Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____	6/28/2021
Yiheng Li	Date

Characterizing the Developmental Trajectories for Cerebellar Volumes

via Penalized Splines and Derivative Analysis

By

Yiheng Li

Master of Science

Biostatistics

Ying Guo, Ph.D.

Advisor

Longchuan Li, Ph.D.

Co - Advisor

Zhaohui Qin, Ph.D.

Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.

Dean of the James T. Laney School of Graduate Studies

Date

Characterizing the Developmental Trajectories for Cerebellar Volumes via Penalized Splines and Derivative Analysis

By

Yiheng Li

B.S., Zhejiang University, 2019

Advisor: Ying Guo, Ph.D.

Co-Advisor: Longchuan Li, Ph.D.

An abstract of

A thesis submitted to the Faculty of the James T. Laney School of Graduate Studies of Emory University in partial fulfillment of the requirements for the degree of

Master of Science

in Biostatistics

2021

Abstract

Characterizing the Developmental Trajectories for Cerebellar Volumes
via Penalized Splines and Derivative Analysis
By Yiheng Li

The human cerebellum has been established an important role in both motor and cognitive functions. Many of these functions change across age, especially during childhood and adolescence. A representation of cerebellar development is the volume of the cerebellum and of various regions of the cerebellum. Past literature has spotted a non-linear trajectory as well as a gender dimorphism of how the cerebellar volumes develop across age. Specifically in this paper, the focus is on how total cerebellum volume (TCV), cerebellar cortex volume (CCV), cerebellum white-matter volume (CWM) and the total intracranial volume (TICV) changes during childhood, adolescence and adulthood in different gender groups. Cross-sectional volumetric data both with and without TICV adjustment from 1608 healthy subjects ranging from 5 to 37 years old is analyzed stratified on sex. To characterize the developmental trajectories of the cerebellar volumes, a flexible penalized cubic spline modeling framework is adopted, which combines the reduced knots of a regression spline (computational efficiency) and the roughness penalty of a smoothing spline (smoothness control of the fitted splines). Derivatives of the fitted trajectories are then approximated by difference quotients and their variances are estimated based on the posterior Bayesian covariance matrix of the smoothing coefficients. Hypothesis tests of whether the non-zero derivatives are significantly different from 0 are derived, through which the age periods of significant increase/decrease of the cerebellar volumes of interest are identified. The TCV of males experiences an increase in childhood and adolescence and reaches its peak at14 years old with minimal change afterward. The CCV of males follows an inverse-U-shaped curve peaking at 13 years. For both female and male CWM, a steady increase is fitted in the studied age range. The TICV adjustment does not affect the general trend of the development of male TCV and male CCV. For TCV and CCV of females, the predicted trajectory is in the shape of 2 inverse-U connected together with 2 peaks around age 12 and at age 28. However, after adjusting for TICV, age is no longer a significant factor affecting TCV and CCV is predicted to show a sustained volume loss.

Characterizing the Developmental Trajectories for Cerebellar Volumes via Penalized Splines and Derivative Analysis

By

Yiheng Li

B.S., Zhejiang University, 2019

Advisor: Ying Guo, Ph.D.

Co-Advisor: Longchuan Li, Ph.D.

A thesis submitted to the Faculty of the James T. Laney School of Graduate Studies of Emory University in partial fulfillment of the requirements for the degree of

Master of Science

in Biostatistics

2021

Acknowledgement

Thanks to Dr. Ying Guo and Dr. Longchuan Li for their guidance and suggestions on this thesis project. Many thanks to Esra Sefik who is a doctoral student in Neuroscience from the GDBBS at Emory. She helped to collect and sort the publicly available HCP data.

Table of Contents

List of Tables

List of Figures

1 Introduction

Accumulating evidence recognizes the cerebellum as a prominent contributor to both motor and cognitive functions involving learning, language, working memory and more (Leiner et al. (1993); Chen and Desmond (2005); Stoodley et al. (2012)). Besides, the cerebellum has also been implicated in several mental health disorders such as autism (Stanfield et al., 2008) and ADHD (Valera et al., 2007), which makes the cerebellum a more intriguing structure to study.

Abundant literature has concluded that the cerebellum is affected by age and this effect varies across different regions of the cerebellum. For instance, Luft et al. (1999) adopt an exponential model and find that total cerebellum volume loss begins around 50 years old and the degree of loss continues to increase until age 65, after which the decline becomes slower. Tiemeier et al. (2010) fit a mixed model including a cubic term of age to quantify the trend of total cerebellar volume and 11 subdivisions. They find that total cerebellum volume follows an inverse U-shaped trajectory peaking at age 11.8 years in females and 15.6 years in males. The anterior region and total grey matter volume of the cerebellum also reach their peaks during the early teen years and then begin to slowly decrease. While the cerebellar white matter volume remains a steady and almost linear increase throughout childhood and adolescence. Bernard et al. (2015) confirm a quadratic inverse-U pattern for the more posterior part of the cerebellum while the anterior cerebellum follows a logarithmic pattern so as the total cerebellar gray matter volume.

However, most of these studies handle the non-linearity of age by adopting sophisticated parametric models such as exponential, quadratic, cubic, or logarithmic. These delicate assumptions require great expertise to make and are inflexible, especially when parametric models with different levels of complexity are compared in terms of fit. Therefore, more flexible and effective regression techniques are required and penalized splines are one of them (Wahba, 1980; Eilers et al., 1996). They combine the reduced knots of regression splines and the roughness penalty of smoothing splines (Reinsch, 1967; Eilers et al., 1996). For penalized splines, the user-defined knots given a fair coverage can capture the changes in the underlying level quite well (Wood, 2017), and the roughness penalty can control the smoothness of the fit, i.e., avoid overfitting. Penalized splines also enjoy computational efficiency compared with smoothing splines, the latter of which place a knot at every unique covariate point. Penalized splines are especially beneficial when nonlinear effects that are difficult to model parametrically are involved, as in the case of modeling the cerebellar volume trajectories, since the ground truth of the non-linear effect of age is unknown and varies across

regions. We will employ the penalized cubic splines to reveal the underlying nonlinear effects of age in the development of the cerebellum.

Besides the inflexibility of the regression methods, most previous studies of characterizing cerebellar trajectories focus on finding the best fit and the interpretation of the volumetric development is drawn by simply looking at the fitted curves, which is heuristic and approximate. A more detailed and statistically meaningful description is required to assess how the cerebellar volumes change across age, inspiring us to look into the derivatives of a penalized cubic spline which represent the rate of cerebellar volume change.

It is clear that the analytical derivatives of spline functions can be exactly expressed in terms of lower order spline functions, as described in De Boor (1972). The asymptotic bias, variance and normality of these analytical estimators have also been derived in Zhou and Wolfe (2000). However, from the perspective of computation, numeric approximation of the derivatives is preferred, among which using the difference quotient is common. For example, Ilyasov (2014) analyzes the growth rate (the first derivative) and growth acceleration (the second derivative) of the economic indexes through differentiation of cubic spline models, and detect the latent trends of the economic dynamic. In this paper, we will also adopt difference quotient to estimate the first derivatives of the fitted penalized cubic spline at a grid of age points, which gives more details on how fast cerebellar volumes change across age.

Once the first derivative estimates are obtained, we develop further hypothesis tests of whether the derivatives are significantly different from 0. The variance of the derivative estimate of the fitted penalized cubic spline is required for this test and the detailed procedure of how to obtain this variance will be given in Section 3.2. This step helps us to find the statistically significant increasing and decreasing periods of the cerebellar volume of interest. Similar intuition can be found in an ecological study by Monteith et al. (2014). They fit a semi-parametric additive model to the hydro-chemical data from 22 lakes and streams, and identify the periods along the fitted trend where the first derivative estimates are significantly different from 0. We won't stop at identifying periods where the cerebellar volume is either increasing or decreasing in a significant manner: The age(s) at which the cerebellar volume of interest reaches its peak(s) is also assessed after the identification of the significantly changing periods.

The data we will analyze is obtained from the HCP Development study (Human Connectome Project, 2021) and the HCP Young Adult study (Human Connectome Project, 2018) which cover the cerebellar volume data from children, adolescents and adults. The volumes of interest will include total cerebellar volume, cerebellar cortex (i.e. gray matter) volume, cerebellar white matter volume, and the corresponding adjusted

volumes after considering the total intracranial volume, along with the total intracranial volume itself. To summarize, this paper aims to characterize the developmental trajectories of the 7 volumes during childhood, adolescence and adulthood with penalized cubic splines and to estimate the rate of change of the fitted trajectories by difference quotient approximation. The following aim is to identify the statistically significant increasing or decreasing periods of the volume of interest and to determine the peak(s). To our knowledge, this study is by far the first to depict how cerebellar volumes change through estimating derivatives of the fitted trajectories and identifying statistically significant changing periods as well peaks.

2 Dataset and Preliminary Analysis

Combining the data from the HCP Development and HCP Young Adult studies (Human Connectome Project, 2021, 2018), a total of 1608 healthy subjects from 5 to 37 years old are included, among which 747 are males and 861 are females. The studied age range covers normal individuals during childhood, adolescence and adulthood. Table 1 lists the mean, standard deviation, median and range of age among males, females and the whole cohort respectively. We also run preliminary tests to check whether age is distributed differently among males and females. Both the independent 2-sample t-test and the Mann–Whitney–Wilcoxon (MWW) test are conducted to test the existence of a significant difference of the mean age and of the age distribution. The p - values are shown in the last column of Table 1 and we can see that the mean age across sex is not significantly different (0.1143) while the age distribution is (0.0274).

Table 1: Demographic characteristics of the study sample, stratified by sex

	Male ($N = 747$, 46.455%)	**Female** ($N = 861$, 53.545%)	**Total** ($N = 1608$)	p **-Value**[a] ($\alpha = 0.05$)
Age (years)				
Mean (SD)	22.384 (7.738)	23.028 (8.597)	22.729 (8.213)	0.1143^{b}
Median [Range]	24 [5, 37]	25 [6, 36]	25 [5, 37]	0.0274^{*cd}

[a] The p -values are derived from testing whether the ages are significantly different in males and females.

[b] The p -value is derived from the independent 2-sample t-test assuming unequal variance.

[c] The p -value is derived from the Mann–Whitney–Wilcoxon test.

[d] p -value $< 0.001^{***}$, p -value $< 0.01^{**}$, p -value $< 0.05^{*}$.

The volumes of interest include total cerebellum volume (TCV), cerebellar cortex volume (CCV), cere-

bellar white matter volume (CWM), estimated total intracranial volume (TICV) and the TICV-adjusted measurements of TCV, CCV, and CWM (aTCV, aCCV and aCWM). The last 3 volumes are relative volumes with adjustment for the general head size. All 7 outcomes are completely recorded for the 1608 subjects along with their age. Table 2 summarizes the mean and standard deviation of the 7 volumes stratified by sex and in the whole cohort. It's apparent that all of the 7 mean volumes are larger in males and we perform an independent 2-sample t-test for each of the 7 volumes to test against the alternative that males have a larger mean volume than that in females. From the p - values listed in the last column of Table 2, all of the 7 volumes have a significantly larger mean for male subjects than those in female subjects.

Table 2: Average volumes of interest, strstified by sex

	Male Mean (SD)	**Female** Mean (SD)	**Total** Mean (SD)	p **-Value**[a] ($\alpha = 0.05$)
TCV[b] **(cm^3)**				
Absolute	154.758 (13.466)	138.133 (12.001)	145.856 (15.167)	$< 2.2 \times 10^{-16***}$
Adjusted[c]	148.724 (10.804)	143.368 (9.589)	145.856 (10.514)	$< 2.2 \times 10^{-16***}$
CCV[b] **(cm^3)**				
Absolute	124.611 (10.982)	111.119 (10.184)	117.387 (12.522)	$< 2.2 \times 10^{-16***}$
Adjusted	120.057 (9.206)	115.070 (8.381)	117.387 (9.117)	$< 2.2 \times 10^{-16***}$
CWM[b] **(cm^3)**				
Absolute	30.147 (4.024)	27.014 (3.120)	28.469 (3.895)	$< 2.2 \times 10^{-16***}$
Adjusted	28.667 (3.464)	28.298 (2.799)	28.469 (3.130)	0.010^{**}
TICV[b] **(cm^3)**	1701.329 (141.516)	1488.715 (144.306)	1587.485 (178.024)	$< 2.2 \times 10^{-16***}$

[a] The p -values are derived from the independent 2-sample t-test of whether each mean volume is significantly larger in males than that in females.

