review of Clinical Psychology*, 4(1), 275–303. <https://doi.org/10.1146/annurev.clinpsy.3.022806.091358>
- Zheng, W., Chee, M. W. L., & Zagorodnov, V. (2009). Improvement of brain segmentation accuracy by optimizing non-uniformity correction using N3. *NeuroImage*, 48(1), 73–83. <https://doi.org/10.1016/j.neuroimage.2009.06.039>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Backhausen, L. L., Fröhner, J. H., Lemaitre, H., Artiges, E., Martinot, M.-L.P., Herting, M. M., Sticca, F., Banaschewski, T., Barker, G. J., Bokde, A. L. W., Desrivières, S., Flor, H., Grigas, A., Garavan, H., Gowland, P., Heinz, A., Brühl, R., Nees, F., Papadopoulos-Orfanos, D., Poustka, L., Hohmann, S., Robinson, L., Walter, H., Winterer, J., Whelan, R., Schumann, G., Martinot, J.-L., Smolka, M. N., Vetter, N. C., & the IMAGEN Consortium (2024). Adolescent to young adult longitudinal development of subcortical volumes in two European sites with four waves. *Human Brain Mapping*, 45(3), e26574. <https://doi.org/10.1002/hbm.26574>