Longitudinal allometry of sulcal morphology in health and schizophrenia

Joost Janssen[1,2,3,4,*], Clara Alloza[1,2,3], Covadonga M. Díaz-Caneja[1,2,3,5], Javier Santonja[1,2], Laura Pina-Camacho[1,2,3,5], Pedro M. Gordaliza[6], Alberto Fernández-Pena[2,6], Noemi González Lois[1,2], Elizabeth E.L. Buimer[4], Neeltje E.M. van Haren[4,7], Wiepke Cahn[4], Eduardo Vieta[3,9], Josefina Castro-Fornieles[3,10], Miquel Bernardo[3,8], Mara Parellada[1,2,3,5], Celso Arango[1,2,3,5], René S. Kahn[4,11], Hilleke E. Hulshoff Pol[4], Hugo G. Schnack[4]

[1]Department of Child and Adolescent Psychiatry, Institute of Psychiatry and Mental Health, Hospital General Universitario Gregorio Marañón, Madrid, Spain.
[2]Instituto de Investigación Sanitaria Gregorio Marañón (IiSGM), Madrid, Spain.
[3]Ciber del Área de Salud Mental (CIBERSAM).
[4]Department of Psychiatry, UMCU Brain Center, University Medical Center Utrecht, Utrecht, The Netherlands.
[5]School of Medicine, Universidad Complutense, Madrid, Spain.
[6]Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, Madrid, Spain.
[7]Department of Child and Adolescent Psychiatry/Psychology, Erasmus University Medical Centre, Sophia Children's Hospital, Rotterdam, The Netherlands.
[8]Barcelona Clinic Schizophrenia Unit, Hospital Clinic of Barcelona, Neuroscience Institute, IDIBAPS, University of Barcelona, Barcelona, Spain.
[9]Bipolar Disorders Unit, Clinical Institute of Neurosciences, Hospital Clínic, University of Barcelona, IDIBAPS, Barcelona, Spain.
[10]Department of Child and Adolescent Psychiatry and Psychology, Clinical Institute of Neurosciences, Hospital Clínic, University of Barcelona, IDIBAPS, Barcelona, Spain
[11]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, United States

Abstract

Scaling between subcomponents of cortical folding and total brain volume (TBV) in healthy individuals (HI) is allometric, i.e. non-linear. It is unclear whether this is also true in individuals with schizophrenia (SZ) or first-episode psychosis (FEP). The current study first confirmed normative allometric scaling norms in HI using discovery and replication samples. Cross-sectional and longitudinal diagnostic differences in folding subcomponents were then assessed using an allometric analytic framework.

Structural imaging from a longitudinal (sample 1: HI and SZ, $n_{HI\ Baseline}$ = 298, $n_{SZ\ Baseline}$ = 169, $n_{HI\ Follow-up}$ = 293, $n_{SZ\ Follow-up}$ = 168, a total of 1087 images, all individuals ≥ 2 images, age 16-69 years) and a cross-sectional sample (sample 2: n_{HI} = 61 and n_{FEP} = 89, age 10-30 years) is leveraged to calculate global folding and its nested subcomponents: sulcation index (SI, total sulcal/cortical hull area) and determinants of sulcal area; sulcal length and sulcal depth.

Scaling of the SI, sulcal area, and sulcal length with TBV in SZ and FEP was allometric and did not differ from HI. Longitudinal age trajectories demonstrated steeper loss of SI and sulcal area through adulthood in SZ. Longitudinal allometric analysis revealed that both annual change in SI and sulcal area was significantly stronger related to change in TBV in SZ compared to HI.

Our results detail the first evidence of the disproportionate contribution of changes in SI and sulcal area to TBV changes in SZ. Longitudinal allometric analysis of sulcal morphology provides deeper insight into lifespan trajectories of cortical folding in SZ.

Introduction

The degree of folding of the human cortex is an important marker of early-life neurodevelopmental processes (Llinares-Benadero and Borrell, 2019). Early-life neurodevelopmental processes are thought to be aberrant in first-episode psychosis (FEP) and schizophrenia (SZ). This notion is corroborated by consistently reported average reductions in the degree of folding and sulcation of FEP and SZ relative to healthy individuals (HI) (Penttilä et al., 2009; Palaniyappan et al., 2011; Janssen et al., 2014; Lavoie et al., 2014; Nanda et al., 2014; Nesvåg et al., 2014; Cachia et al., 2015; Gay et al., 2017; MacKinley et al., 2019; Dazzan et al., 2020). However, the degree of folding is not static throughout life in HI and reduces postnatally, during adolescence, and beyond (Alemán-Gómez et al., 2013; Mutlu et al., 2013; Li et al., 2014; Cao et al., 2017).

Longitudinal studies are necessary to establish whether SZ is associated with abnormalities in folding reductions and whether the pattern of SZ-related abnormalities may be dynamic over time. In addition, longitudinal studies can clarify whether SZ-related folding abnormalities can also occur at later stages of the illness during late adulthood and old age (Mirakhur et al., 2009; Raznahan et al., 2011; Li et al., 2014).

The in vivo degree of folding can be estimated using the gyrification index or the sulcation index (SI) (Shimony et al., 2016). These metrics are composite scores that combine two components: hull surface area (exposed cortex) and outer cortex in the case of the gyrification index and sulcal surface area (non-exposed cortex) in the case of SI. Sulcal surface area can in turn can be decomposed into sulcal depth and length allowing for a complete decomposition of the SI into folding subcomponents (see Figure 1) (Mangin et al., 2004; Alemán-Gómez et al., 2013). Assessing the SI and its subcomponents in a longitudinal lifespan study of HI and individuals with SZ allows for a fine-grained assessment of change in the degree of folding across the lifespan and whether SZ contributes distinctively to the folding subcomponents.
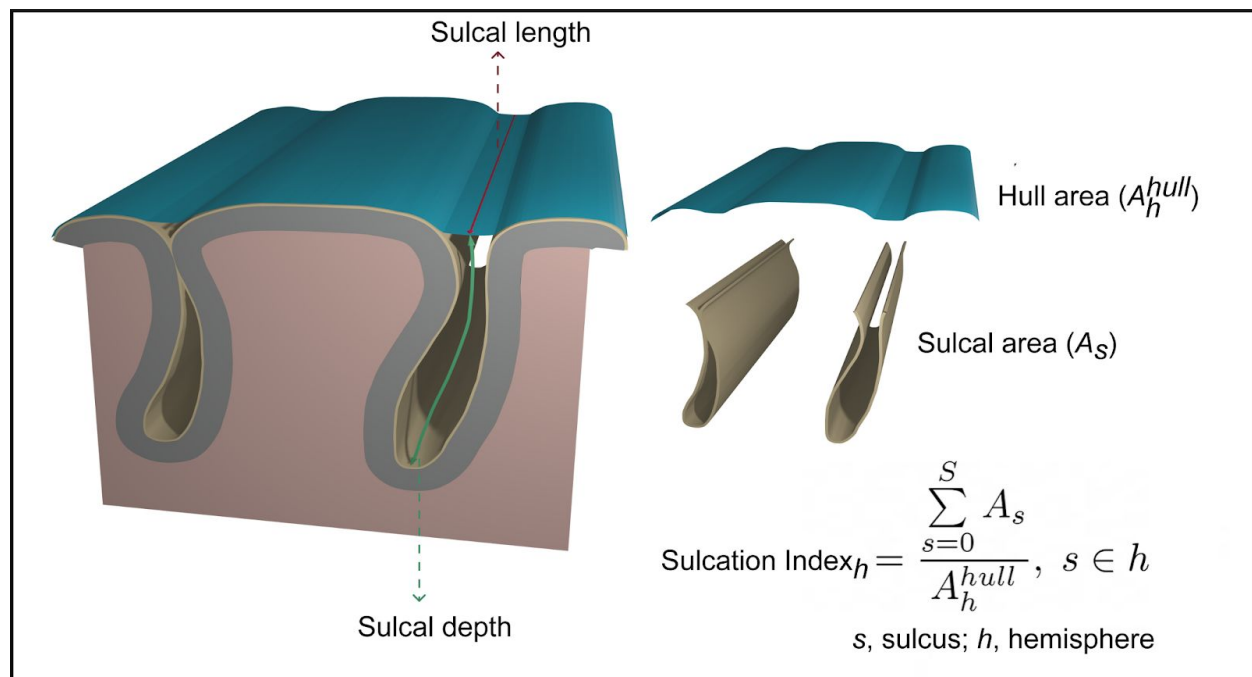
Reduced total brain volume (TBV) is a robust neurobiological feature of psychotic disorders (Haijma et al., 2013). Since smaller brain sizes are associated with a lower degree of folding and early-life folding growth positively correlates with brain volume growth, the question arises whether reduced folding in psychotic disorders may be merely due to a global decrease in brain size (Toro et al., 2008). Accounting for brain size in comparisons of folding metrics between individuals with psychotic disorders and HI is usually done using traditional methods (e.g. including TBV as a nuisance variable). However, these approaches assume an identical linear association between folding and TBV across groups. However, the relationship between brain metrics may be allometric, i.e. non-linear (Im et al., 2008; Toro et al., 2008; Barnes et al., 2010; Reardon et al., 2016; de Jong et al., 2017; Mankiw et al., 2017; Jäncke et al., 2019). Using a HI sample in an allometric framework allows for providing normative scaling norms between folding subcomponents and TBV. These norms are necessary for determining whether the

effect of SZ on subcomponents of folding is due to scaling differences or whether folding abnormalities are present independently from group differences in TBV.

Allometric scaling laws predict that the static relationship between brain volume and areal measures (sulcal and hull area) in a log-log framework should be ⅔ while the relationship with linear measures (sulcal depth and length) should be ⅓ (Toro et al., 2008; Fish et al., 2017). However, the only study so far to assess scaling of brain volume with sulcal folding metrics in neurotypical children show that larger brains have a higher SI due to a disproportionate increase in sulcal length but not depth (Fish et al., 2017). While interesting, it is unclear whether the observed allometric scaling also affects longitudinal trajectories of folding metrics and modulates potential differences in trajectories between SZ and HI.

In this study we first set out to replicate the previously reported normative scaling norms between TBV and the folding metrics in two independent samples of HI. Thereafter, we assess separately in each sample whether SZ or FEP groups differ from the HI group in allometric scaling. We then compare longitudinal age trajectories of folding metrics between SZ and HI. Finally we apply we longitudinal allometric scaling to examine how change in folding subcomponents relate to change in TBV and whether this scaling differs between SZ and health.

Figure 1. Graphical representation of SI and its components. Total sulcal surface is defined as the sum of the area of all sulcal median meshes. Hull area is derived by calculating the area of the brain mesh defined via a morphologic closing of the brain mask which excludes sulcal areas during definition of the exposed cortical surface area. Total sulcal length is defined as the sum of the length of all median sulcal meshes. Geodesic sulcal depth is the shortest distance along the surface of the brain from the point to where the brain surface makes contact with the hull surface.

$$\text{Sulcation Index}_h = \frac{\sum\limits_{s=0}^{S} A_s}{A_h^{hull}}, \quad s \in h$$

Hull area (A_h^{hull})

Sulcal area (A_s)

s, sulcus; h, hemisphere

Materials and Methods

Our study uses two independent samples (sample 1, sample 2). Detailed clinical, demographic and imaging measures for both samples are summarized in Table 1 and 2.