[b] TCV – total cerebellum volume, CCV – cerebellar cortex volume, CWM – cerebellar white matter volume, TICV – estimated total intracranial volume.

[c] The corresponding volume after adjusting the estimated total intracranial volume (TICV).

This situation coincides with previous findings on gender dimorphism of cerebellar volumes. Larger cerebellar volumes in male adults have been reported by several studies, consistent with the overall larger brain size (Cho et al., 1999; Raz et al., 2001; Szabó et al., 2003). In children, adolescents and young adults, the model-based total cerebellar volume is 10% to 13% larger in males depending on the age of comparison and

this dimorphism remains significant after adjusting the total brain volume (Tiemeier et al., 2010). Besides the volume difference, late peaks of the male cerebellum have also been noted in a cross-sectional case (Caviness Jr et al., 1996) and a longitudinal case (Tiemeier et al., 2010). Therefore, to investigate a potentially different time course of cerebellar volumes in both shape and height, the following statistical analysis will be performed separately for males and females.

3 Statistical Methods

All statistical analyses were performed using R version 4.0.5 (RStudio Team, 2021), with the penalized cubic splines fitted in package mgcv (Wood (2011), Wood (2017)). A significance level of 0.05 is adopted for all of the hypothesis tests unless mentioned otherwise.

3.1 Building Developmental Trajectories of the Cerebellar Volumes

For a given scatter plot of each volume outcome, the model we apply is penalized cubic splines, which is a compromise between cubic regression splines and smoothing splines. We will gradually build our way to the form of a penalized cubic spline from the basis of spline modeling.

The form of a truncated power basis of degree p is given in (1) with K knots at $k_1, k_2, ..., k_K$, where $(x-k)_+^p$ equals $(x-k)^p$ when $x > k$ and 0 otherwise. The linear combination of the truncated power basis is called a spline, which is the $f(x)$ in formula (2). $f(x)$ is a piecewise degree-p polynomial and has a continuous $(p-1)_{th}$ derivative. If the knots are selected to cover the observations well, splines are flexible to fit a wide variety of shapes. Higher values of p lead to smoother functions, but generally $p = 3$ gives a curve smooth enough that human eyes cannot tell compared to higher order splines (Ruppert et al., 2003).

$$1, x, x^2, ..., x^p, (x-k_1)_+^p, (x-k_2)_+^p, ..., (x-k_K)_+^p \tag{1}$$

$$f(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + ... + \alpha_p x^p + \sum_{m=1}^{K} \alpha_{pm}(x-k_m)_+^p \tag{2}$$

One problem with the basic splines is that they tend to have high variance at the boundaries of the observations, especially when p is large. A remedy is to force the piecewise polynomial function to have a lower degree to the left of the leftmost knot, and to the right of the rightmost knot, which is how natural

splines work. The definition of a natural spline is as follows (given the same set of knots at $k_1, k_2, ..., k_K$):

1. $f(x)$ is a polynomial of degree p on each of the intervals $[k_1, k_2], [k_2, k_3], ..., [k_{K-1}, k_K]$.
2. $f(x)$ has continuous $(p-1)_{th}$ derivatives at knots $k_1, k_2, ..., k_K$.
3. $f(x)$ is a polynomial of degree $\frac{(p-1)}{2}$ on $(-\infty, k_1]$ and $[k_K, \infty)$.

The first 2 components agree with the basic splines in formula (2) while the last requirement of a lower order outside the range of observations distinguishes natural splines. We will be using natural cubic splines (natural splines with $p = 3$) which are linear beyond the boundary knots. An appealing feature of a natural cubic spline is that it arises as the solution of an optimization problem described in equation (3), where $(x_i, y_i), i = 1, ..., n$ are a set of observations and $f(x)$ includes all that have continuous second derivatives.

$$\min_f \{\sum_{i=1}^{n}(y_i - f(x_i))^2 + \lambda \int (f''(x))^2 dx\} \tag{3}$$

It turns out that the unique solution to equation (3) is a natural cubic spline with knots at every unique observation $x_1, ..., x_n$. The resulting $f(x)$ is called a smoothing spline (Reinsch, 1967). The objective function in equation (3) trades off the squared error of f over (x_i, y_i) and a penalty term that grows larger as the second derivatives of f become more wiggly. $\lambda > 0$ is a smoothing parameter that controls the trade-off between fidelity to the data and roughness of the function estimate (Wikipedia contributors, 2020). If $f(x)$ is restricted to a natural spline but with user-defined knots, and the penalty term in equation (3) is removed, the solution reduces to a least squared estimator, which is a regression spline. Regression splines have the potential to be over-fitted or over-wiggly and this problem is eased by the regularization of the integrated square second derivative of smoothing splines.

Although smoothing splines enjoy nice theoretical properties, the only substantial problem is that they have as many free parameters as there are data to be smoothed since the knots are placed at every unique observation. Actually, due to the existence of λ, the resulting smoothing spline is always much smoother than n degrees of freedom would suggest (Wood, 2017). This inspires the combination of the freely chosen degree of freedom (i.e., number of knots) of a regression spline and the roughness penalty of a smoothing spline, which gives penalized splines (Wahba, 1980). Specifically, penalized cubic splines are defined as the solution to the optimization problem in equation (4), where the number and positions of the knots are freely chosen in advance by the user and K is usually much smaller than n:

$$\begin{cases} \min\limits_{f}\{\sum_{i=1}^{n}(y_i - f(x_i))^2 + \lambda \int (f''(x))^2 dx\} \\ f \text{ is a natural cubic spline with pre-defined knots at } k_1, ..., k_K \end{cases} \tag{4}$$

We can see that the only difference between a penalized cubic spline and a smoothing spline defined by equation (3) is that penalized splines place more constraints on the form of f to be a simpler natural cubic spline with a reduced number of knots. The pre-defined and reduced knot selection resembles a regression spline and the penalty term regulates the roughness as in a smoothing spline. To guarantee a satisfying fit, the covariate values in the knot set should be arranged to nicely cover the distribution of covariate values in the original data set (Wood, 2017).

Up to this point, it's time to write the penalized cubic spline models we use to characterize the developmental trajectories of the cerebellar volumes of interest. Separately for males and females, and for TCV, CCV, CWM, aTCV, aCCV, aCWM, and TICV, denote y_i as the corresponding volume of the i_{th} subject. We assume that y_i is represented by a function of the subject's age plus a Gaussian random error with mean 0, as expressed in formula (5). $f(Age_i)$ is a natural cubic spline representing the conditional mean of y_i given the subject's age and it's estimated following the optimization framework of a penalized cubic spline described in (4). The number of knots is chosen to be 10 and they evenly cover the quantiles of the age distribution, i.e., the knots are placed at the 0%, 11.111%, 22.222%, 33.33%, 44.444%, 55.556%, 66.667%, 77.778%, 88.889%, and 100% quantiles of age. Note that the penalized cubic splines are fitted separately for males and females, so the positions of the 10 knots are different. The smoothing parameter λ is selected using the restricted maximum likelihood (REML) method (Wood, 2011).

$$y_i = f(Age_i) + \epsilon_i,\ \epsilon_i \overset{iid}{\sim} N(0, \sigma^2) \tag{5}$$

In fact, it's possible to write out the explicit solution of the penalized cubic splines defined in (4) if we write out an equivalent form of the objective function. For a 10_{th} order natural splines with the knots fixed as in the last paragraph, the corresponding basis functions are noted as $g_1, ..., g_{10}$. Then $f(x_i) = \sum_{m=1}^{10} \beta_m g_m(x_i)$ and the design matrix $\boldsymbol{X} \in R^{n\times 10}$ can be written as $\boldsymbol{X}_{ij} = g_j(x_i)$. The first term of the objective function now becomes $\sum_{i=1}^{n}(y_i - f(x_i))^2 = ||\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}||^2$, where $\boldsymbol{Y} = (y_1, ..., y_n)^T, \boldsymbol{\beta} = (\beta_1, ..., \beta_{10})^T$. Then we define a matrix $\boldsymbol{S} \in R^{10\times 10}$ where $\boldsymbol{S}_{ij} = \int g_i''(x) g_j''(x) dx$. $\boldsymbol{S}$ is called the roughness penalty coefficient matrix and it can be proved that the optimization problem in (4) is equiva-

lent to finding $\boldsymbol{\beta}$ to minimize the objective function in (6) (Wood, 2017), which has an explicit solution shown in equation (7). Now the estimated mean curves of each volume of interest can be expressed as $(\hat{f}(Age_1), ..., \hat{f}(Age_n))^T = \boldsymbol{X}\hat{\boldsymbol{\beta}}$, corresponding to the estimated developmental trajectories of the 7 volumes that will appear in the figures of the fitted splines as black solid lines.

$$||\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}||^2 + \boldsymbol{\beta}^T\boldsymbol{S}\boldsymbol{\beta} \tag{6}$$

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X} + \lambda\boldsymbol{S})^{-1}\boldsymbol{X}^T\boldsymbol{Y} \tag{7}$$

The effective degree of freedom (edf) and the significance of the smooth term $\hat{f}(Age)$ will also be given. Edf of the smooth is defined as the trace of the influence matrix $\boldsymbol{F}$ where $\boldsymbol{F} = (\boldsymbol{X}^T\boldsymbol{X}+\lambda\boldsymbol{S})^{-1}\boldsymbol{X}^T\boldsymbol{X}$, which represents the complexity of the estimated developmental trajectories. $edf = 1$ is equivalent to a straight line, $edf = 2$ is equivalent to a quadratic curve, etc., with higher edfs describing more wiggly patterns of how the volumes of interest change across age (Wood and Wood, 2015). The significance of the smooth terms of age will be reported by p - values of the tests that $\hat{f}(Age) = 0$. The p - values are approximate in the sense that the components of the test statistic are weighted by the iterative fitting weights (Wood and Wood, 2015). How to obtain the theoretical p - values is described in Wood (2013).

3.2 Estimating Rate of Volume Change

Once the mean volumes conditioned on age are obtained from the penalized cubic spline models, we move on to estimate the first derivatives of the fitted penalized cubic splines. The definition of the first derivative of a smooth function g at a point x is $g^{'}(x) = \lim_{h\to 0}\frac{g(x+h)-g(x)}{h}$, which indicates estimating $g^{'}(x)$ using $\frac{g(x+h)-g(x)}{h}$ with a sufficiently small h. $g(x+h)-g(x)$ is a finite difference and $\frac{g(x+h)-g(x)}{h}$ is known as the difference quotient. It can be proved that the error between the difference quotient estimator and the true derivative is $O(h)$ if g has continuous second derivatives around x (Iserles, 2009). This means the error will converge to 0 as h tends to 0, which validates the approximation. We will refer to h as step size afterwards. Since the penalized cubic splines possess continuous second derivatives within the range of the observations, it's natural to consider estimating the first derivatives of our fitted splines using quotients.