Table 1. Demographic, cognitive and clinical characteristics of individuals with schizophrenia and healthy individuals for sample 1.

	Schizophrenia patients			Healthy individuals		
Number of scans	**Baseline** 169	**Follow-up 1** 168	**Follow-up 2** 50	**Baseline** 298	**Follow-up 1** 293	**Follow-up 2** 109
Demographics						
Sex, n (%), males	131 (77.51)	-	-	167 (56)	-	-
Age, years: mean (SD)	29.86 (9.27)	34 (9.55)	34.4 (5.75)	30.42 (10.96)	34.41 (10.96)	33.96 (8.10)
Education, years: mean (SD)	20.21 (25.7)	19.75 (24.9)	NA	21.75 (25)	20.19 (22.88)	NA
Cognitive and clinical variables						
Full-scale IQ total: mean (SD)	97.86 (16.5)	103 (20.22)	100.42 (18.58)	110.34 (15.74)	113.9 (16.5)	115.62 (17.6)
Age of onset, years: mean (SD)[1]	22.13 (5.28)	-	-			
Duration of illness, years: mean, (SD)[2]	6.32 (6.63)	11.04 (7.7)	11.84 (3.91)			
PANSS score, total: mean (SD)	63.34 (19.26)	50.18 (14.2)	56.2 (13.97)			
Total Positive symptoms	14.96 (5.82)	12.37 (4.62)	13.26 (3.83)			
Total Negative symptoms	16.27 (6.22)	12.87 (5.64)	15.22 (5.45)			
Total General symptoms	31.22 (11.2)	24.93 (7.02)	27.71 (7.28)			
Antipsychotics[3]						
Without medication, n (%)	7 (4.14)	7 (4.16)	NA			
Exclusively typical, n (%)	23 (13.6)	12 (7.14)	NA			
Exclusively atypical, n (%)	89 (52.66)	92 (54.76)	NA			
Both, n (%)	4 (2.36)	2 (1.19)	NA			
Cumulative dose exposure (mg): mean (SD)	275837.8 (397345.9)	945754.9 (1285851)	NA			
Missing, n (%)	103 (61.54)	102 (60.71)	50 (100)			

The total amount of scans is 1087. [1]Age of onset is the age at which the first positive symptom occurs. [2]Duration of illness is calculated based on the date of the appearance of first positive symptoms of schizophrenia until the date of scan. [3]Antipsychotic treatment data is not available for all the patients included (82 patients had medication information). Cumulative doses were calculated at the time of scan and given in chlorpromazine equivalents using standard conversion factors and estimated by the daily doses of each of the antipsychotics used by the patient.

Table 2. Demographic, cognitive and clinical variables of individuals with schizophrenia or first-episode psychosis and healthy individuals of samples 1 (at baseline) and 2.

	Schizophrenia patients (n=169)	Healthy individuals (n=298)
Sociodemographics		
Sex, n (%), males	131 (77.51)	167 (56)
Age, years: mean (SD)	29.86 (9.27)	30.42 (10.96)
Range		
Education, years: mean (SD)	20.21 (25.7)	21.75 (25)
Cognitive and clinical variables		
Full-scale IQ total: mean (SD)	97.86 (16.5)	110.34 (15.74)
Age of onset, years: mean (SD)[1]	22.13 (5.28)	
Duration of illness, years: mean, (SD)[2]	6.32 (6.63)	
PANSS score, total: mean (SD)	63.34 (19.26)	
Total Positive symptoms	14.96 (5.82)	
Total Negative symptoms	16.27 (6.22)	
Total General symptoms	31.22 (11.2)	
Antipsychotics[3]		
Without medication, n (%)	7 (4.14)	
Exclusively typical, n (%)	23 (13.6)	
Exclusively atypical, n (%)	89 (52.66)	
Both, n (%)	4 (2.36)	
Cumulative dose exposure at scan (mg): mean (SD)	275837.8 (397345.9)	

	FEP patients (n=89)	Healthy individuals (n=61)	Missing, n (%)
			103 (61.54)
Sociodemographics			
Sex, n (%), males	56 (70)	36 (62)	
Age, years: mean (SD)	21.55 (4.66)	22.85 (5.14)	
Cognitive and clinical variables			
Estimated IQ total: mean (SD)	92.40 (15.23)	106.58 (11.58)	
Age of onset, years: mean (SD)[1]	20.83 (4.64)		
Duration of illness, days: mean, (SD)[2]	157.88 (109.6)		
PANSS score, total: mean (SD)	74.7 (21.68)		
Total Positive symptoms	18.20 (7.6)		
Total Negative symptoms	18.61 (8.15)		
Total General symptoms	37.89 (11.41)		
Antipsychotic use (AP)[3]			
Use of AP at scan, n (%)	74 (93%)		
Cumulative dose at scan (mg): mean (SD)	33330.95 (40504.16)		

FEP, first-episode psychosis; SD, standard deviation.; IQ, intelligence quotient; PANSS, Positive and Negative Symptoms Scale; AP, antipsychotic use. [1]Age of onset is the age at which the first positive symptoms occurred. [2]Duration of illness is calculated based on the date of the appearance of first positive symptoms until the date of scan. [3]Antipsychotic treatment data is not available for all the patients included (82 patients had medication information). Cumulative doses were calculated per time of scan and given in chlorpromazine equivalents using standard conversion factors and estimated by the daily doses of each of the antipsychotics used by the patient.

Sample 1: schizophrenia and healthy individuals (longitudinal)

From a large longitudinal sample of individuals with schizophrenia and HI aged 16-70 years at baseline we included individuals who had T1-weighted magnetic resonance imaging (MRI) scan acquisitions at two or three time points. Detailed information regarding diagnostic criteria and clinical and cognitive assessments of the Utrecht Schizophrenia project and the Genetic Risk and Outcome of Psychosis (GROUP) consortium are described in (Hulshoff Pol et al., 2001; Korver et al., 2012; Kubota et al., 2015). Detailed information on the current subsample can be found in (Janssen et al., 2020). Details on the MRI quality assessment (both FreeSurfer and BrainVISA) and time between scans can be found in the Supplemental text and SFigures 1 and 2. The final dataset consisted of 1087 scans from 298 HI and 168 patients. Demographic, cognitive and clinical information of the participants at all time points is summarized in Table 1 and information at baseline is summarized in Table 2 (left column). The study was approved by the institutional review board and was conducted according to the provisions of the World Medical Association Declaration of Helsinki. Written informed consent was obtained from all participants and also from parents/legal guardians for children under 18 years of age.

Sample 2: first-episode psychosis and healthy individuals (cross-sectional)

The sample is a subset of participants with available imaging data from a larger multi-site, naturalistic study including participants aged 7-35 years with a first episode of psychosis, the PEPs Project. The main objective of the PEPs project was to identify the several gene-by-environment interactions involved in the risk of psychosis in a

naturalistic cohort of first-episode psychosis to develop a predictive model of psychosis. Most of the centers were integrated in the well-recognized Spanish network of research in Mental Health named "Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM)". Detailed information about the study design, recruitment procedures, in- and exclusion criteria, clinical and cognitive assessments, are provided elsewhere (Bernardo et al., 2013; Pina-Camacho et al., 2016). For the current study only participants scanned at a single site (Barcelona) were included (n=182) of which a further 32 participants were excluded based on image quality. Detailed demographic, cognitive and clinical information for the final sample is summarized in Table 2 (right column). Details on the MRI quality assessment can be found in the Supplemental text . The study was approved by the institutional review board and was conducted according to the provisions of the World Medical Association Declaration of Helsinki. Written informed consent was obtained from all participants and from parents/legal guardians for children under 18 years of age.

MRI acquisition

Image acquisition

For sample 1, two scanners (same vendor, field strength and acquisition protocol) were used. Participants were scanned twice on either a Philips Intera or Achieva 1.5 T and a T1-weighted, 3-dimensional, fast-field echo scan with 160-180 1.2 mm contiguous coronal slices (echo time [TE], 4.6 ms; repetition time [TR], 30 ms; flip angle, 30°; field of view [FOV], 256 mm; in-plane voxel size, 1x1 mm^2) was acquired. All included participants had their baseline and follow-up scan on the same scanner. For

sample 2, participants were scanned on a Siemens Trio TIM 3 T and a T1-weighted, 3-dimensional, gradient echo scan with 240 1 mm contiguous sagittal slices (echo time [TE], 2.98 ms; repetition time [TR], 2300 ms; flip angle, 9°; field of view [FOV], 256 mm; in-plane voxel size, 1x1 mm^2) was acquired.

Image analysis

TBV, SI, total sulcal and hull area, global average sulcal depth, and total sulcal length were assessed (see Figure 1). All images were analyzed using the FreeSurfer analysis suite (v5.1) with default settings to provide total gray and white matter volumes which were summed to calculate TBV (Dale et al., 1999; Fischl et al., 1999; Fischl, 2012). For all images, sulcal segmentation and identification was performed with BrainVISA software (v4.5) using the Morphologist Toolbox and Mindboggle software using default settings (Mangin et al., 2004, 2010; Klein et al., 2017). After importing Freesurfer's ribbon image into BrainVISA each sulcus is segmented with the cortical sulci corresponding to the crevasse bottoms of the "landscape," the altitude of which is defined by image intensity. A spherical hull surface was generated from a smooth envelope that wrapped around the hemisphere but did not encroach into the sulci, a morphological isotropic closing of 6 mm was applied to ensure boundary smoothness. The median sulcal surface spans the entire space contained in a sulcus, from the fundus to its intersection with the hull. For each fold, sulcal area is defined as the total surface area of the medial sulcal surface, sulcal length is measured on the hull and is defined as the distance of the intersection between the median sulcal surface and the hull. All sulcal nodes belonging to a hemisphere were relabeled to one label (see

SFigure 1) and then BrainVISA software was used to calculate hemispheric sulcal area and length. FreeSurfer output was imported into Mindboggle software for calculating geodesic sulcal fundi depth. All metrics were measured in the native space of the participant's images and left and right hemisphere values were either summed or averaged. FreeSurfer, BrainVISA and Mindboggle derived measurements have been validated via histological and manual measurements and have shown good test–retest reliability (Rosas et al., 2002; Kuperberg et al., 2003; Pizzagalli et al., 2020).

Statistics

All analyses were performed in R (https://cran.rstudio.com/).

Allometric analyses at baseline: samples 1 and 2

Normative scaling coefficients from healthy individuals

We first regressed out the effects of age (age, age^2, age^3), sex, and scanner (only for sample 1), by computing at each timepoint separately a linear regression model with each metric m as dependent variable and the covariates as predictors. To control for collinearity among the polynomials of age we made use of the poly()-function in R which created three uncorrelated variables to model age effects (so-called orthogonal polynomials) (34) for metric m (Chambers and Hastie, 1992):

$$m = \beta_0 + \beta_1(age) + \beta_2(sex) + \beta_3(scanner) \qquad (1)$$

We then saved the residuals and added the mean back in for interpretability.