We divide the age range into a uniform grid with 1000 grid points, on which the first derivatives with respect to age will be estimated. The number of grid points is chosen to be large since we want to draw a

detailed picture of the derivatives, even though ages are integers in the data with around 30 unique values. This will help the detection of significantly changing periods in the next subsection. Denote the 1000 grid points of age as $a_1, ..., a_{1000}$ and the step size is set to be $h = 10^{-7}$. We estimate the volume derivatives to age as follows, where $\hat{d}(x)$ refers to the estimated derivative at age x. This process is repeated for each fitted penalized cubic spline of the 7 outcomes separately for males and females:

1. Predict the mean volumes at $a_1, ..., a_{1000}$ from the fitted penalized cubic splines:

$$(\hat{f}(a_1), ..., \hat{f}(a_{1000}))^T = \boldsymbol{X}_a \hat{\boldsymbol{\beta}} \tag{8}$$

   where $\boldsymbol{X}_a \in R^{1000\times 10}$ is the design matrix under the supplied value of covariates $\{a_1, ..., a_{1000}\}$, and $\hat{\boldsymbol{\beta}}$ is given in equation (7).

2. Predict the mean volumes at $a_1 + h, ..., a_{1000} + h$ from the fitted penalized cubic splines, i.e., move the grid points forward with a step size of h and obtain new predictions:

$$(\hat{f}(a_1 + h), .., \hat{f}(a_{1000} + h))^T = \boldsymbol{X}_{a+h} \hat{\boldsymbol{\beta}} \tag{9}$$

   where $\boldsymbol{X}_{a+h} \in R^{1000\times 10}$ is the design matrix under the age set $\{a_1 + h, ..., a_{1000} + h\}$.

3. The derivative of the volume at age a_i is estimated to be $\hat{d}(a_i) = \frac{\hat{f}(a_i+h)-\hat{f}(a_i)}{h}, i = 1, ..., 1000$. Write the derivative estimates in matrix form:

$$\hat{\boldsymbol{d}} = (\hat{d}(a_1), ..., \hat{d}(a_{1000}))^T = \frac{(\boldsymbol{X}_{a+h} - \boldsymbol{X}_a)}{h} \hat{\boldsymbol{\beta}} \tag{10}$$

The matrix notation of the derivative estimates facilitates calculating its covariance matrix since $\hat{\boldsymbol{d}}$ is essentially a linear transformation of the coefficient estimates. Marra and Wood (2012) derive a posterior Bayesian covariance matrix for $\hat{\boldsymbol{\beta}}$ which roots from Nychka (1988). This covariance matrix results from adopting a particular Bayesian model of the smoothing process (Wood and Wood, 2015) and we note it as $\boldsymbol{V}_B$. An estimate of $\boldsymbol{V}_B$ is shown in equation (11) where $\hat{\sigma}^2$ is the residual sum of squares for the fitted model, divided by the effective residual degrees of freedom (Wood, 2017).

$$\hat{\boldsymbol{V}_{\boldsymbol{\beta}}} = (\boldsymbol{X}^T \boldsymbol{X} + \lambda \boldsymbol{S})^{-1} \hat{\sigma}^2 \tag{11}$$

Naturally, we can estimate the covariance matrix of $\hat{\boldsymbol{d}}$ as shown in equation (12), the diagonal of which gives the variance estimates of the fist derivatives.

$$\hat{\boldsymbol{Cov}}(\hat{\boldsymbol{d}}) = \frac{(\boldsymbol{X}_{a+h} - \boldsymbol{X}_a)(\boldsymbol{X}_{a+h} - \boldsymbol{X}_a)^T}{h^2}\hat{\boldsymbol{V}}_{\boldsymbol{\beta}} \tag{12}$$

From the estimated Bayesian covariance matrix of the coefficient estimates in (11), we can also obtain the point-wise Bayesian confidence intervals of $f(Age)$ according to Marra and Wood (2012) and we will plot those confidence bands in the figures of the fitted splines. It turns out that these point-wise Bayesian confidence intervals have surprisingly good coverage by a frequentist criterion, provided that coverage is measured as an average across the observation points (Marra and Wood, 2012). From simulations by Wahba (1983) and the theoretical proof by Nychka (1988), the average coverage probability (ACP):

$$ACP = \frac{1}{n}\sum_{i=1}^{n} P(f(Age_i) \in BI_\alpha(Age_i))$$

is close to the nominal level $1-\alpha$, where $BI_\alpha(Age)$ indicates the $(1-\alpha)100\%$ Bayesian interval for $f(Age)$ (Marra and Wood, 2012).

3.3 Detection of Significantly Changing Periods and Peaks

The rate of change estimates in $\hat{\boldsymbol{d}}$ yield positive/negative (increasing/decreasing) trend at some age points and the estimated variance of those difference quotients give the uncertainty measurement of the rate of change. For those derivative estimates that are not 0, we want to include the uncertainty information to test the hypothesis that whether each of them is significantly different from 0, which helps detect statistically valid increasing or decreasing periods of the volumes.

Since the estimated variance of the difference quotients are readily available in equation (12) and the random errors of the penalized cubic splines in formula (5) are assumed to be Gaussian, the 95% confidence interval of $f^{'}(a_i)$ can be reasonably estimated to be:

$$\hat{d}(a_i) \pm Z_{0.975} \times diag\{\hat{\boldsymbol{Cov}}(\hat{\boldsymbol{d}})\}_i,\ i = 1, ..., 1000 \tag{13}$$

where $Z_{0.975}$ is the 97.5% quantile of a standard normal distribution and $diag\{\hat{\boldsymbol{Cov}}(\hat{\boldsymbol{d}})\}_i$ is the i_{th} diagonal element of $\hat{\boldsymbol{Cov}}(\hat{\boldsymbol{d}})$, which is the estimated variance of $\hat{d}(a_i)$. This confidence interval is calculated for

each difference quotient $\hat{d}(a_1), ..., \hat{d}(a_{1000})$. For any difference quotient that is positive (negative), if its 95% confidence interval does not contain 0, we will reject the null and claim a significantly increasing (decreasing) trend at its corresponding age. Consecutive age points in $\{a_1, ..., a_{1000}\}$, at which a significantly increasing (decreasing) trend is claimed, form an increasing (decreasing) age period of the volume of interest.

$\hat{\boldsymbol{d}}$ will be plotted along with the confidence intervals described in (13), and the derivative estimates at the significantly increasing (decreasing) age period(s) will be marked in red (blue). The identified significant increasing or decreasing periods of the cerebellar volumes will also be marked in the same manner in the figures of the fitted splines.

Note that the confidence intervals in (13) are point-wise rather than simultaneous, similarly to the confidence intervals of $f(Age)$ mentioned above. Since they are also derived based on the posterior Bayesian covariance matrix estimate ($\hat{\boldsymbol{V}}_B$) in (11), those confidence intervals of the derivatives are Bayesian confidence intervals as well. The point-wise nature of the confidence intervals only allow for a point-wise inspection of the hypothesis tests of the derivatives. Multiple testing of the rate of volume change can be achieved only if the confidence intervals of the derivatives are simultaneous.

After detecting the significantly changing age periods for each of the 7 outcomes, we seek to find the age at which the estimated developmental trajectory is at its peak. In this paper, a peak is defined as the local maximum of the estimated developmental trajectory sandwiched between a significantly increasing age period before and a significantly decreasing age period after. Conceptually, any local maximum of the inverse-U in the trajectory will be classified as a peak, given that there exists a significant increasing period on the left and a significant decreasing period on the right part of the inverse-U. Each peak represents a significant "bump" in the development of the cerebellar volume of interest. If no peak is identified, the global maximum of the estimated mean volume will be reported. We will mark the peak(s) or maximum, along with the corresponding peak or maximum volume(s) in the figures of the fitted splines.

To evaluate the robustness of the identified significantly changing period(s) and the peak(s)/maximum, 1000 times of stratified resampling is performed separately for males and females: such robustness analysis can be very useful when dealing with noisy and complex data sets. Each time, we randomly select 80% of the samples within each age group. The reason for treating age as a stratifier is to retain the age distributions shown in Figure A.1 (males) and Figure A.1 (females). The process of characterizing the developmental trajectories as penalized cubic splines, estimating the rate of volume change, detecting significant changing period(s) of age and peak(s)/maximum is repeated for each resample. The number of peaks identified in the

1000 resamples will be summarized in proportions. The proportions of the original increasing/decreasing period(s) and peak(s)/maximum re-identified in the 1000 resamples will also be reported. To allow for an acceptable amount of bias, we also check the proportions of the fluctuated increasing/decreasing period(s) and peak(s)/maximum identified in the original data. For example, if [7,10] is identified as a significantly increasing period in the original data, we also check the proportion of any ranges contained by [6,11] being identified as a significantly increasing period in the 1000 resamples. If 12 years old is identified as a peak originally, we also check the proportion of 11, 12, or 13 years being identified as a peak in the 1000 resamples.

4 Results

To facilitate reading, abbreviations of the volume outcomes are restated: TCV – total cerebellum volume, CCV – cerebellar cortex volume, CWM – cerebellar white matter volume, TICV – estimated total intracranial volume; aTCV – TICV-adjusted total cerebellum volume, aCCV – TICV-adjusted cerebellar cortex volume, aCWM – TICV-adjusted cerebellar white matter volume.

4.1 Developmental Trajectories of the Cerebellar Volumes and Trend Assessment

The effective degrees of freedom (edf) of the fitted penalized cubic splines are listed in Table 3 separately for male and female subjects and for each of the volume outcomes. We can see that except for the TICV-adjusted TCV of females, all fitted trajectories behave more sophisticated patterns than linear (edf larger than 1), indicating the non-linear effect of age on the total cerebellum volume (only the unadjusted), cerebellar cortex volume, cerebellar white matter volume and total intracranial volume. We notice that adjusting TICV seems to explain some non-linearity of the cerebellar volume growth for females. For example, the edf of the TICV-adjusted CCV trajectory is 1.719 (close to quadratic), which is largely reduced compared with that of the absolute CCV trajectory (6.141, close to a 6-degree polynomial). But this reduction of the non-linearity due to TICV adjustment is not as obvious in males.

From the p - values displayed in the same table, the smooth term of age is significant for all but the TICV-adjusted TCV for both males and females. This reflects that after adjusting for the total intracranial volume, age is no longer a significant factor affecting the volumetric development of the cerebellum which is intuitive: The total intracranial volume is closely related to the general head size with which the cerebellum scales. Under the current data, regressing out the head size seems to eliminate the age effect on the volume of

the cerebellum as a whole, while age maintains a significant contribution to the volume growth for cerebellar gray matter and white matter even after adjusting for the head size. Due to the large p - value of 0.363, we will not include the TICV-adjusted TCV for females in the following trend assessment.

Table 3: Edf and significance of the smooth term, stratified by sex.

	Effective Degree of Freedom		**p -Value**[a] ($\alpha = 0.05$)	
	Male	**Female**	**Male**	**Female**
TCV (cm^3)				
Absolute	4.997	6.071	0.0015^{**}	$4.44 \times 10^{-6***}$
Adjusted	4.063	1.002	0.0786	0.363
CCV (cm^3)				
Absolute	5.101	6.141	0.0001^{***}	$< 2 \times 10^{-16***}$
Adjusted	4.394	1.719	$1.02 \times 10^{-6***}$	$2.48 \times 10^{-7***}$
CWM (cm^3)				
Absolute	4.044	4.334	$< 2 \times 10^{-16***}$	$< 2 \times 10^{-16***}$
Adjusted	3.602	2.599	$< 2 \times 10^{-16***}$	$< 2 \times 10^{-16***}$
TICV (cm^3)	3.042	5.201	0.0057^{**}	$< 2 \times 10^{-16***}$

[a] The p - values are derived from an approximate F-test of whether the smooth term of age is significant (or equivalently, whether is 0) in the penalized cubic spline model for each volume of interest in each sex.

Figure 1 illustrates the estimated growth curve of the 6 cerebellar volumes for male subjects estimated from the penalized cubic spline models. Figure 2 gives the cerebellar growth curves for females. Absolute volumes and their corresponding TICV-adjusted volumes are displayed side-by-side. The fitted volumetric developmental trajectories are presented in the Appendix after adjusting the scale of the y-axis to include all the data points in the plot. Please refer to Figure A.3 and A.4.