We converted the residualized metrics from HI (samples 1 and 2 separately) at baseline to log values and applied orthogonal regression to calculate normative baseline allometric scaling coefficients of log transformed values of TBV with SI, sulcal

and hull area, sulcal depth, and length. In our scenario, both dependent and independent variables have similar measurement error. Therefore, in our case non-orthogonal regression may lead to biased regression estimates as only the measurement error of the independent variable is taken into account. In orthogonal regression the measurement error of both dependent and independent variables are accounted for, leading to higher accuracy of regression estimates (Madansky, 1959). This method allowed us to assess whether the normative baseline allometric scaling relationships between the log transformed TBV and folding subcomponents were similar in the discovery sample (sample 1) and the replication sample (sample 2) and those reported by Fish et al. (2017) and whether they were similar compared to the expected scaling coefficients. The t for the difference between the observed and expected coefficients was calculated as follows:

$$t = \frac{(\beta_1 - \beta_{expected})}{se_\beta} \sim T(n_{healthy\ individuals} - 1) \tag{2}$$

where β_1 is the scaling coefficient for the HI group, $\beta_{expected}$ is the scaling coefficient as expected by scaling laws, se_β is the standard error of the slope of the HI group given by bootstrapping and n is the number of participants. For each folding metric, scaling coefficients below and above the expected values were considered to indicate hypo- and hyperallometry, respectively. Coefficients that were statistically indistinguishable from the null indicated isometric scaling. The *mcr* R package was used for orthogonal regression.

Comparison of allometric scaling coefficients between healthy individuals and SZ (sample 1) and FEP (sample 2)

We converted the residualized metrics from SZ and FEP patients (sample 1 and 2 separately) to log values and applied orthogonal regression to compare baseline allometric scaling coefficients between HI and SZ (sample 1) and FEP (sample 2) separately. Significance of the slope was determined through 95% bias-corrected and accelerated bootstrap (BCa) confidence intervals generated using bootstrapping (n=1000) (Efron and Tibshirani, 1993). If the slopes of the SZ and healthy participant groups were both significant we calculated a t for the difference between the slopes of the patient and healthy participant groups, as follows:

$$t = \frac{(\beta_{patients} - \beta_{participants})}{se_{\beta}} \sim T(n_{patients} - 1) \tag{3}$$

where $\beta_{healthy\ individuals}$ is the scaling coefficient for the HI group, $\beta_{patients}$ is the coefficient for the SZ group, se_{β} is the standard error of the slope of the SZ group given by the bootstrap and $n_{patients}$ is the number of patients.

In order to directly compare the results of our allometric approach to traditional methods of controlling for effects of TBV we ran linear models comparing groups with expressing the residualized fold metrics as a fraction of TBV, i.e. normalization:

$$residuals\ m/TBV = \beta_0 + \beta_1 \times diagnosis \tag{4}$$

and with adding TBV as a covariate,

$$residuals\ m = \beta_0 + (\beta_1 \times diagnosis) + (\beta_2 \times TBV) \tag{5}$$

where m is the metric.

Longitudinal age trajectories in SZ and healthy individuals (sample 1)

We used generalized additive mixed models (GAMMs) to investigate the longitudinal trajectories of SI, sulcal and hull area, sulcal depth, and length over the age

range (Wood, SN, 2006). We specified cubic regression splines with shrinkage (a type of penalized splines) as smooth terms and set k = 4 as the number of knots of the spline. GAMM models were implemented to examine diagnosis (i.e., SZ vs HI), the age×diagnosis interaction, and TBV, while controlling for sex and scanner (as fixed effects) and the random effect of the individual. Including a dummy variable to account for scanner type as a random effect in the models did not change the results. In order to account for the potential effect of a non-linear relationship between cortical folding metrics and TBV on age trajectories we reran the GAMM models but now transformed TBV and the folding metrics to log values. In these models we also tested whether the log(TBV)×diagnosis interaction was significant. If this interaction was not significant it was omitted from the model. To better understand the age×diagnosis interaction, GAMM estimates for age were also implemented and visualized for patients and HI separately. We used the *mgcv* R package for modeling of the GAMMs. Finally, we plotted a difference fit (predicted SZ fit - predicted HI fit) with its 95% confidence interval against age. The difference fit allows for assessment of where along the age range the difference between SZ and HI becomes significantly different from zero.

Longitudinal allometric analysis: the relationship between annual change in fold metrics and annual change in brain volume in SZ and healthy individuals (sample 1)

In these analyses we assessed whether yearly changes in folding metrics scaled with yearly changes in TBV and whether this scaling differed between the SZ and HI groups. For this, we used the residuals from baseline and the nearest available

follow-up measurement from each subject to determine the symmetrized annual percent change (APC) for each subject and metric m:

$$APC_m(a) \ = 100 \ \frac{1}{\underline{m}} \frac{\Delta m}{\Delta a} \ = 100 \ \frac{2}{a_{i-1} - a_i} \frac{m_{i-1} - m_i}{m_{i-1} + m_i} \ , \qquad\qquad (6)$$

where Δm represents change in metric m, obtained at two different time points, i and $i-1$, $i \in [1, 2, 3]$ (up to three measures per subject in this study) and Δa the time lapse between them in years; a is the subject's age in years at a specific time point. Finally, $\overline{m}$ represents the average of the metric's value at the two time points. We then calculated longitudinal allometric scaling coefficients of the APC of TBV versus the APCs of the folding metrics using orthogonal regression in HI and individuals with SZ. Significance of the slope was determined through 95% bias-corrected and accelerated bootstrap (BCa) confidence intervals generated using bootstrapping (n=1000) (Efron and Tibshirani, 1993). If the slopes of the SZ and HI groups were both significant we calculated a t for the difference between the slopes of the patient and healthy participant group using the previously described formula (3).

In all analyses we controlled for multiple comparisons using Bonferroni correction.

Results

Allometric analyses at baseline: sample 1 and 2

*Normative scaling coefficients from healthy individuals*

As can be seen in Figure 2 the cross-sectional allometric findings replicated across the two independent samples and replicated the normative findings reported by Fish et al. (2017). In the healthy participant groups of sample 1 and 2, a larger SI was associated with a larger TBV. In both samples there was hyperallometric scaling of SI, sulcal area and length and isometric scaling of hull area. Sulcal depth showed either hypoallometric scaling (sample 1) or isometric scaling (sample 2).

Comparison of baseline allometric scaling coefficients between healthy individuals and SZ (sample 1) and FEP (sample 2)

Allometric analyses indicated that in both the SZ and FEP patient groups the SI, sulcal and hull area, and sulcal length were in proportion with normative decreases in TBV (see Figure 2). Allometric analysis clearly demonstrated that SI, sulcal area, and sulcal length scaled non-linearly with TBV in HI, SZ and FEP groups. Traditional methods of "controlling for" TBV differences, which assume linear scaling between the fold metrics and TBV, were not consistent with the results of our allometric analysis (see Table 3). There were no significant associations between the residualized brain metrics, CPZ equivalents, and illness duration. In the group of SZ patients, estimated IQ explained about 9% of the variance of the brain metrics, except for sulcal depth (see SFigure 3).

Figure 2. Orthogonal regression fits ± 95% confidence interval and the scaling coefficient for observed log–log allometric scaling of residualized folding metrics relative to residualized total brain volume in the healthy individuals and patient groups at baseline (samples 1 (left) and 2 (right) separately). "Expected" values (black line) represent the expected scaling (from scaling laws) coefficient. ***=p<0.001 denotes significance after Bonferroni correction for the difference between (1) the scaling of the healthy individuals groups and the expected scaling and (2) the healthy individuals groups and patient groups. Values were residualized for scanner (only sample 1), age and sex.

Sample 1: Schizophrenia
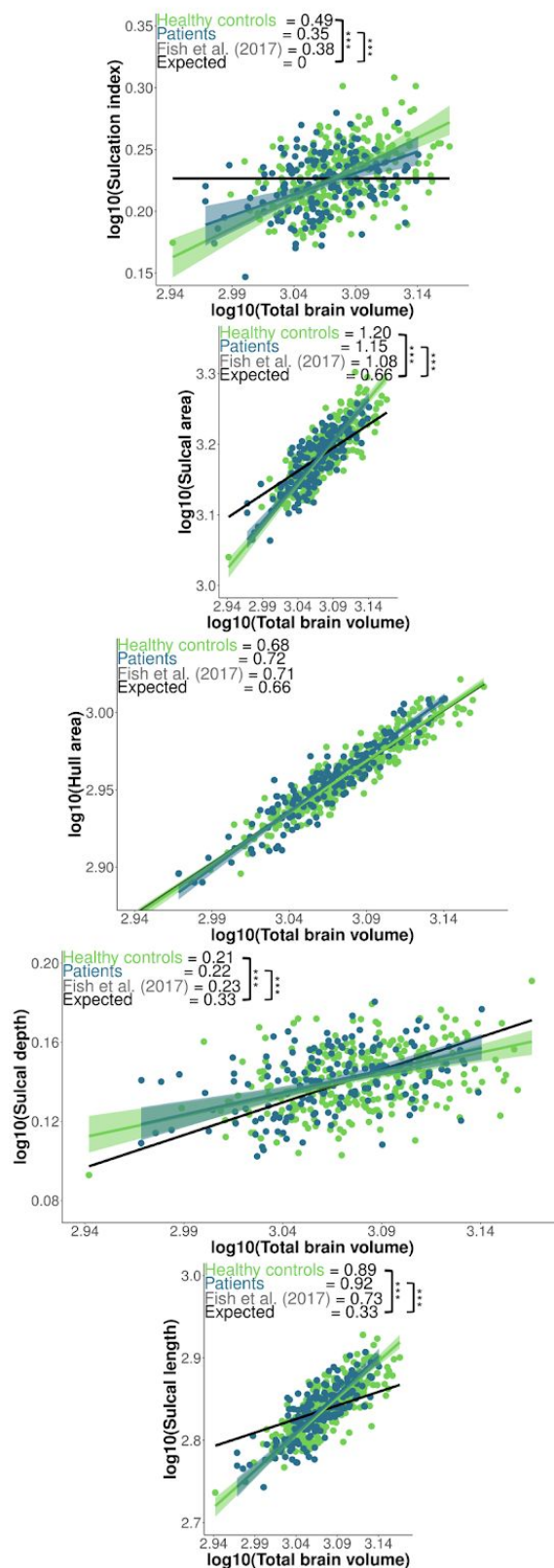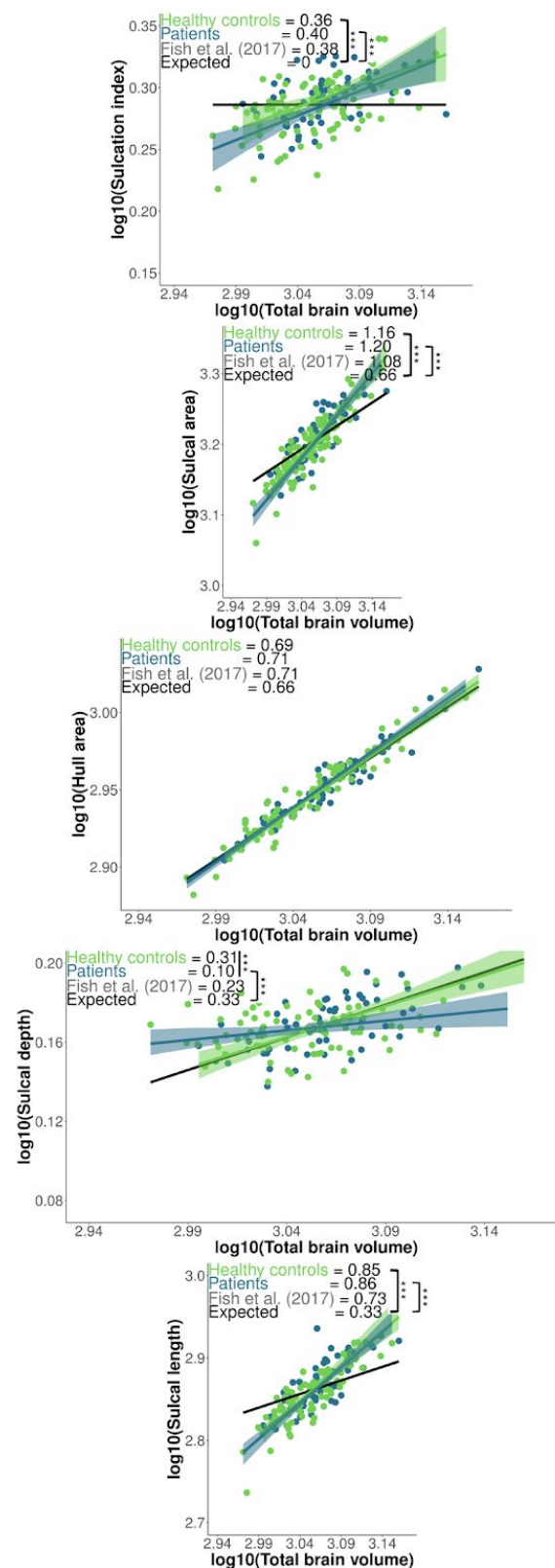
Sample 2: First-episode psychosis