For male subjects, we can roughly tell that TCV, aTCV, CCV and aCCV increase until some point during adolescence and begin to fluctuate or decrease afterward. While for CWM and aCWM, the increase seems consistent during the whole studied age range (5 to 37 years). Further analysis of the rate of change produces the derivative estimates of the cerebellar trajectories in Figure 3. They measure how fast the cerebellar

volumes are changing at specified ages. Following the test procedure described in Section 3.3, the positive (negative) derivative estimates claimed to be significantly different from 0 are marked in red (blue) with the corresponding age periods labeled on the x-axis. According to the information provided in Figure 3, the significantly changing periods of age are also marked on the estimated trajectories in Figure 1.

For male TCV, the penalized cubic spline predicts an increase from 5 to 14 years old, and the derivative analysis determines the increase from 5 to 12 years to be significant. Maximum TCV is reached at 14 years old, followed by insignificant volume fluctuation to 37 years. Male aTCV develops in a similar pattern with the only difference of increasing significantly from 5 to 11 years old. For both CCV and aCCV, an inverse-U-shaped trajectory along with a single peak is detected. CCV rises significantly from 5 to 12 years and experiences a significant drop from 15 to 16 years with a peak at 13 years old. The significant increase of aCCV happens during [7,9], peaking at the age of 12, followed by a continuous volume loss afterward in which [15,20] is the significantly decreasing age period. While for cerebellar white matter volumes, the time-course is much simpler with a steady increase throughout the whole age range (5 to 37 years). The increase is significant during [5,24] and the maximum value is reached at 27 years old.

For females, results from a similar derivative analysis are given in Figure 4 and the periods of significant changes are also marked in the predicted developmental trajectories in Figure 2. The TICV-adjusted TCV is not presented since age does not affect aTCV significantly. For female TCV, the predicted trajectory exhibits 2 inverse-U shapes connected together (i.e., increase-decrease-increase-decrease). The significant increase happens from 6 to 10 years as well as from 25 to 26 years. The significant decrease also appears in 2 periods which are [19,20] and [29,33]. The local maximum of the 2 inverse Us are the identified 2 peaks at age 12 and age 28 respectively. The growth of the female cerebellar cortex follows a similar trend as the total cerebellum with 2 inverse Us connected together. The 2 significant increasing periods remain as [6,10] and [25,26], and CCV decreases significantly on a longer period of [15,17]-[18,20]. The second significantly decreasing period is still [29,33]. The peaks of CCV are identified as age 11 and 28 years old. However, after adjusting for TICV, aCCV is predicted to show a sustained volume loss across the age range of 6 to 36 years (with the decrease in [8,27] being significant). For female cerebellar white matter volumes, the trend of the time-courses resembles that of male CWMs, which increases steadily throughout the whole age range (6 to 36 years). For the absolute CWM, the increase is significant during [6,15] and [22,24] and the maximum value is reached at 28 years old. After TICV adjustment, a significant increase is detected in the age period of 6 to 33 years and reaches its maximum at 36 years old.

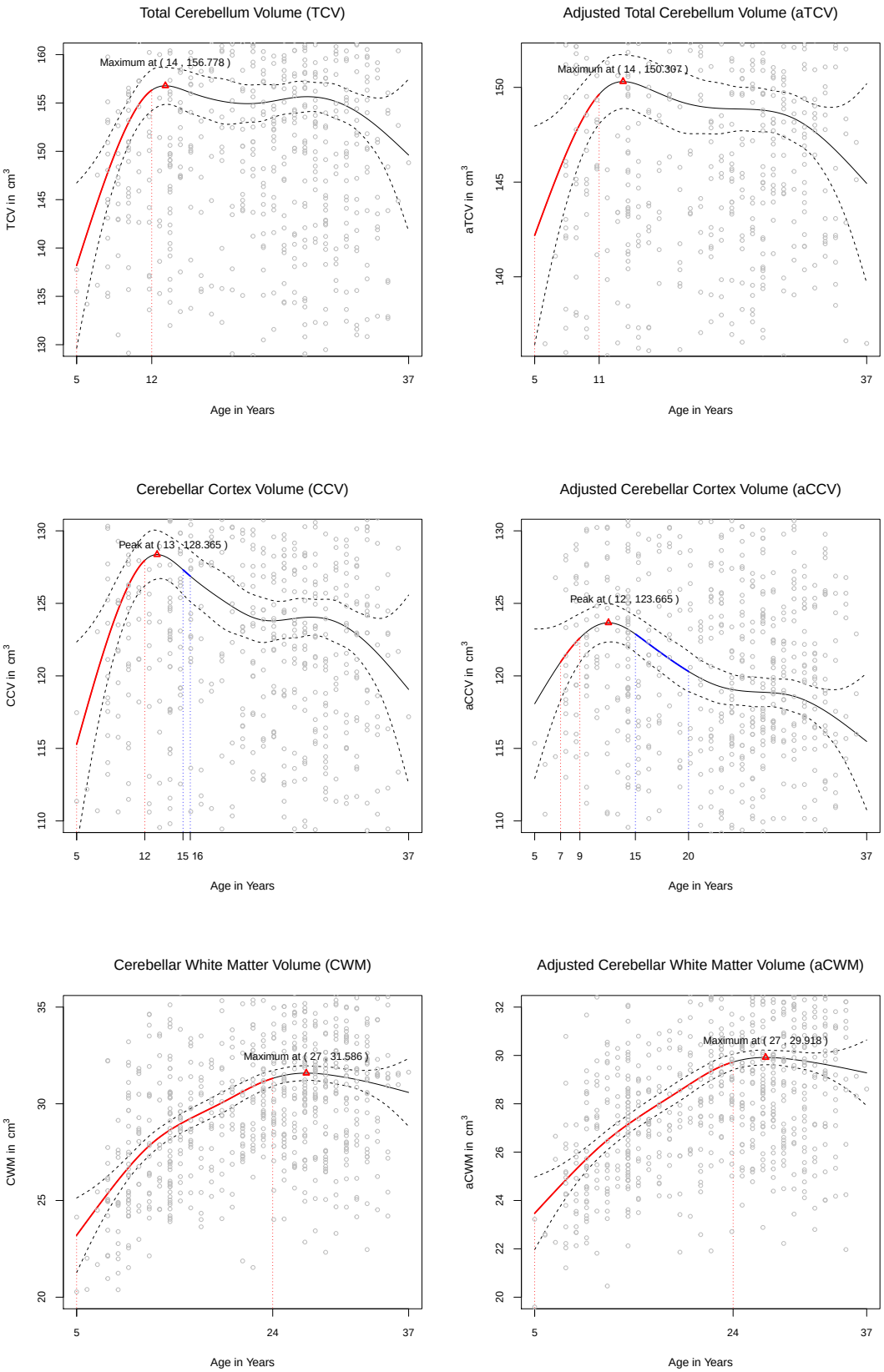

Figure 1: Estimated developmental trajectories for the cerebellar volumes in male subjects. Black solid lines: mean volumes estimated by penalized cubic splines; Dashed bands: point-wise Bayesian confidence intervals; Grey hollow dots: original data samples; Red triangles: peaks/maximum.

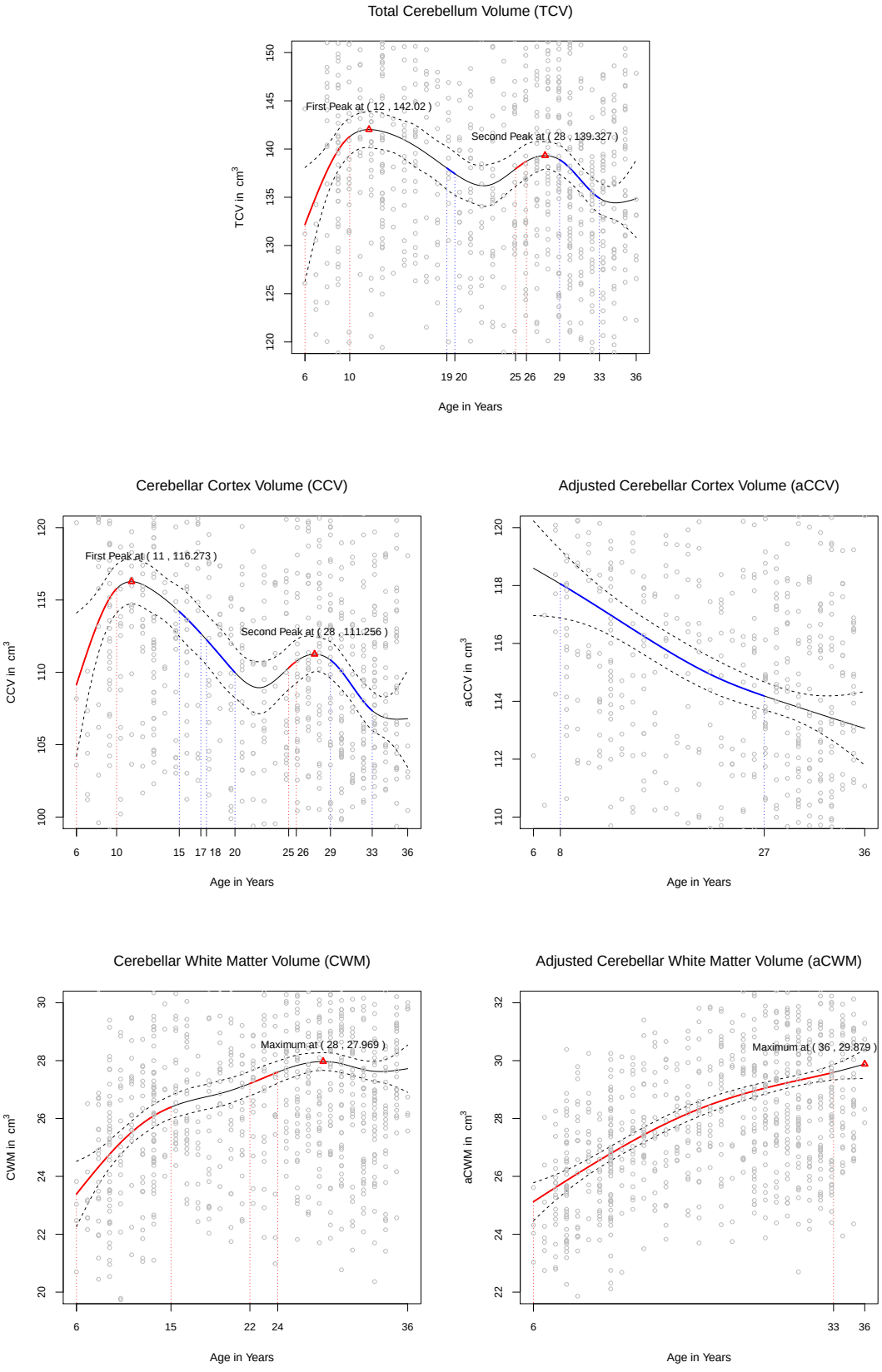

Figure 2: Estimated developmental trajectories for the cerebellar volumes in female subjects.

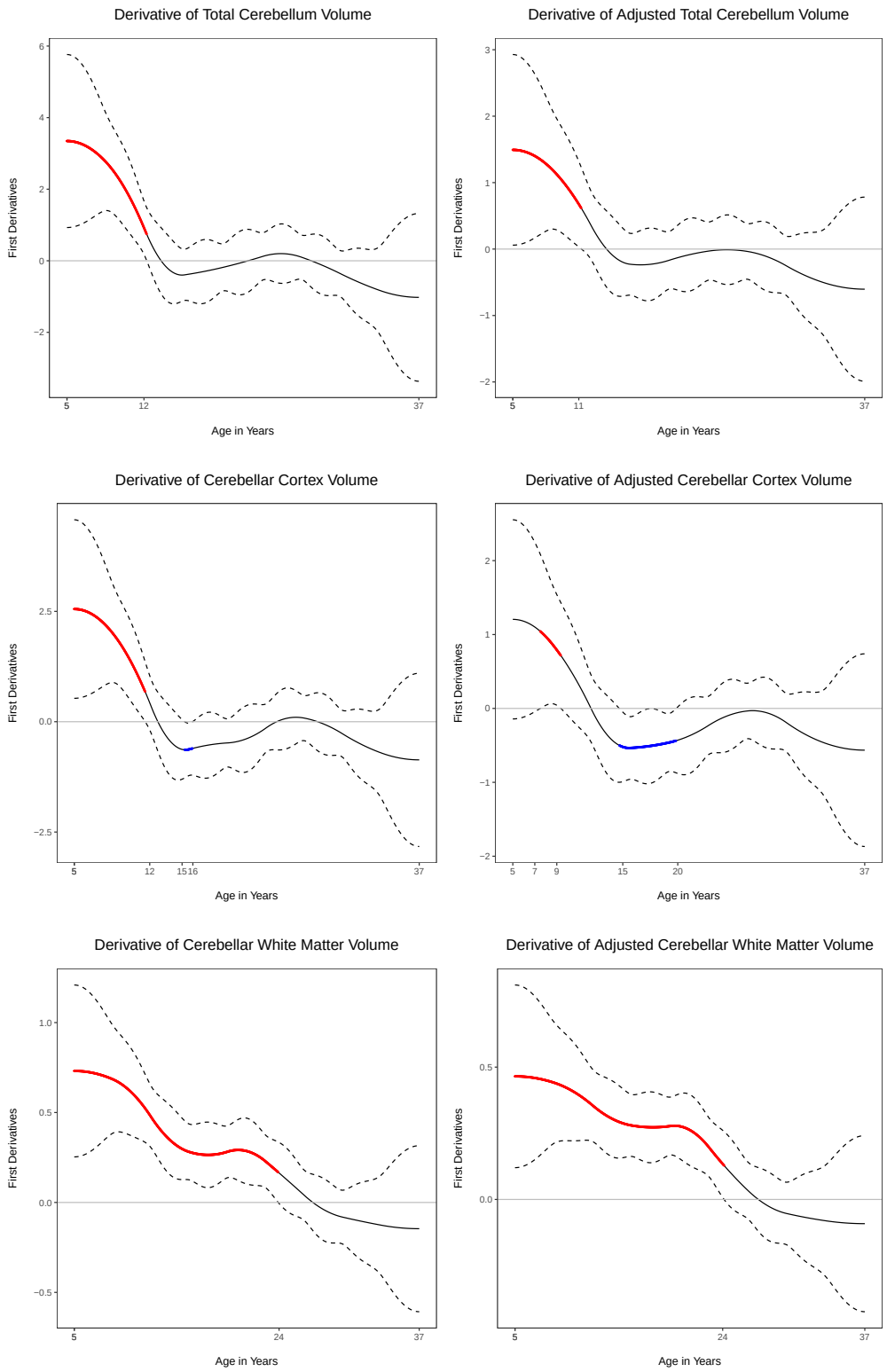

Figure 3: Derivative estimates of the developmental trajectories for the cerebellar volumes in male subjects. Black solid lines: estimated rate of change by difference quotients; Dashed bands: point-wise Bayesian confidence intervals; Grey horizontal line: derivative $= 0$; Derivative estimates marked in red (blue): of which the confidence intervals do not contain 0, indicating a significantly increasing (decreasing) period of the volume of interest on the corresponding age range.

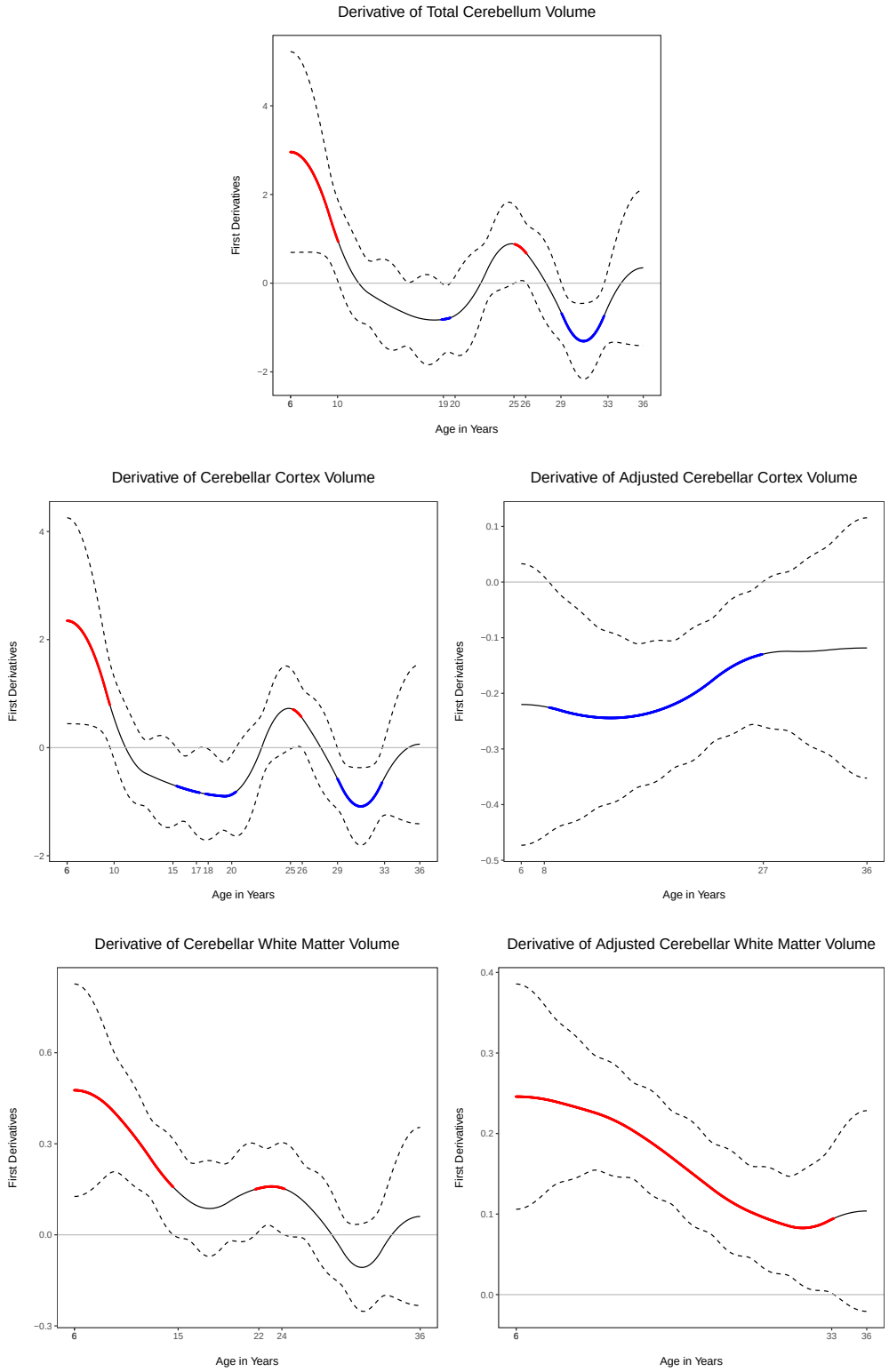

Figure 4: Derivative estimates of the developmental trajectories for the cerebellar volumes in female subjects.

The estimated mean volumetric trajectories of TICV and the analysis of its rate of change are performed for TICV as well with the results shown in Figure 5 and 6. Male TICV increases significantly during 6 to 14 years old and experiences minimal change afterward with the maximum volume attained at age 27. The trend of female TICV growth looks like 2 inverse-U connected together as for female TCV and CCV, but the derivative analysis determines [8,11] as the only significantly increasing period and 2 periods of significant decrease which is [15,19] and [29,33]. That's why only 1 peak at age 13 is identified by the definition of a peak mentioned in Section 3.3.

Based on the current data and the penalized cubic spline modeling, we spot an obvious gender dimorphism of how the cerebellum and its grey matter grow during the age range covered in this study (5-37 years). For TCV, the connected inverse-U pattern in female subjects seems different from how male TCV grows: Male TCV exhibits a single period of significant increase from childhood to adolescence without any other significant volume change. For CCV, a similar difference exists between genders. It's also interesting how different the adjustment of TICV is in explaining the change of TCV and CCV in different gender groups. In male subjects, the general patterns of the trend of TCV and CCV are not changed after adjusting for TICV. While for female subjects, the age effect on TCV is eliminated after adjusting for TICV; and the age effect on aCCV follows a steady decrease instead of a connected inverse-U before TICV adjustment. Although further robustness check of the several "bump"s for female TCV and CCV is required (provided later), we can sense the gender dimorphism of cerebellar development.

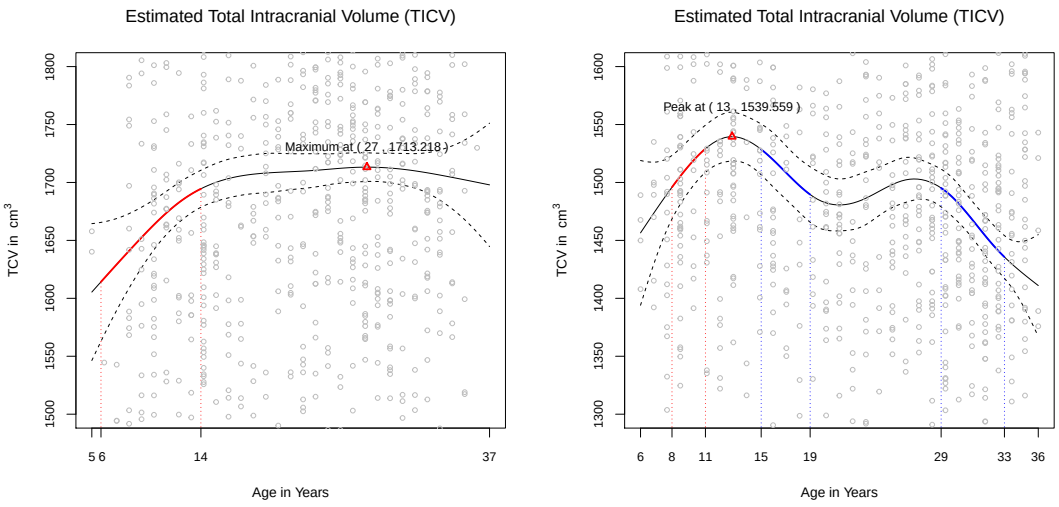

Figure 5: Estimated developmental trajectories for the total intracranial volume (TICV). Left: male subjects; Right: female subjects.

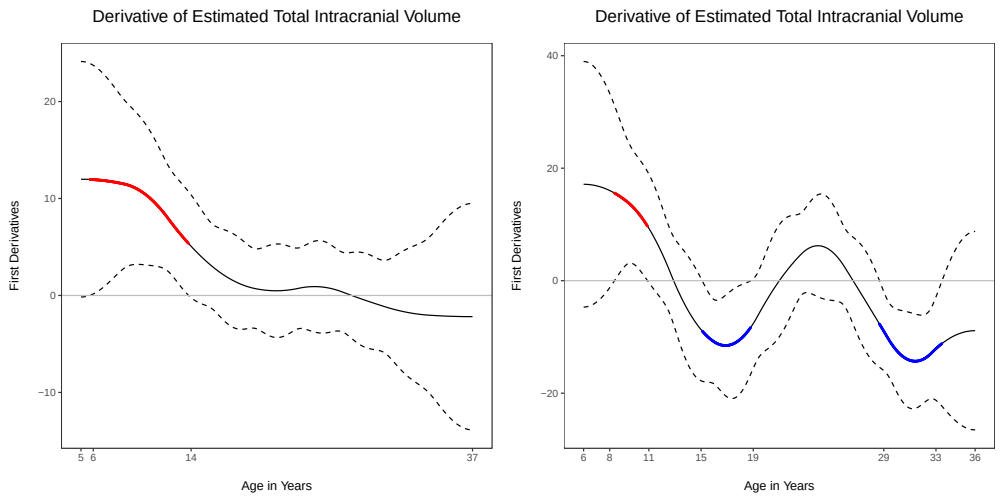

Figure 6: Derivative estimates of the developmental trajectories for total intracranial volume (TICV). Left: male subjects; Right: female subjects.