Table 3. Comparing different methods of controlling for brain size in analysis of diagnostic differences at baseline.

	Schizophrenia vs control	First-episode psychosis vs control
Sulcation Index		
Normalization	↑	—
Covarying	—	↓
Allometry	—	—
Sulcal area		
Normalization	—	—
Covarying	—	—
Allometry	—	—
Hull area		
Normalization	↑	↑
Covarying	↑	—
Allometry	—	—
Sulcal depth		
Normalization	↑	—
Covarying	—	—
Allometry	—	↓
Sulcal length		
Normalization	↑	—
Covarying	—	—
Allometry	—	—

Folding metrics were compared between each patient group and their respective healthy individuals group using three different methods to control for total brain volume variation: normalizing by total brain volume, covarying by total brain volume, and within the allometric framework. Arrows indicate presence and direction of statistically significant group difference at $p < 0.05$.

Longitudinal age trajectories in SZ and healthy individuals (sample 1)

There was no main effect of diagnosis for any of the metrics (see Table 4). The healthy control and SZ group showed loss of TBV, sulcal area, and sulcal depth and

length across the lifespan (see Figure 3A). After Bonferroni correction the age×diagnosis interaction was significant for SI and sulcal area (see Table 4). Trajectory differences in SI were larger than zero from 32 years onwards approximately (approximate patient group deficit at 32 years and at the maximum age: -0.02 and -0.08, respectively) and for sulcal area from 44 years onwards, approximately (approximate patient group deficit at 44 years and at the maximum age: 18 and 35cm^2, respectively), see Figure 3B.

Table 4. Generalized Additive Mixed Model (GAMM) estimates for diagnosis and age*diagnosis for each brain metric in sample 1.

Brain metrics	GAMM estimates							
Total brain volume	**estimate**		**se**		**t**		**p**	
Diagnosis	-44.48		9.24		-4.82		**<0.001**	
	edf		**ref.df**		**F**		**p**	
Age:patients	0.92		3		704.97		0.05	
Sulcation Index	**estimate**		**se**		**t**		**p**	
	TBV	**log**	**TBV**	**log**	**TBV**	**log**	**TBV**	**log**
Diagnosis	-0.02	<-0.01	0.01	<0.01	-2.11	-2.05	0.04	0.04
	edf		**ref.df**		**F**		**p**	
Age:patients	1.34	1.39	3	3	642.66	741.41	**<0.001**	**<0.001**
Sulcal area	**estimate**		**se**		**t**		**p**	
Diagnosis	-8.36	<-0.01	7.86	<0.01	-1.06	-1.03	0.29	0.30
	edf		**ref.df**		**F**		**p**	
Age:patients	1.13	1.21	3	3	358.05	494.11	**<0.01**	**<0.01**
Hull area	**estimate**		**se**		**t**		**p**	
Diagnosis	3.01	<0.01	1.96	<0.01	1.54	1.51	0.12	0.13
	edf		**ref.df**		**F**		**p**	
Age:patients	1.21	2.71	3	3	942.27	1039.09	0.01	0.01
Sulcal depth	**estimate**		**se**		**t**		**p**	
Diagnosis	<0.01	<0.01	<0.01	<0.01	0.36	0.41	0.72	0.68
	edf		**ref.df**		**F**		**p**	
Age:patients	1.11	2.34	3	3	13.64	21.75	0.03	0.03
Sulcal length	**estimate**		**se**		**t**		**p**	
Diagnosis	-0.61	<-0.01	3.03	<0.01	-0.20	-0.21	0.84	0.83
	edf		**ref.df**		**F**		**p**	
Age:patients	0.16	0.42	3	3	0.84	6.63	0.26	0.20

Two types of models were implemented: 1. 'TBV' with TBV, sex and scanner as covariates and 2. 'log' with log(TBV), sex, and scanner as covariates and predicting

log(brain metric). Smooth function (edf), degrees of freedom (Ref.df), F-statistic are given. P-value in bold denotes significance after Bonferroni correction. Note diagnosis is coded as an ordered factor (healthy individuals=0; patients=1). TBV, Total Brain Volume.

Figure 3. **A**. Age trajectories of total brain volume and the five folding metrics in healthy individuals and individuals with schizophrenia from sample 1. Generalized Additive Mixed Model (GAMM) fits ± 95% confidence intervals are shown over the data points for all participants in the corresponding color, with data at each of the time points connected for each participant. Sex, total brain volume, and scanner were included in the models as covariates (except for the model for total brain volume) and the individual was modeled as a random effect. Sulcation Index and sulcal area showed a significant age $\times$ diagnosis interaction. **B**. Difference fits ± 95% confidence interval for Sulcation Index and sulcal area. The plots display at approximately what age the mean difference between the predicted fits of the healthy participant and patient groups becomes different from zero and show that the mean difference between the diagnostic groups increases with age.
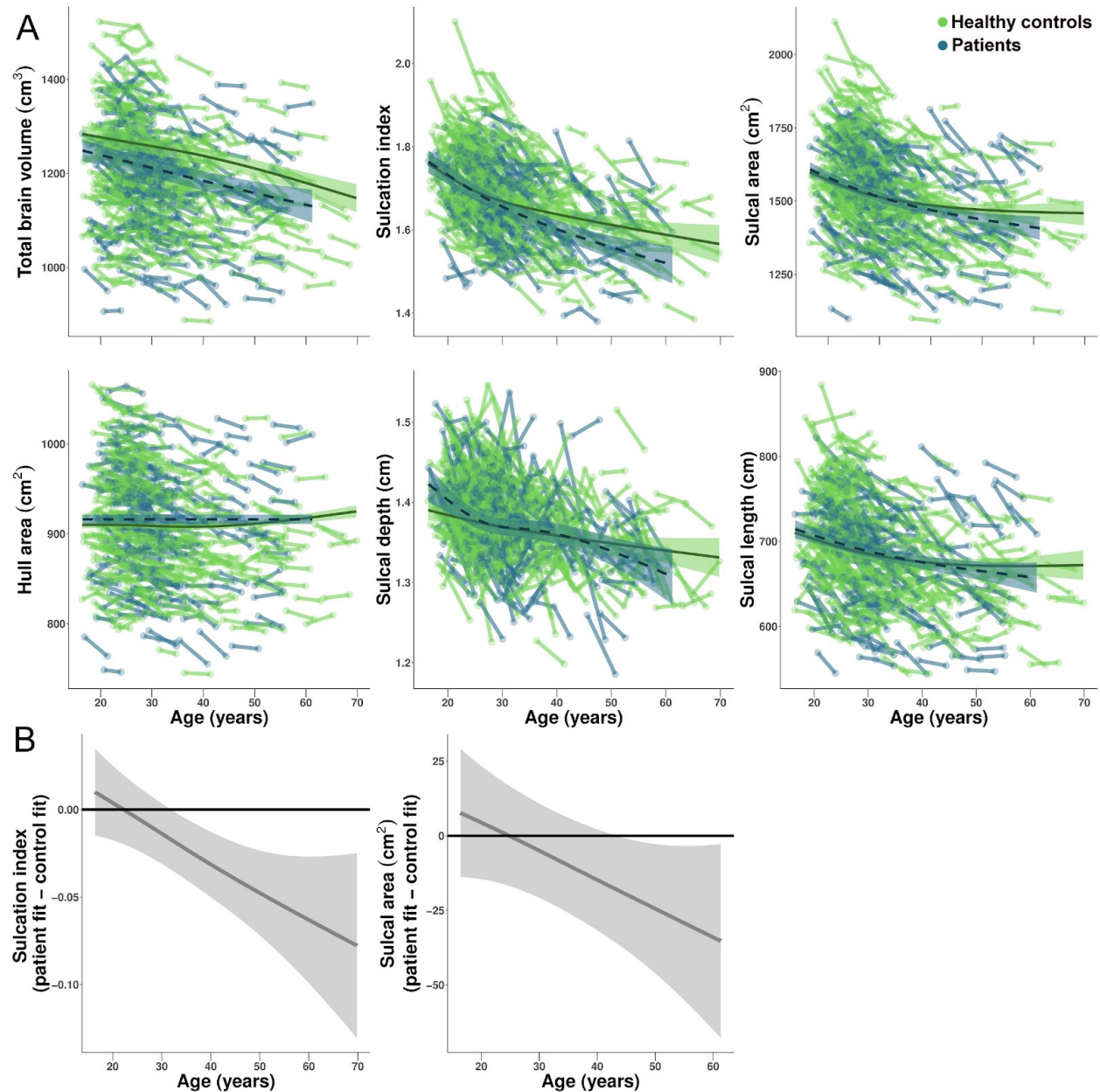
Longitudinal allometric analysis: the relationship between annual change in fold metrics and annual change in brain volume in SZ and healthy individuals

The APC of the hull area showed the strongest scaling with the APC of TBV, with no significant differences between the patient and control groups (see Figure 4). In the patient group, the scaling coefficient for the APC of SI and the APC of total brain volume

(0.60) was similar to the scaling coefficient for the APC of sulcal area (0.63) and the

APC of total brain volume. The patient group showed a significantly stronger scaling,

i.e., steeper slope for change in SI and sulcal area compared to the HI group (both p's

<0.001). There were no significant relationships for APCs of sulcal depth and length.