4.2 Resampling Results: Robustness Check

The proportions of identifying 2 peaks, 1 peak and the maximum in 1000 resamples are shown in Table 4, which gives us a rough idea of whether the shapes of the developmental trajectories presented in Figure 1 and 2 are robust. Cerebellar white volumes are not included since the interest is mainly on situations more complicated than a steady increase.

For Male TCV, aTCV and TICV, the trend of no peaks detected (i.e., only the maximum is reported) is retained well with 0.951, 0.953 and 1 re-identification probabilities. 95% of the resamples do not yield a peak for female aCCV, which supports the steady decreasing trend we have identified.

Compared with trends we haven't identified any peaks, those with peak(s) identified are less robust, especially when more than 1 peak is identified. The single inverse-U (i.e., 1 peak) shape of male CCV, male aCCV, and female TICV is re-identified in 54.1%, 46.5%, and 64.9% of the 1000 resamples. The trend of a connected inverse-U with 2 peaks for female TCV and CCV is only identified for 21.6% and 19.1% of the time. Since peaks are closely related to the identified periods of significant increase and decrease, we also check the detection of significantly changing periods as summarized in Table 5 (male) and Table 6 (female).

For male TCV, the significant increase on [5,12] is pretty robust with 87.8% of the resamples detecting it as a significantly increasing period. The probability of detecting any significant increase of male TCV within the limits of [5,13] even reaches 1. Half of the time, [5,11] is detected as a significantly increasing

period of male aTCV and this probability rises to 0.772 when loosening the limits. In 44.3% and 24% of the time, age 14 is identified as a peak/maximum for male TCV and aTCV. And the proportion of detecting any peak/maximum at age 13, 14, or 15 is 87.2% and 77.2% for male TCV and aTCV respectively. Generally, we conclude that the total cerebellum volume (with or without TICV adjustment) for males is highly likely to experience a significant increase within the age range of 5 to 12 and reaches its peak at around 13-15 years old. Following a similar analysis, the cerebellar cortex volume of males seems to be significantly increased in some age range within [5,13] but the significant decrease identified afterward is not very robust. For the TICV-adjusted cerebellar cortex volume of males, it's the significant increase before adolescence that is not robust, but it's likely to be decreasing significantly at some point between [14,21]. Generally, a peak/maximum arises for male cerebellar cortex volume at around 11-13 years old. For the total intracranial volume in males, a significant increase in some range within [5,15] is likely to appear with a stable level afterward (no significant peak/maximum).

For female TCV, the first significantly increasing period [6,10] is more robust than the second one identified [25,26]. While the significant decrease in [29,33] is more robust than that in [19,20] with a high proportion of 97.4% resamples detecting a significant decrease within the limits of [28, 34]. For the 2 peaks detected, the first one at 12 years old is more robust than the second (28 years). Therefore, we conclude that the total cerebellum volume for females experiences a robust and significant increase between 6 and 11 years old and a decrease within the range of [28,34], with a robust peak/maximum at around 12 years. For female CCV, the only significant period of change with a re-identification probability higher than 0.5 is [29,33] (loosening the limits gives a higher probability of 0.979). And the first peak of 11 years possesses higher robustness compared with the second one (28 years). For female TICV, it exhibits a significant increase that is fairly robust within the range of 7 to 12 years, and decreases significantly around ranges of [14,20] and [28,34] that are pretty robust, with the peak/maximum attained robustly at 13 years old.

Table 4: Stratified resampling results – shape of curve, stratified by sex.

	Male			Female		
	2 Peaks	**1 Peak**	**Maximum**	**2 Peaks**	**1 Peak**	**Maximum**
TCV						
Absolute	0	0.049[a]	0.951	0.216	0.592	0.192
Adjusted	0	0.047	0.953			
CCV						
Absolute	0	0.541	0.459	0.191	0.350	0.459
Adjusted	0	0.465	0.535	0.014	0.036	0.950
TICV	0	0	1	0.029	0.649	0.322

[a] Proportion out of the 1000 stratified resamples.

Table 5: Stratified resampling results – significantly increasing or decreasing periods, male.

	Increasing		Decreasing		Peaks/Maximum	
	Original[a]	**Fluctuated**[a]	**Original**	**Fluctuated**	**Original**	**Fluctuated**
TCV						
Absolute	[5,12] (0.878)	[5,13][b] (1)	\	\	14 (0.443)	13-15[c] (0.872)
Adjusted	[5,11] (0.545)	[5,12][b] (0.722)	\	\	14 (0.240)	13-15 (0.722)
CCV						
Absolute	[5,12] (0.384)	[5,13][b] (0.999)	[15,16] (0.244)	[14,17] (0.395)	13 (0.800)	12-14 (0.999)
Adjusted	[7,9] (0.053)	[6,10] (0.292)	[15,20] (0.191)	[14,21] (0.993)	12 (0.721)	11-13 (0.833)
TICV	[6,14] (0.081)	[5,15] (0.839)	\	\	27 (0.16)	26-28 (0.477)

[a] ”Original” refers to the significantly increasing(decreasing) period(s), or the peak(s)/maximum identified in the original data. ”Fluctuated” refers to the interval with both limits expanded 1 from the original interval (for the ”Increasing” and ”Decreasing” columns), or expanded from the original value $\mp$ 1 (for the ”Peaks/Maximum” column). The fluctuated interval refers to detecting any significantly changing period(s) within the limits of this interval, not that the fluctuated interval itself is detected as a significantly changing period.

[b] The fluctuated lower limit remains 5 since the youngest male subject is of age 5.

[c] 13-15 represents that the peak or maximum is identified at age 13, 14, or 15.

Table 6: Stratified resampling results – significantly increasing or decreasing periods, female.

	Increasing		Decreasing		Peaks/Maximum	
	Original[a]	**Fluctuated**[a]	**Original**	**Fluctuated**	**Original**	**Fluctuated**
TCV						
	[6,10] (0.542)	[6,11][b] (0.794)	[19,20] (0.137)	[18,21] (0.305)	12 (0.607)	11-13 (0.899)
Absolute	[25,26] (0.157)	[24,27] (0.262)	[29,33] (0.539)	[28,34] (0.974)	28 (0.249)	27-29 (0.297)
CCV						
			[15,17] (0.144)	[14,18] (0.305)		
	[6,10] (0.260)	[6,11][b] (0.493)	[18,20] (0.141)	[17,21] (0.354)	11 (0.738)	10-12 (0.936)
Absolute	[25,26] (0.130)	[24,27] (0.226)	[29,33](0.730)	[28,34](0.979)	28 (0.171)	27-29 (0.239)
Adjusted	\	\	[8,27] (0.029)	[7,28] (0.314)	\	\
			[15,19] (0.209)	[14,20] (0.922)		
TICV	[8,11] (0.157)	[7,12] (0.674)	[29,33] (0.479)	[28,34] (0.956)	13 (0.939)	12-14 (0.996)

[a] "Original" refers to the significantly increasing(decreasing) period(s), or the peak(s)/maximum identified in the original data. "Fluctuated" refers to the interval with both limits expanded 1 from the original interval (for the "Increasing" and "Decreasing" columns), or expanded from the original value $\mp$ 1 (for the "Peaks/Maximum" column). The fluctuated interval refers to detecting any significantly changing period(s) within the limits of this interval, not that the fluctuated interval itself is detected as a significantly changing period.

[b] The fluctuated lower limit remains 6 since the youngest female subject is of age 6.

5 Conclusions and Discussion

This paper adopts a flexible penalized cubic spline modeling framework to characterize the volumetric developmental trajectories of the cerebellum, cerebellar cortex and cerebellar white matter in healthy people, both with and without adjustment for total intracranial volume. The targeted cohort covers an age range of 5 to 37 years old across childhood, adolescence and adulthood. Compared with the traditional parametric modeling framework, a penalized spline fits for a wide variety of trends with different degrees of freedom under 1 general model, free of rigid parametric assumptions of the model form, which makes it a perfect tool to reveal the underlying growth curve of cerebellar volumes. A derivative analysis is followed which uses different quotients approximation to estimate the rate of change of the volume of interest. Variances of the derivatives are derived and further hypothesis tests are proposed to detect age periods of significant volume change. The downstream derivative analysis distinguishes our study, for it provides the significance of the increases and decreases that have appeared in the predicted trajectories, which helps detect statistically

significant age periods of volume change besides simply interpreting the general trend of growth out of the estimated trajectories. Based on the significantly increasing and decreasing periods of age identified, peak(s)/maximum of the volume of interest are given as well. Stratified resampling is performed to check the robustness of the periods of significant change and the peak(s)/maximum.

To summarize, the total cerebellum volume of males is highly likely to experience a significant increase within the age range of 5 to 12 and reaches its peak at around 13-15 years old with minimal change afterward. The cerebellar cortex volume of males seems to be significantly increased in some age range within [5,13], followed by a gradual decrease (i.e., an inverse-U shape) which is not very robust indicated by the resampling result. For the TICV-adjusted cerebellar cortex volume of males, an inverse-U shape is followed as well with a peak at 12 years old, and a significant and robust decrease within the range of [14,21]. For the total intracranial volume in males, a significant increase in some range within [5,15] is likely to appear followed by a stable level afterward (no significant peak/maximum). For the total cerebellum volume and cerebellar cortex volume (both unadjusted) of females, the predicted trajectory is in the shape of 2 inverse-U connected together (i.e., increase-decrease-increase-decrease). This kind of cerebellar developmental curve is rarely seen in previous literature. For the absolute total cerebellum volume, the significant increase happens from 6 to 10 years old as well as from 25 to 26 years. The significant decrease also appears in 2 periods which are [19,20] and [29,33]. The local maximum of the 2 inverse Us are the identified 2 peaks at age 12 and age 28 respectively. Resampling results indicate that the significant change in [6,10], [29,33] and the first peak of 12 years are relatively more robust. For the absolute cerebellar cortex volume, the 2 significant increasing periods remain as [6,10] and [25,26], and it decreases significantly on a longer period of [15,17]-[18,20]. The second significantly decreasing period is still [29,33], which is the only robust period of change. The identified peaks are reached at age 11 and 28 years old between which the peak at 11 years is more robust. However, after adjusting for TICV, aCCV is predicted to show a sustained volume loss across the age range of 6 to 36 years. The total intracranial volume in females exhibits a significant increase that is fairly robust within the range of 7 to 12 years, and decreases significantly around robust ranges of [14,20] and [28,34], with the peak/maximum attained robustly at 13 years old. The cerebellar white matter volumes for both males and females act differently in the growth pattern with a steady increase throughout the age range covered.

Based on the current data, besides the difference of the mean level of cerebellar volumes (refer to Table 2), an obvious gender dimorphism of the shape of the developmental trajectories for the cerebellum and its grey matter exist (refer to the trajectories shown in Figure 1 and 2 to see how different the predicted trends are

for male and female subjects). This dimorphism coincides with past literature on sex differences in cerebellar volume trajectories (Tiemeier et al., 2010; Bernard et al., 2015; Sussman et al., 2016).

Penalized cubic splines aided by derivative analysis for the studied cohort provide an accurate characterization of the growth of cerebellar volumes, which can help understand normal development among children, adolescents and adults. Therefore, an exciting future work can be using this accurate characterization as a reference for detecting atypical cerebellar developmental patterns among patients with certain neurodevelopmental disorders. Several other future directions are also suggested by this paper. Since the confidence intervals of the rate of change are given point-wisely, a possible improvement is to derive a simultaneous confidence interval for the first derivatives, which requires further theoretical inspection. Also if it's possible to acquire measurements of other covariates besides age such as nutrient intake, a generalized additive model (Hastie and Tibshirani, 1987) can be considered to model the effects from various factors on cerebellar volumetric development. One limitation of the study is related to merging the data from the HCP Development and HCP Young Adult studies. The age range covered in the HCP Development study is 5 to 21 years old and 22 to 37 years old for the HCP Young Adult study. It happens that the saddle points of the connected inverse-Us from female TCV and female CCV locate around the age (21 and 22 years) at which the 2 HCP studies overlap. Therefore, it's possible that merging the 2 large-scale studies partially explains the saddles of the connected inverse-Us from the developmental trajectories of female TCV and female CCV.

A Appendix

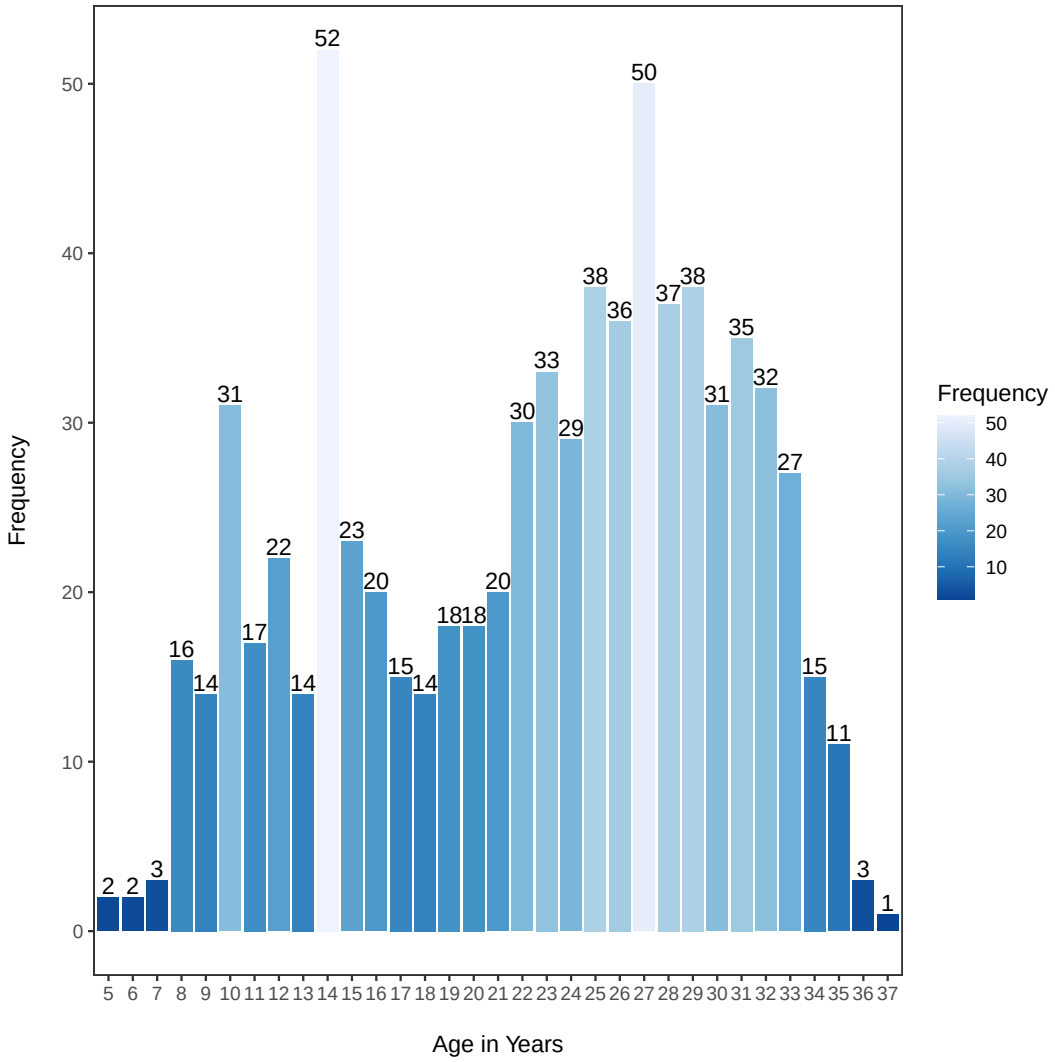

Figure A.1: Age distribution in males

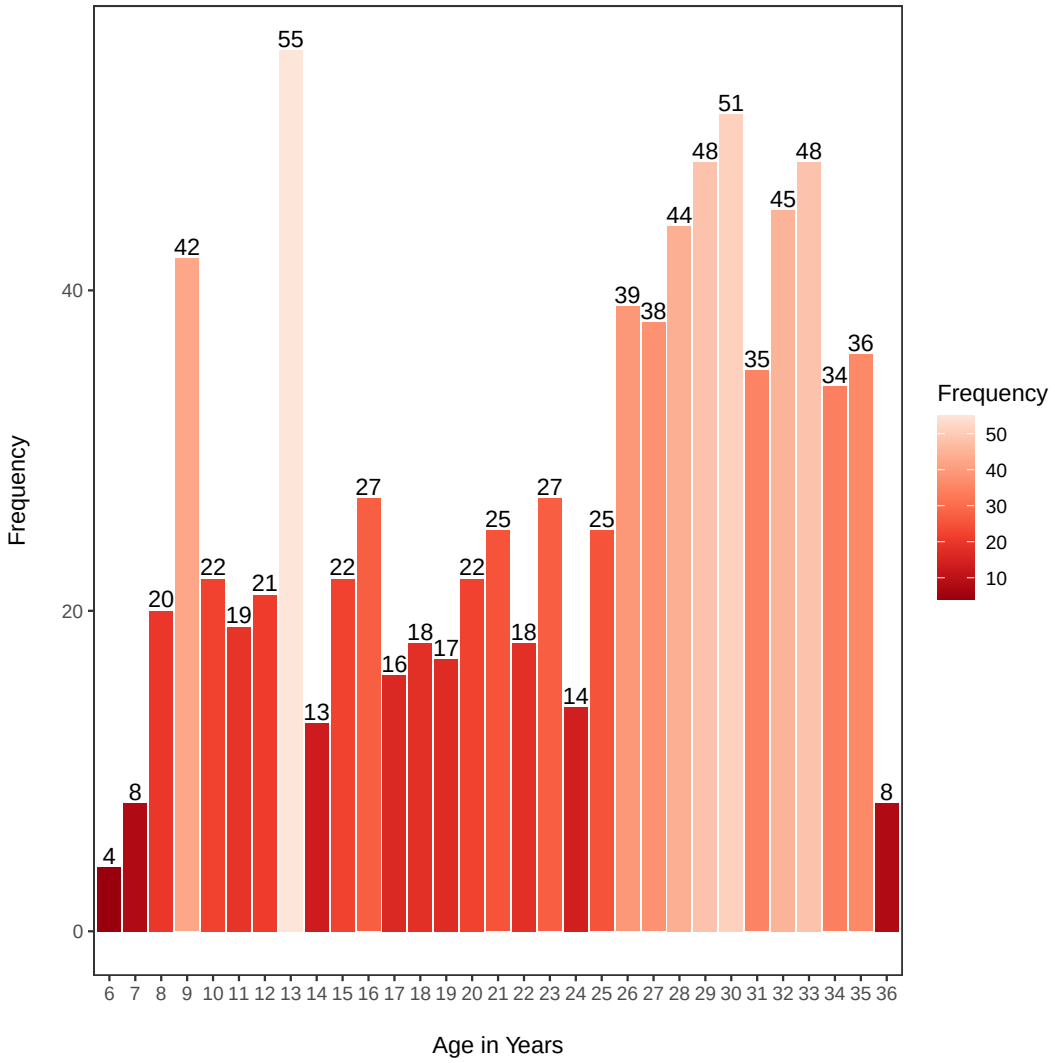

Figure A.2: Age distribution in females

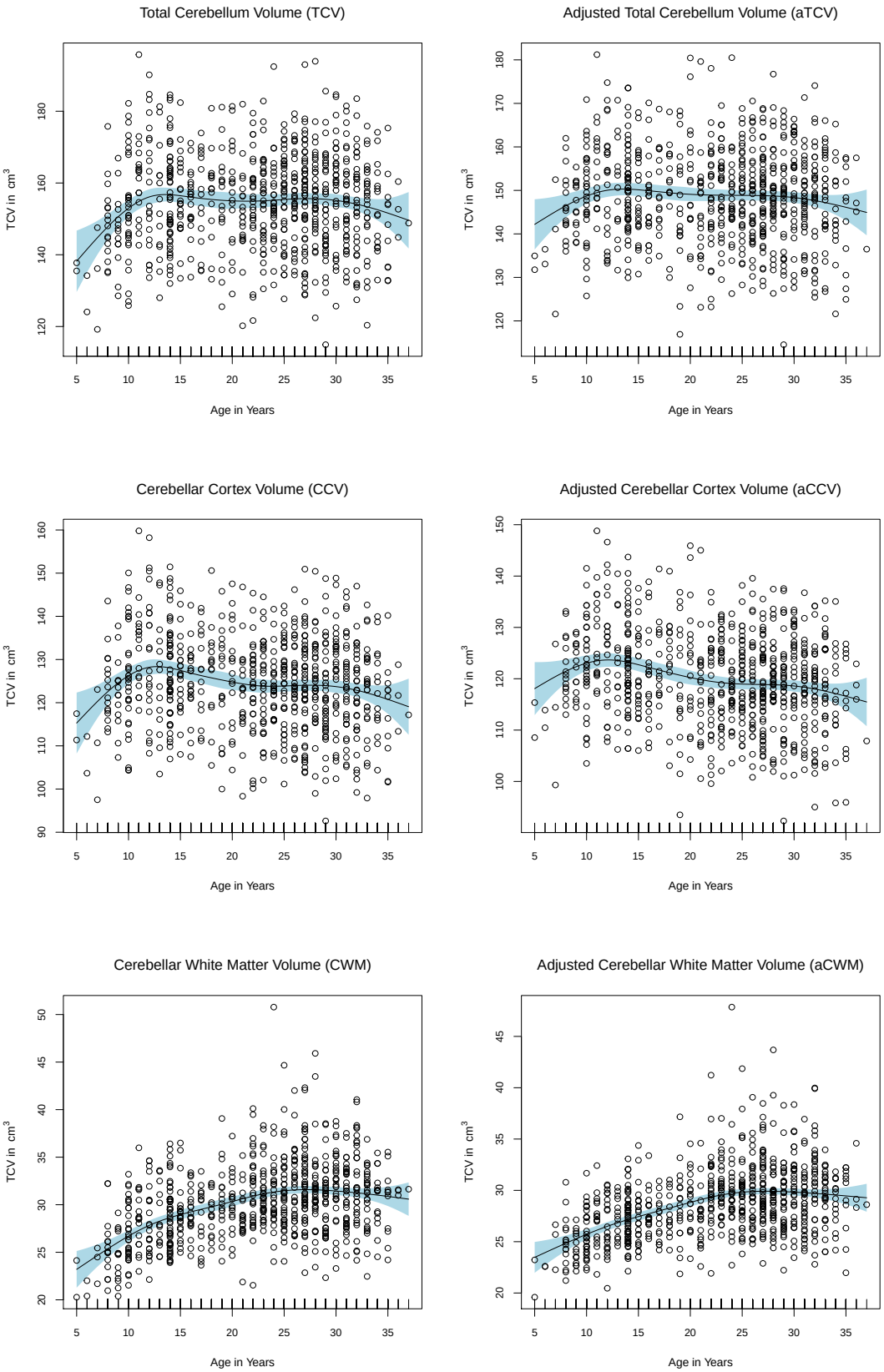

Figure A.3: Estimated developmental trajectories for the cerebellar volumes in male subjects. Black solid lines: mean volumes estimated by penalized cubic splines; Light-blue bands: point-wise Bayesian confidence intervals; Black hollow dots: original data samples.