Figure 4. Orthogonal regression fits ± 95% confidence interval and density maps for observed longitudinal allometric scaling of annual percent change (APC) of the folding metrics relative to the APC of total brain volume in the healthy individuals (HI) and the schizophrenia patients in sample 1. The formula for calculating whether the scaling coefficient differed between the HI and patient groups is given. ***=p<0.001 denotes significance after Bonferroni correction for the difference in scaling between the HI group and the patient group. APCs were estimated from values residualized for age, sex, and scanner.
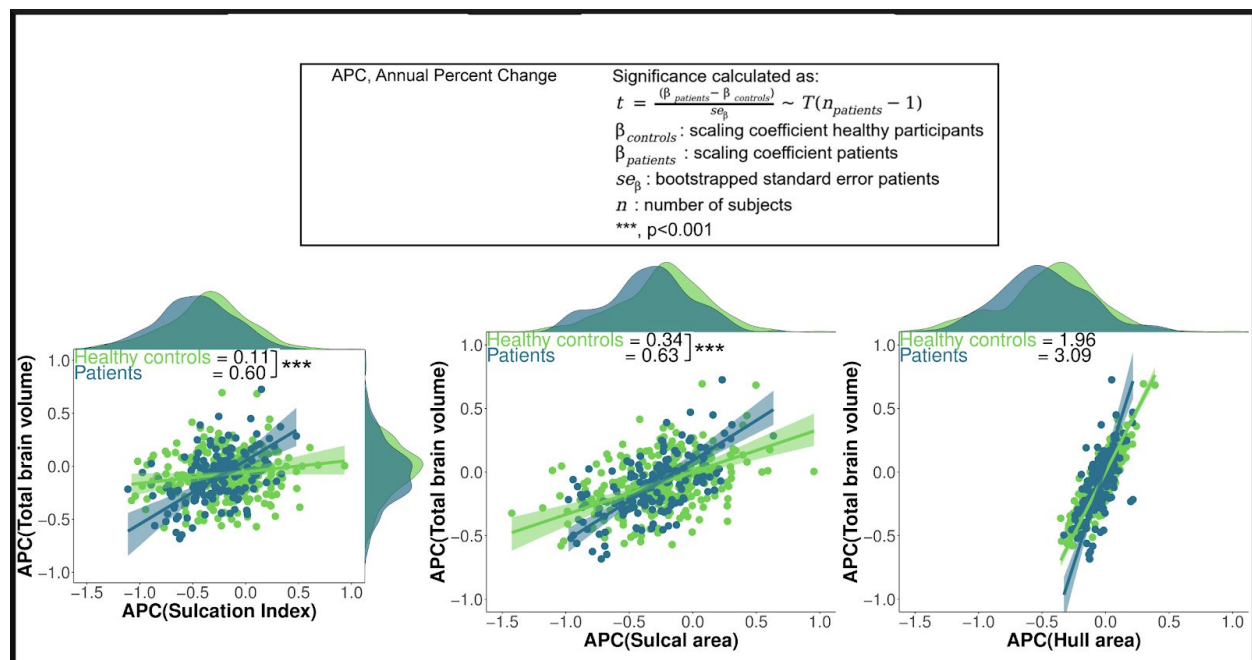
Discussion

Cross-sectional scaling of the Sulcation Index (SI), sulcal area, and sulcal length with total brain volume (TBV) was significantly higher than expected, i.e. hyperallometric, in healthy individuals (HI), individuals with schizophrenia (SZ) and those with first-episode psychosis (FEP). There were no diagnostic differences in scaling. Longitudinal analyses demonstrated steeper loss of SI and sulcal area through adulthood and beyond in SZ. Finally, we assessed the longitudinal allometric scaling between the annual percent change (APC) of the folding metrics and the APC of TBV. The SZ group showed a significantly stronger scaling between the APC in SI and sulcal area with the APC of TBV compared to the HI group. This finding complemented our previous finding indicating that TBV changes over time are more strongly associated with changes in SI and sulcal area in SZ compared to HI.

Allometric analyses at baseline: sample 1 and 2

Normative scaling coefficients in healthy individuals

The hyperallometric scaling of the SI with TBV was driven by a disproportionate expansion of sulcal area, with increasing TBV accompanied by a proportionate expansion of hull area. Allometric analysis of the sulcal depth and length revealed that disproportionate expansion of sulcal area with increasing TBV arises through a hyperallometric scaling relationship between TBV and sulcal length. The disproportionate hyperallometric scaling between TBV and sulcal length is striking and may be explained by an association between increased TBV and twistier sulci (Germanaud et al., 2012). Whether increased sulcal length in larger brains is also due

to the presence of more ('new') sulci is difficult to assess. Currently, no common unified classification framework for sulci exists making it challenging to determine whether a sulcus is 'new' or not (Ono et al., 1990; Mangin et al., 2010; Mikhael et al., 2018). With respect to the underlying neurobiological mechanisms, it remains unclear what produces the hyperallometric scaling of TBV with sulcal surface area. Possible explanations include an expansion of the progenitor cell pool size producing an increased number of neurons during early-life neurodevelopment and/or as lineage regulation of the different types of cortical progenitors which mediate the tangential dispersion of radially migrating neurons (Rakic, 2009; Ronan et al., 2014; Llinares-Benadero and Borrell, 2019). The robustness of our normative allometric findings are evidenced by the replication in an independent sample and replicating Fish et al. (2017).

Comparison of baseline allometric scaling coefficients between healthy individuals and SZ (sample 1) and FEP (sample 2)

Scaling relationships in FEP and SZ are in line with what is expected from the corresponding normative samples. Thus, disproportionate allometric scaling between TBV and SI, sulcal area, and length was also present in individuals with SZ and FEP. In other words, diagnostic differences in global sulcation and its determinants are proportional to the overall effect of diagnosis on TBV, with no added effect of diagnosis per se on these metrics. This suggests that while a diagnosis of SZ or FEP is associated with decreased TBV, global aspects of cortical folding such as SI and its subcomponents are not influenced by the illness. Interestingly, results diverged when using classical methods for controlling for TBV (i.e. normalization and covarying) which

do not account for nonlinear relationships between the folding metrics and TBV indicating that accounting for allometry matters when comparing global metrics of cortical folding between HI and individuals with SZ or FEP.

Longitudinal age trajectories in SZ and healthy individuals (sample 1)

Our longitudinal analyses, to the best of our knowledge, provide the first evidence for diverging lifespan trajectories of SI and sulcal area in SZ when compared to a HI group. After adjusting analyses for the effect of potential non-linear relationships between cortical folding components and TBV on age trajectories, the SZ group demonstrated a steeper SI decline from approximately 32 years onwards and a steeper sulcal area decline from approximately 44 years onwards when compared to the HI group. Our results are in line with increasing *in vivo* evidence for sulcal abnormalities associated with SZ. The amount of cortical surface area depends on cortical curvature. In SZ, sulcal curvature may be disproportionately decreased relative to gyral curvature (Wagstyl et al., 2016). In addition, age-dependent decreases in cortical thickness are more present in sulci than in gyri in HI and this effect may be more pronounced in SZ (Vandekar et al., 2015; Wagstyl et al., 2016). Supragranular layers (layers I-III) are thicker in sulci and thinner in gyri and thus possibly more vulnerable to excessive cortical thinning in SZ (Wagstyl et al., 2016); finally, there is consistent functional as well as post-mortem evidence of supragranular layer pathology in SZ (Harrison and Weinberger, 2005; Williams and Boksa, 2010). Changes in hull surface area may be more sensitive to changes in gyri as compared to sulci and there were no differences in the age trajectories of hull surface area between the SZ and HI groups. Taken together,

our results point to abnormal global sulcal changes over time in the SZ group. This may be explained by a disproportionate demyelination in SZ (Mighdoll et al., 2015; Kochunov et al., 2016; Kelly et al., 2018). Tertiary sulci are the latest sulci to develop postnatally and their development is associated with cortical expansion and myelination (Miller et al., 2021). Aging and its associated demyelination may reverse this process, decreasing the prominence of the tertiary sulci leading to a reduction in sulcal area and SI. Nevertheless, future longitudinal studies assessing gyral- and sulcal-specific morphology are necessary to determine whether aberrant morphological changes over time are excessive in sulci.

Longitudinal allometric analysis: the relationship between annual change in fold metrics and annual change in brain volume in SZ and healthy individuals

This study is the first to show the scaling between annual percent change in subcomponents of SI and change in TBV in HI and individuals with SZ. We do not interpret the reported scaling coefficients from this longitudinal analysis as being hypo-, iso-, or hyperallometric as there are, to the best of our knowledge, no prior studies presenting normative longitudinal scaling coefficients for folding subcomponents have been published. Scaling coefficients for SI and sulcal area were higher in the SZ group compared to the HI group, i.e. TBV changes over time are more strongly associated with changes in SI and sulcal area in SZ compared to HI. This complements our findings of the group differences in age trajectories. Loss of SI and sulcal area over time was higher in SZ and this loss has an atypical strong association with change in TBV. As such it may be that loss of SI and sulcal area, rather than hull area, sulcal depth and length, are important contributors to the disproportionate loss of TBV over time in SZ.

A number of limitations should be taken into account when interpreting these findings. In sample 1 antipsychotics usage information was heterogeneous and unfortunately limited. Thus, the obtained correlations between measures of chlorpromazine equivalents and brain metrics reported in this study are approximate and the results must be interpreted cautiously. Recent work demonstrated a power law relation between average global cortical thickness, total exposed cortical surface area and total cortical surface area which describes the degree of folding (Wang et al., 2016). Whether this power law also predicts the relationship between the metrics used in the current study should be addressed in future studies. We assessed global but not regional (e.g. at the level of the individual sulcus) folding. By examining global aspects of cortical folding we address brain-wide geometrical organization and as such provide a building block for future studies examining more local characteristics of folding.

References

Alemán-Gómez Y, Janssen J, Schnack H, Balaban E, Pina-Camacho L, Alfaro-Almagro F, Castro-Fornieles J, Otero S, Baeza I, Moreno D, Bargalló N, Parellada M, Arango C, Desco M (2013) The human cerebral cortex flattens during adolescence. J Neurosci Off J Soc Neurosci 33:15004–15010.

Barnes J, Ridgway GR, Bartlett J, Henley SMD, Lehmann M, Hobbs N, Clarkson MJ, MacManus DG, Ourselin S, Fox NC (2010) Head size, age and gender adjustment in MRI studies: a necessary nuisance? NeuroImage 53:1244–1255.

Bernardo M, Bioque M, Parellada M, Saiz Ruiz J, Cuesta MJ, Llerena A, Sanjuán J, Castro-Fornieles J, Arango C, Cabrera B, PEPs Group (2013) Assessing clinical and functional outcomes in a gene-environment interaction study in first episode of psychosis (PEPs). Rev Psiquiatr Salud Ment 6:4–16.

Cachia A, Amad A, Brunelin J, Krebs M-O, Plaze M, Thomas P, Jardri R (2015) Deviations in cortex sulcation associated with visual hallucinations in schizophrenia. Mol Psychiatry 20:1101–1107.

Cao B, Mwangi B, Passos IC, Wu M-J, Keser Z, Zunta-Soares GB, Xu D, Hasan KM, Soares JC (2017) Lifespan Gyrification Trajectories of Human Brain in Healthy Individuals and Patients with Major Psychiatric Disorders. Sci Rep 7:511.

Chambers JM, Hastie TJ (1992) Statistical models. S. Pacific Grove. California, USA: Wadsworth & Brooks/Cole.

Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis. I. Segmentation and surface reconstruction. NeuroImage 9:179–194.

Dazzan P et al. (2020) Symptom Remission and Brain Cortical Networks at First Clinical

Presentation of Psychosis: The OPTiMiSE Study. Schizophr Bull.

de Jong LW, Vidal J-S, Forsberg LE, Zijdenbos AP, Haight T, Alzheimer's Disease Neuroimaging Initiative, Sigurdsson S, Gudnason V, van Buchem MA, Launer LJ (2017) Allometric scaling of brain regions to intra-cranial volume: An epidemiological MRI study. Hum Brain Mapp 38:151–164.

Efron B, Tibshirani RJ (1993) An Introduction to the Bootstrap. New York: Chapman and Hall.

Fischl B (2012) FreeSurfer. NeuroImage 62:774–781.

Fischl B, Sereno MI, Dale AM (1999) Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. NeuroImage 9:195–207.

Fish AM, Cachia A, Fischer C, Mankiw C, Reardon PK, Clasen LS, Blumenthal JD, Greenstein D, Giedd JN, Mangin J-F, Raznahan A (2017) Influences of Brain Size, Sex, and Sex Chromosome Complement on the Architecture of Human Cortical Folding. Cereb Cortex N Y N 1991 27:5557–5567.

Gay O, Plaze M, Oppenheim C, Gaillard R, Olié J-P, Krebs M-O, Cachia A (2017) Cognitive control deficit in patients with first-episode schizophrenia is associated with complex deviations of early brain development. J Psychiatry Neurosci JPN 42:87–94.

Germanaud D, Lefèvre J, Toro R, Fischer C, Dubois J, Hertz-Pannier L, Mangin J-F (2012) Larger is twistier: spectral analysis of gyrification (SPANGY) applied to adult brain size polymorphism. NeuroImage 63:1257–1272.

Haijma SV, Van Haren N, Cahn W, Koolschijn PCMP, Hulshoff Pol HE, Kahn RS (2013) Brain volumes in schizophrenia: a meta-analysis in over 18 000 subjects. Schizophr Bull 39:1129–1138.

Harrison PJ, Weinberger DR (2005) Schizophrenia genes, gene expression, and neuropathology: on the matter of their convergence. Mol Psychiatry 10:40–68; image 5.

Hulshoff Pol HE, Schnack HG, Mandl RC, van Haren NE, Koning H, Collins DL, Evans AC, Kahn RS (2001) Focal gray matter density changes in schizophrenia. Arch Gen Psychiatry 58:1118–1125.

Im K, Lee J-M, Seo SW, Hyung Kim S, Kim SI, Na DL (2008) Sulcal morphology changes and their relationship with cortical thickness and gyral white matter volume in mild cognitive impairment and Alzheimer's disease. NeuroImage 43:103–113.

Jäncke L, Liem F, Merillat S (2019) Scaling of brain compartments to brain size. Neuroreport 30:573–579.

Janssen J, Alemán-Gómez Y, Schnack H, Balaban E, Pina-Camacho L, Alfaro-Almagro F, Castro-Fornieles J, Otero S, Baeza I, Moreno D, Bargalló N, Parellada M, Arango C, Desco M (2014) Cortical morphology of adolescents with bipolar disorder and with schizophrenia. Schizophr Res 158:91–99.

Janssen J, Díaz-Caneja CM, Alloza C, Schippers A, de Hoyos L, Santonja J, Gordaliza PM, Buimer EEL, van Haren NEM, Cahn W, Arango C, Kahn RS, Hulshoff Pol HE, Schnack HG (2020) Dissimilarity in Sulcal Width Patterns in the Cortex can be Used to Identify Patients With Schizophrenia With Extreme Deficits in Cognitive Performance. Schizophr Bull.

Kelly S et al. (2018) Widespread white matter microstructural differences in schizophrenia across 4322 individuals: results from the ENIGMA Schizophrenia DTI Working Group. Mol Psychiatry 23:1261–1269.

Klein A, Ghosh SS, Bao FS, Giard J, Häme Y, Stavsky E, Lee N, Rossa B, Reuter M,

Chaibub Neto E, Keshavan A (2017) Mindboggling morphometry of human brains. PLoS Comput Biol 13:e1005350.

Kochunov P, Rowland LM, Fieremans E, Veraart J, Jahanshad N, Eskandar G, Du X, Muellerklein F, Savransky A, Shukla D, Sampath H, Thompson PM, Hong LE (2016) Diffusion-weighted imaging uncovers likely sources of processing-speed deficits in schizophrenia. Proc Natl Acad Sci U S A 113:13504–13509.

Korver N, Quee PJ, Boos HBM, Simons CJP, de Haan L, GROUP investigators (2012) Genetic Risk and Outcome of Psychosis (GROUP), a multi-site longitudinal cohort study focused on gene-environment interaction: objectives, sample characteristics, recruitment and assessment methods. Int J Methods Psychiatr Res 21:205–221.

Kubota M, van Haren NEM, Haijma SV, Schnack HG, Cahn W, Hulshoff Pol HE, Kahn RS (2015) Association of IQ Changes and Progressive Brain Changes in Patients With Schizophrenia. JAMA Psychiatry 72:803–812.

Kuperberg GR, Broome MR, McGuire PK, David AS, Eddy M, Ozawa F, Goff D, West WC, Williams SCR, van der Kouwe AJW, Salat DH, Dale AM, Fischl B (2003) Regionally localized thinning of the cerebral cortex in schizophrenia. Arch Gen Psychiatry 60:878–888.

Lavoie S, Bartholomeuz CF, Nelson B, Lin A, McGorry PD, Velakoulis D, Whittle SL, Yung AR, Pantelis C, Wood SJ (2014) Sulcogyral pattern and sulcal count of the orbitofrontal cortex in individuals at ultra high risk for psychosis. Schizophr Res 154:93–99.

Li G, Wang L, Shi F, Lyall AE, Lin W, Gilmore JH, Shen D (2014) Mapping longitudinal development of local cortical gyrification in infants from birth to 2 years of age. J

Neurosci Off J Soc Neurosci 34:4228–4238.

Llinares-Benadero C, Borrell V (2019) Deconstructing cortical folding: genetic, cellular and mechanical determinants. Nat Rev Neurosci 20:161–176.

MacKinley ML, Sabesan P, Palaniyappan L (2019) Deviant cortical sulcation related to schizophrenia, but not cognitive deficits, likely predate brain development in the second trimester. bioRxiv:868133.

Madansky A (1959) The Fitting of Straight Lines When both Variables are Subject to Error. J Am Stat Assoc 54:173–205.

Mangin J-F, Jouvent E, Cachia A (2010) In-vivo measurement of cortical morphology: means and meanings. Curr Opin Neurol 23:359–367.

Mangin J-F, Rivière D, Cachia A, Duchesnay E, Cointepas Y, Papadopoulos-Orfanos D, Scifo P, Ochiai T, Brunelle F, Régis J (2004) A framework to study the cortical folding patterns. NeuroImage 23 Suppl 1:S129-138.

Mankiw C, Park MTM, Reardon PK, Fish AM, Clasen LS, Greenstein D, Giedd JN, Blumenthal JD, Lerch JP, Chakravarty MM, Raznahan A (2017) Allometric Analysis Detects Brain Size-Independent Effects of Sex and Sex Chromosome Complement on Human Cerebellar Organization. J Neurosci Off J Soc Neurosci 37:5221–5231.

Mighdoll MI, Tao R, Kleinman JE, Hyde TM (2015) Myelin, myelin-related disorders, and psychosis. Schizophr Res 161:85–93.

Mikhael S, Hoogendoorn C, Valdes-Hernandez M, Pernet C (2018) A critical analysis of neuroanatomical software protocols reveals clinically relevant differences in parcellation schemes. NeuroImage 170:348–364.

Miller JA, Voorhies WI, Lurie DJ, D'Esposito M, Weiner KS (2021) Overlooked Tertiary

Sulci Serve as a Meso-Scale Link between Microstructural and Functional Properties of Human Lateral Prefrontal Cortex. J Neurosci Off J Soc Neurosci 41:2229–2244.

Mirakhur A, Moorhead TWJ, Stanfield AC, McKirdy J, Sussmann JED, Hall J, Lawrie SM, Johnstone EC, McIntosh AM (2009) Changes in gyrification over 4 years in bipolar disorder and their association with the brain-derived neurotrophic factor valine(66) methionine variant. Biol Psychiatry 66:293–297.

Mutlu AK, Schneider M, Debbané M, Badoud D, Eliez S, Schaer M (2013) Sex differences in thickness, and folding developments throughout the cortex. NeuroImage 82:200–207.

Nanda P, Tandon N, Mathew IT, Giakoumatos CI, Abhishekh HA, Clementz BA, Pearlson GD, Sweeney J, Tamminga CA, Keshavan MS (2014) Local gyrification index in probands with psychotic disorders and their first-degree relatives. Biol Psychiatry 76:447–455.

Nesvåg R, Schaer M, Haukvik UK, Westlye LT, Rimol LM, Lange EH, Hartberg CB, Ottet M-C, Melle I, Andreassen OA, Jönsson EG, Agartz I, Eliez S (2014) Reduced brain cortical folding in schizophrenia revealed in two independent samples. Schizophr Res 152:333–338.

Ono M, Kubik S, Abernathey CD (1990) Atlas of the Cerebral Sulci. Stuttgart, New York: Georg Thieme Verlag.

Palaniyappan L, Mallikarjun P, Joseph V, White TP, Liddle PF (2011) Folding of the prefrontal cortex in schizophrenia: regional differences in gyrification. Biol Psychiatry 69:974–979.

Penttilä J, Cachia A, Martinot J-L, Ringuenet D, Wessa M, Houenou J, Galinowski A,

Bellivier F, Gallarda T, Duchesnay E, Artiges E, Leboyer M, Olié J-P, Mangin J-F, Paillère-Martinot M-L (2009) Cortical folding difference between patients with early-onset and patients with intermediate-onset bipolar disorder. Bipolar Disord 11:361–370.

Pina-Camacho L, Del Rey-Mejías Á, Janssen J, Bioque M, González-Pinto A, Arango C, Lobo A, Sarró S, Desco M, Sanjuan J, Lacalle-Aurioles M, Cuesta MJ, Saiz-Ruiz J, Bernardo M, Parellada M, PEPs Group (2016) Age at First Episode Modulates Diagnosis-Related Structural Brain Abnormalities in Psychosis. Schizophr Bull 42:344–357.

Pizzagalli F, Auzias G, Yang Q, Mathias SR, Faskowitz J, Boyd JD, Amini A, Rivière D, McMahon KL, de Zubicaray GI, Martin NG, Mangin J-F, Glahn DC, Blangero J, Wright MJ, Thompson PM, Kochunov P, Jahanshad N (2020) The reliability and heritability of cortical folds and their genetic correlations across hemispheres. Commun Biol 3:510.

Rakic P (2009) Evolution of the neocortex: a perspective from developmental biology. Nat Rev Neurosci 10:724–735.

Raznahan A, Shaw P, Lalonde F, Stockman M, Wallace GL, Greenstein D, Clasen L, Gogtay N, Giedd JN (2011) How does your cortex grow? J Neurosci Off J Soc Neurosci 31:7174–7177.