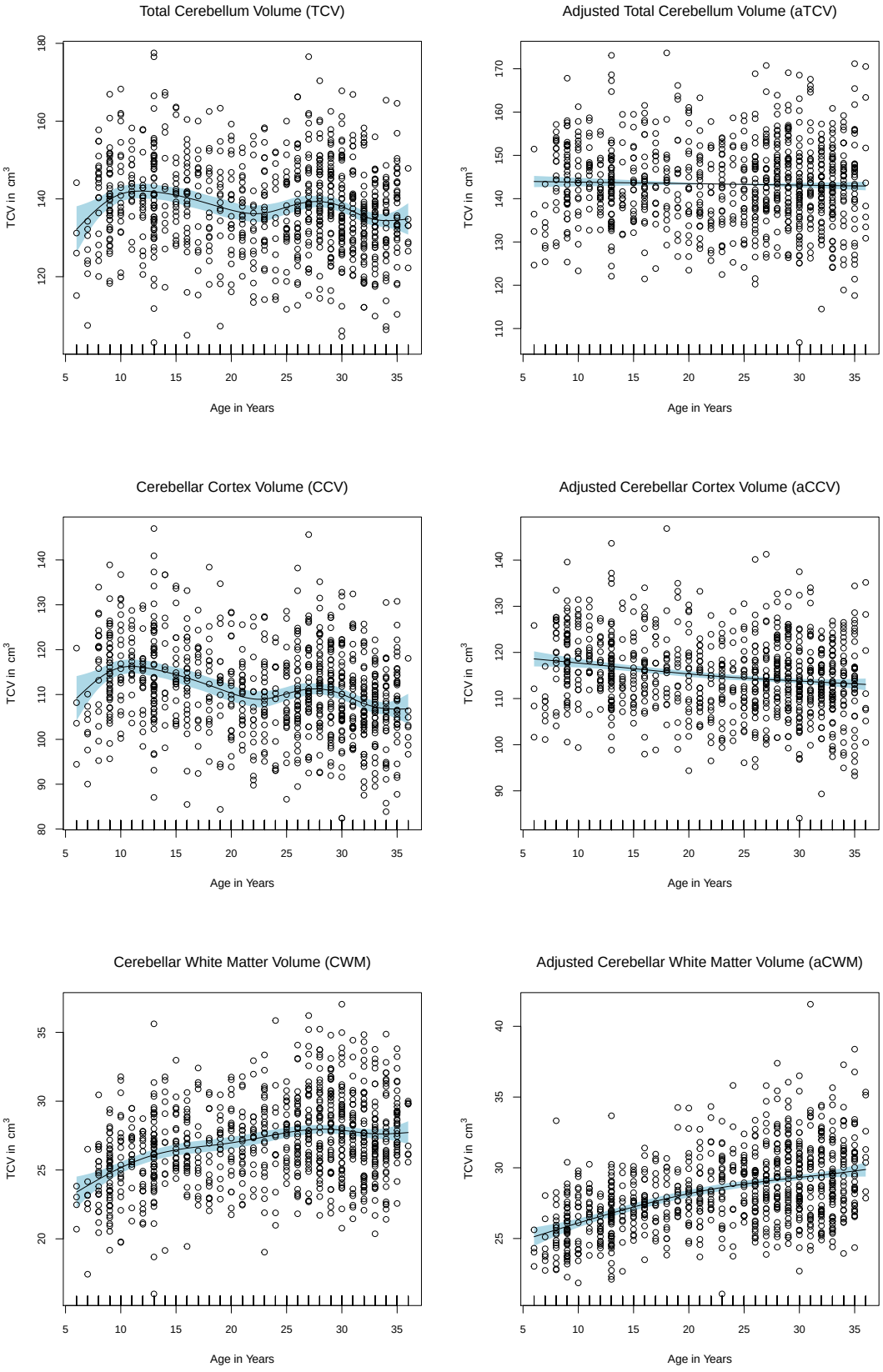

Figure A.4: Estimated developmental trajectories for the cerebellar volumes in female subjects.

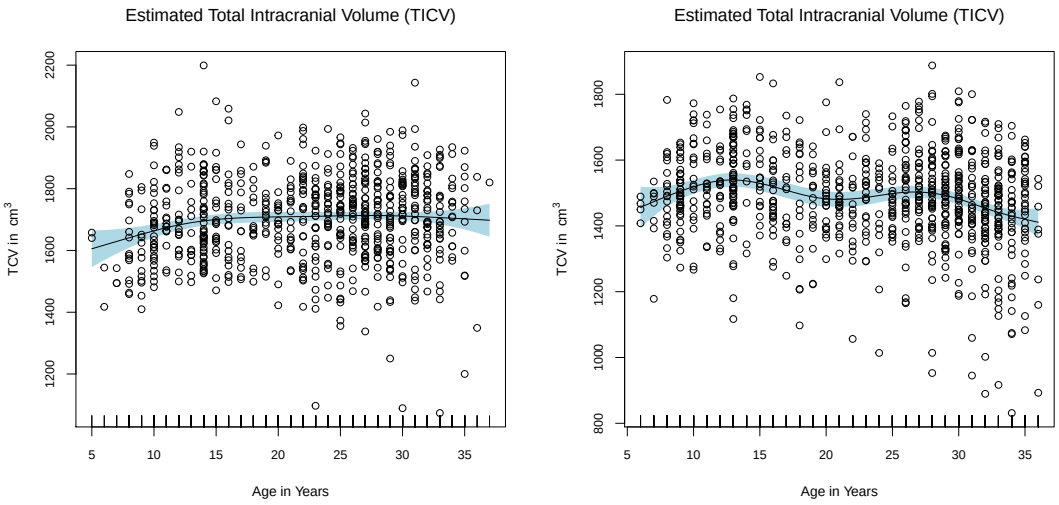

Figure A.5: Estimated developmental trajectories for the total intracranial volume (TICV). Left: male subjects; Right: female subjects.

References

Bernard, J. A., Leopold, D. R., Calhoun, V. D., and Mittal, V. A. (2015). Regional cerebellar volume and cognitive function from adolescence to late middle age. *Human brain mapping*, 36(3):1102–1120.

Caviness Jr, V., Kennedy, D., Richelme, C., Rademacher, J., and Filipek, P. (1996). The human brain age 7–11 years: a volumetric analysis based on magnetic resonance images. *Cerebral cortex*, 6(5):726–736.

Chen, S. A. and Desmond, J. E. (2005). Temporal dynamics of cerebro-cerebellar network recruitment during a cognitive task. *Neuropsychologia*, 43(9):1227–1237.

Cho, T. H., Lee, N. J., Uhm, C.-S., Kim, H., Suh, Y.-S., et al. (1999). Magnetic resonance image-based cerebellar volumetry in healthy korean adults. *Neuroscience letters*, 270(3):149–152.

De Boor, C. (1972). On calculating with b-splines. *Journal of Approximation theory*, 6(1):50–62.

Eilers, P. H., Marx, B. D., et al. (1996). Flexible smoothing with b-splines and penalties. *Statistical science*, 11(2):89–121.

Hastie, T. and Tibshirani, R. (1987). Generalized additive models: some applications. *Journal of the American Statistical Association*, 82(398):371–386.

Human Connectome Project (2018). Hcp young adult. `https://www.humanconnectome.org/study/hcp-young-adult`. [Online; accessed 23-June-2021].

Human Connectome Project (2021). Hcp development. `https://www.humanconnectome.org/study/hcp-lifespan-development`. [Online; accessed 23-June-2021].

Ilyasov, R. H. (2014). About the method of analysis of economic correlations by differentiation of spline models. *Modern Applied Science*, 8(5):197.

Iserles, A. (2009). *A first course in the numerical analysis of differential equations*. Number 44. Cambridge university press.

Leiner, H. C., Leiner, A. L., and Dow, R. S. (1993). Cognitive and language functions of the human cerebellum. *Trends in neurosciences*, 16(11):444–447.

Luft, A. R., Skalej, M., Schulz, J. B., Welte, D., Kolb, R., Bürk, K., Klockgether, T., and Voigt, K. (1999). Patterns of age-related shrinkage in cerebellum and brainstem observed in vivo using three-dimensional mri volumetry. *Cerebral Cortex*, 9(7):712–721.

Marra, G. and Wood, S. N. (2012). Coverage properties of confidence intervals for generalized additive model components. *Scandinavian Journal of Statistics*, 39(1):53–74.

Monteith, D., Evans, C., Henrys, P., Simpson, G., and Malcolm, I. (2014). Trends in the hydrochemistry of acid-sensitive surface waters in the uk 1988–2008. *Ecological Indicators*, 37:287–303.

Nychka, D. (1988). Bayesian confidence intervals for smoothing splines. *Journal of the American Statistical Association*, 83(404):1134–1143.

Raz, N., Gunning-Dixon, F., Head, D., Williamson, A., and Acker, J. D. (2001). Age and sex differences in the cerebellum and the ventral pons: a prospective mr study of healthy adults. *American Journal of Neuroradiology*, 22(6):1161–1167.

Reinsch, C. H. (1967). Smoothing by spline functions. *Numerische mathematik*, 10(3):177–183.

RStudio Team (2021). *RStudio: Integrated Development Environment for R*. RStudio, PBC, Boston, MA.

Ruppert, D., Wand, M. P., and Carroll, R. J. (2003). *Semiparametric regression*. Number 12. Cambridge university press.

Stanfield, A. C., McIntosh, A. M., Spencer, M. D., Philip, R., Gaur, S., and Lawrie, S. M. (2008). Towards a neuroanatomy of autism: a systematic review and meta-analysis of structural magnetic resonance imaging studies. *European psychiatry*, 23(4):289–299.

Stoodley, C. J., Valera, E. M., and Schmahmann, J. D. (2012). Functional topography of the cerebellum for motor and cognitive tasks: an fmri study. *Neuroimage*, 59(2):1560–1570.

Sussman, D., Leung, R. C., Chakravarty, M. M., Lerch, J. P., and Taylor, M. J. (2016). The developing human brain: age-related changes in cortical, subcortical, and cerebellar anatomy. *Brain and behavior*, 6(4):e00457.

Szabó, C. Á., Lancaster, J. L., Xiong, J., Cook, C., and Fox, P. (2003). Mr imaging volumetry of subcortical structures and cerebellar hemispheres in normal persons. *American Journal of Neuroradiology*, 24(4):644–647.

Tiemeier, H., Lenroot, R. K., Greenstein, D. K., Tran, L., Pierson, R., and Giedd, J. N. (2010). Cerebellum development during childhood and adolescence: a longitudinal morphometric mri study. *Neuroimage*, 49(1):63–70.

Valera, E. M., Faraone, S. V., Murray, K. E., and Seidman, L. J. (2007). Meta-analysis of structural imaging findings in attention-deficit/hyperactivity disorder. *Biological psychiatry*, 61(12):1361–1369.

Wahba, G. (1980). *Spline bases, regularization, and generalized cross validation for solving approximation problems with large quantities of noisy data*. University of WISCONSIN.

Wahba, G. (1983). Bayesian "confidence intervals" for the cross-validated smoothing spline. *Journal of the Royal Statistical Society: Series B (Methodological)*, 45(1):133–150.

Wikipedia contributors (2020). Smoothing spline — Wikipedia, the free encyclopedia. `https://en.wikipedia.org/w/index.php?title=Smoothing_spline&oldid=956810800`. [Online; accessed 24-June-2021].

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(1):3–36.

Wood, S. N. (2013). On p-values for smooth components of an extended generalized additive model. *Biometrika*, 100(1):221–228.

Wood, S. N. (2017). *Generalized additive models: an introduction with R*. CRC press.

Wood, S. N. and Wood, M. S. (2015). Package 'mgcv'. *R package version*, 1:29.

Zhou, S. and Wolfe, D. A. (2000). On derivative estimation in spline regression. *Statistica Sinica*, pages 93–108.