Reardon PK, Clasen L, Giedd JN, Blumenthal J, Lerch JP, Chakravarty MM, Raznahan A (2016) An Allometric Analysis of Sex and Sex Chromosome Dosage Effects on Subcortical Anatomy in Humans. J Neurosci Off J Soc Neurosci 36:2438–2448.

Ronan L, Voets N, Rua C, Alexander-Bloch A, Hough M, Mackay C, Crow TJ, James A, Giedd JN, Fletcher PC (2014) Differential tangential expansion as a mechanism for

cortical gyrification. Cereb Cortex N Y N 1991 24:2219–2228.

Rosas HD, Liu AK, Hersch S, Glessner M, Ferrante RJ, Salat DH, van der Kouwe A, Jenkins BG, Dale AM, Fischl B (2002) Regional and progressive thinning of the cortical ribbon in Huntington's disease. Neurology 58:695–701.

Shimony JS, Smyser CD, Wideman G, Alexopoulos D, Hill J, Harwell J, Dierker D, Van Essen DC, Inder TE, Neil JJ (2016) Comparison of cortical folding measures for evaluation of developing human brain. NeuroImage 125:780–790.

Toro R, Perron M, Pike B, Richer L, Veillette S, Pausova Z, Paus T (2008) Brain size and folding of the human cerebral cortex. Cereb Cortex N Y N 1991 18:2352–2357.

Vandekar SN, Shinohara RT, Raznahan A, Roalf DR, Ross M, DeLeo N, Ruparel K, Verma R, Wolf DH, Gur RC, Gur RE, Satterthwaite TD (2015) Topologically dissociable patterns of development of the human cerebral cortex. J Neurosci Off J Soc Neurosci 35:599–609.

Wagstyl K, Ronan L, Whitaker KJ, Goodyer IM, Roberts N, Crow TJ, Fletcher PC (2016) Multiple markers of cortical morphology reveal evidence of supragranular thinning in schizophrenia. Transl Psychiatry 6:e780.

Wang Y, Necus J, Kaiser M, Mota B (2016) Universality in human cortical folding in health and disease. Proc Natl Acad Sci U S A 113:12820–12825.

Williams S, Boksa P (2010) Gamma oscillations and schizophrenia. J Psychiatry Neurosci JPN 35:75–77.

Wood, SN (2006) Generalized additive models: an introduction with R. Chapman Hall/CRC.

Supplement

Longitudinal allometrics of sulcal morphology in health and schizophrenia

Authors: Joost Janssen[1,2,3,4,*], Clara Alloza[1,2,3], Covadonga M. Díaz-Caneja[1,2,3,5], Javier Santonja[1,2], Laura Pina-Camacho[1,2,3,5], Pedro M. Gordaliza[6], Alberto Fernández-Pena[2,6,], Noemi González Lois[1,2], Elizabeth E.L. Buimer[4], Neeltje E.M. van Haren[4,7], Wiepke Cahn[4], Eduardo Vieta[3,9], Josefina Castro-Fornieles[3,10], Miquel Bernardo[3,8], Mara Parellada[1,2,3,5], Celso Arango[1,2,3,5], René S. Kahn[4,11], Hilleke E. Hulshoff Pol[4], Hugo G. Schnack[4]

[1]Department of Child and Adolescent Psychiatry, Institute of Psychiatry and Mental Health, Hospital General Universitario Gregorio Marañón, Madrid, Spain.
[2]Instituto de Investigación Sanitaria Gregorio Marañón (IiSGM), Madrid, Spain.
[3]Ciber del Área de Salud Mental (CIBERSAM).
[4]Department of Psychiatry, UMCU Brain Center, University Medical Center Utrecht, Utrecht, The Netherlands.
[5]School of Medicine, Universidad Complutense, Madrid, Spain.
[6]Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, Madrid, Spain.
[7]Department of Child and Adolescent Psychiatry/Psychology, Erasmus University Medical Centre, Sophia Children's Hospital, Rotterdam, The Netherlands.
[8]Barcelona Clinic Schizophrenia Unit, Hospital Clinic of Barcelona, Neuroscience Institute, IDIBAPS, University of Barcelona, Barcelona, Spain.
[9]Bipolar Disorders Unit, Clinical Institute of Neurosciences, Hospital Clínic, University of Barcelona, IDIBAPS, Barcelona, Spain.
[10]Department of Child and Adolescent Psychiatry and Psychology, Clinical Institute of Neurosciences, Hospital Clínic, University of Barcelona, IDIBAPS, Barcelona, Spain
[11]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, United States

Table of contents

Image quality assurance

Sample 1

After the completion of the preprocessing pipeline for all T1 MRI scans in FreeSurfer, we rigorously assessed the quality of the data using a combination of visual inspection and the examination of several quantitative metrics following recent recommendations from the literature (Rosen et al., 2018). First, we calculated five quality measurements based on the proposed by the Preprocessed Connectome Project (http://preprocessed-connectomes-project.org/quality-assessment-protocol/) to identify images that were unusable: signal-to-noise ratio (SNR), contrast to noise ratio (CNR), foreground to background Energy Ratio (FBER) and percent artefact Voxels (Artifacts) and Entropy Focus Criterion (Entropy). As in (Vandekar et al., 2015) we defined the threshold for outliers as [mean-(2.698*SD)] for SNR, CNR and FBER metrics; and Artifacts and Entropy were only tested for [mean-(2.698*SD)] based on the exclusion criteria explained by ENIGMA´s QA protocol. Next, for each scan the whole brain mean cortical thickness, total surface area, total white matter volume, total grey matter volume, subcortical grey matter volume and intracranial volume were calculated. Then, we summed for each FreeSurfer directory the amount of ENIGMA outliers over all measures and calculated the mean+(2.698*SD) for the amount of outliers over the whole sample and designate those above the threshold as outliers. These quality control steps resulted in the below exclusions.

Following the exclusion of participants that met our exclusion criteria (585 scans) a sample of 1239 scans was obtained. After the evaluation of the computed quantitative measures of image quality, 152 scans were considered of insufficient quality and were

excluded from our analysis (see SFigure 1A). These excluded scans were validated by manually assessing image quality after preprocessing to assure that all preprocessing steps worked and to avoid the propagation of error along the analysis. The parameters that were most useful for objectively detect artefacts visually were agreed on between researchers, e.g. incorrect sulcal labelling or insufficient quality of sulcal segmentation resulting in gross anatomical abnormalities. These visual checks were performed in BrainVisa. Those scans with poor initial sulcal labelling were manually edited, which often resulted in better sulcal labelling when manually inspected (see SFigure 1B for examples of "good", "medium" and "bad" quality sulcal labelling after first preprocessing of the images). In addition, 14 scans were excluded after this manual inspection.

Sample 2

QA information for sample 2 can be found in (Pina-Camacho et al., 2016). Briefly, manual edits, mainly for errors in skull-stripping, were made. Thereafter, no outliers for FreeSurfer processed images were detected. However, manual inspection of the BrainVisa (see above) output led to the exclusion of 32 participants.

Sulcation Index and Gyrification Index

The sulcation index (SI) is the ratio between total sulcal surface area and hull area:

$$SI_h = \frac{\sum_{s=0}^{S} A_s}{A_h^{hull}}, \quad s \in h$$

where h yields for hemisphere, s for sulcus, and A is area, being A_h^{hull} the area of the brain hull for hemisphere h. The SI differs from the gyrification index (GI) (Zilles et al., 1988)which is calculated as:

$$GI_h = \frac{A_h}{A_h^{hull}}$$

where h yields for hemisphere, and A is area, being A_h^{hull} the area of the cortex for hemisphere h. We made a diagram which gives additional information about the differences between SI and GI:

Sulcation Index	Gyrification Index
$$SI_h = \frac{\sum_{s=0}^{S} A_s}{A_h^{hull}}, \; s \in h$$	$$GI_h = \frac{A_h}{A_h^{hull}}$$
• Indexes degree of sulcation • Sensitive to sulcal contour metrics: sulcal depth and length • Higher ratio signifies increased sulcation	• Indexes degree of gyrification • Sensitive to gyral contour measures: gyral length, cortical thickness • Higher ratio signifies increased sulcation

The SI's particular sensitivity to global changes in sulcal morphology as opposed to changes in gyri was the reason for assessing SI. A cortex with extensive folding has a large *SI*, whereas a cortex with low degree of folding has a small *SI*. At a constant outer cortex area, the *SI* increases with the number and area of sulcal folds, whereas the *SI* of a lissencephalic cortex is zero.

SFigure 1. **A**) Flowchart summarizing the exclusion criteria process: from an initial sample of 1824 total scans, a total of 737 scans were excluded for subsequent analysis. The final sample comprises 1087 scans: a. indicates the number of included scans; b. the number of excluded scans and c. the exclusion criteria (initial and imaging criteria) applied to the sample. **B**) Illustrates examples of sulcal segmentation of different qualities performed with Brainvisa: good quality of sulcal reconstruction, medium quality of sulcal reconstruction, included after manual sulcal editing; and bad quality scan which shows a lump caused by incorrect sulcal reconstruction which caused exclusion.
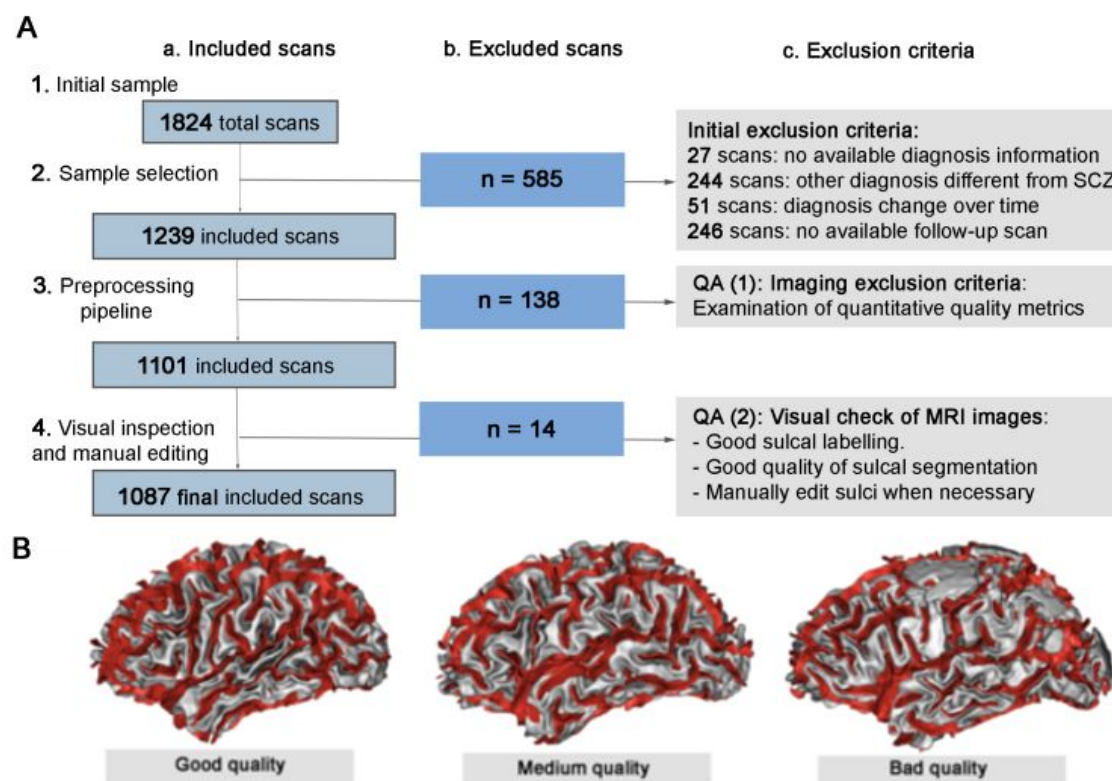
SFigure 2. Age at scans for all participants from sample 1. Each of the 1087 scans obtained is represented by a dot; each of the 466 participants is shown in a different row, with their two or three scans connected by a straight grey line. Healthy individuals and patients are marked separately.
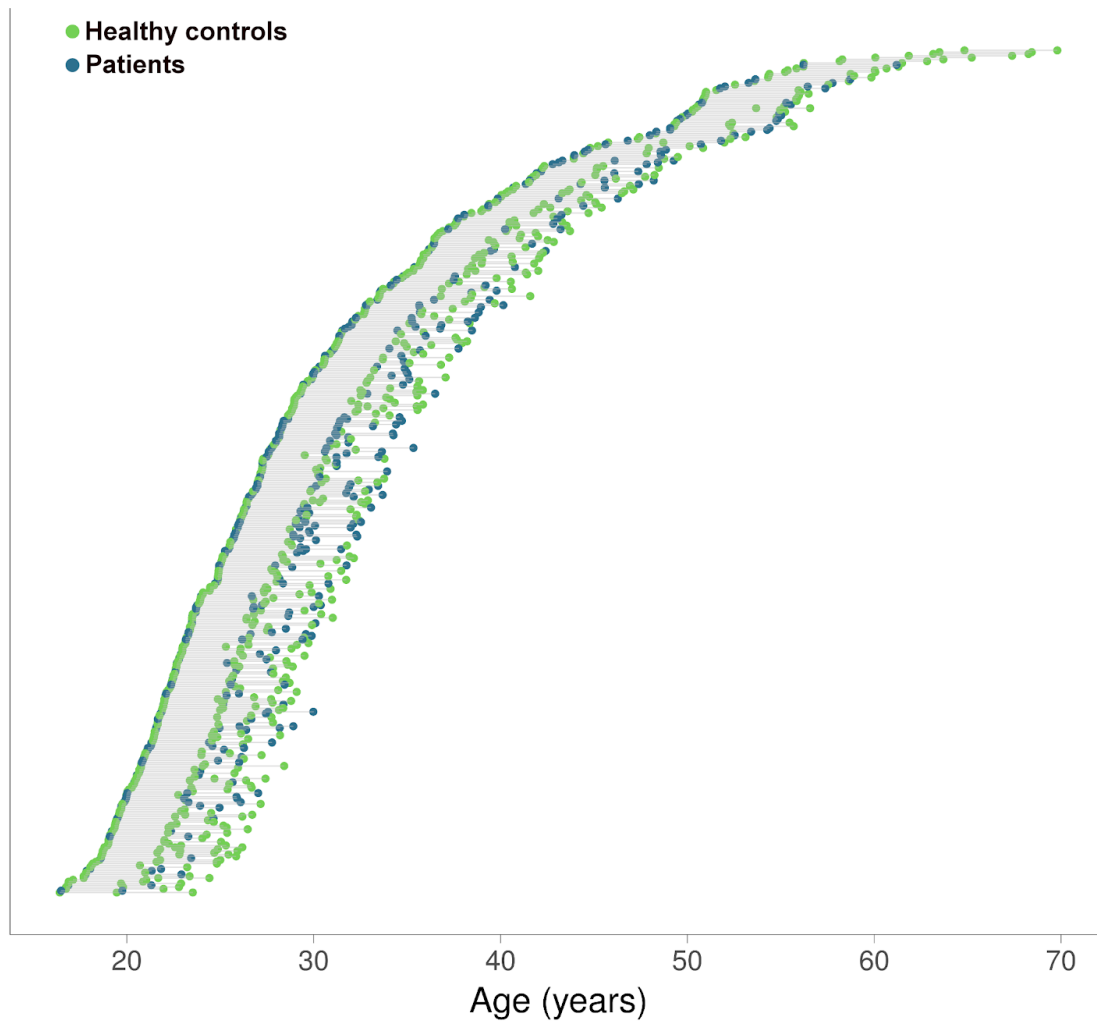
SFigure 3. Regression plots for patients from sample 1 and 2 showing the relationship between the brain measures at baseline and cumulative antipsychotics intake, illness duration, and estimated IQ. Chlorpromazine equivalents and illness duration were first normalized by selecting the optimal transformation method, based on the lowest Pearson p test statistic for normality, among the Yeo-Johnson, exp(x), log10, square-root, arcsinh and orderNorm transformations (bestNormalize package, R). Participants missing data in sample 1: cumulative antipsychotic usage (103), illness duration (60) and estimated IQ (44). Pearson correlations for each pair of variables were not significant after Bonferroni correction across samples. CPZ, chlorpromazine equivalents; IQ, intelligence quotient. CPZ, chlorpromazine; IQ, intelligence quotient.
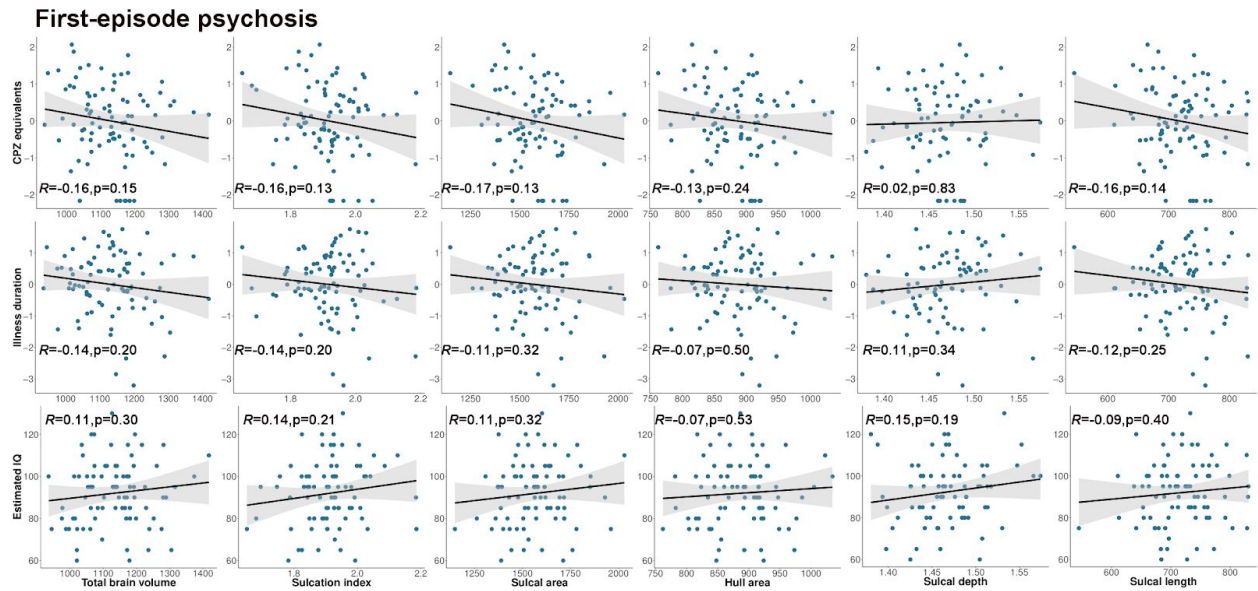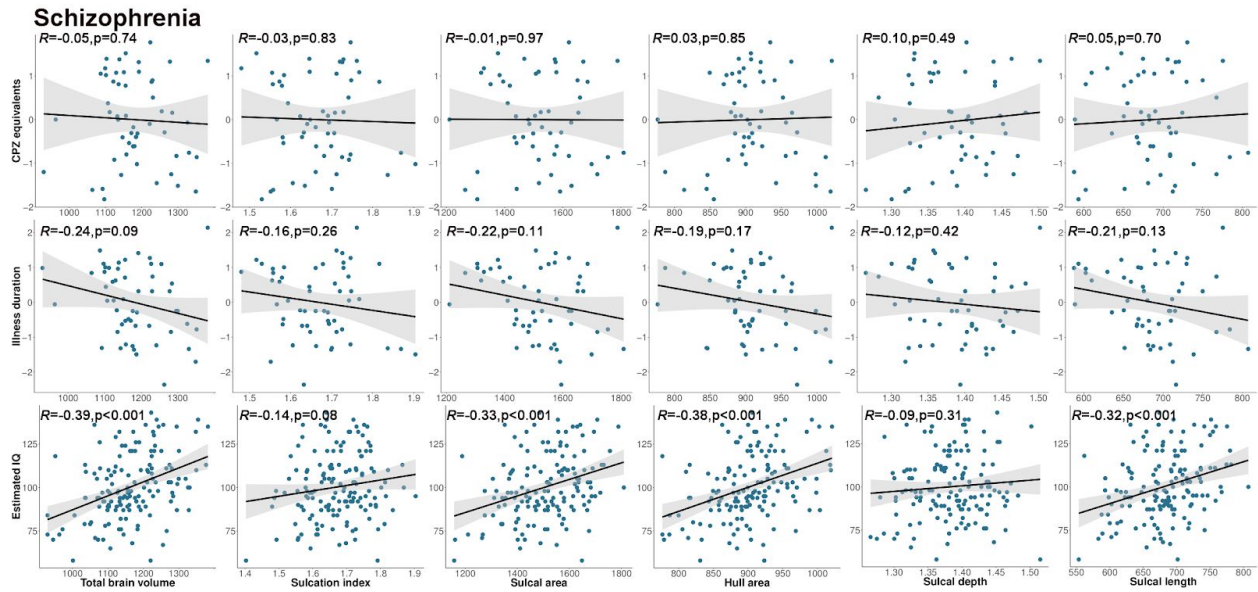
Supplemental references

Pina-Camacho L, Del Rey-Mejías Á, Janssen J, Bioque M, González-Pinto A, Arango C, Lobo A, Sarró S, Desco M, Sanjuan J, Lacalle-Aurioles M, Cuesta MJ, Saiz-Ruiz J, Bernardo M, Parellada M, PEPs Group (2016) Age at First Episode Modulates Diagnosis-Related Structural Brain Abnormalities in Psychosis. Schizophr Bull 42:344–357.

Rosen AFG et al. (2018) Quantitative assessment of structural image quality. NeuroImage 169:407–418.

Vandekar SN, Shinohara RT, Raznahan A, Roalf DR, Ross M, DeLeo N, Ruparel K, Verma R, Wolf DH, Gur RC, Gur RE, Satterthwaite TD (2015) Topologically dissociable patterns of development of the human cerebral cortex. J Neurosci Off J Soc Neurosci 35:599–609.

Zilles K, Armstrong E, Schleicher A, Kretschmann HJ (1988) The human pattern of gyrification in the cerebral cortex. Anat Embryol (Berl) 179:173–179.