Gray matter volume predicts individual body mass index and its development during adolescence

HAIYAN WANG (wanghaiyan2015@ia.ac.cn)

Brainnetome Center and National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China; School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China.

TIANZI JIANG (jiangtz@nlpr.ia.ac.cn)

Brainnetome Center and National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China; School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China; CAS Center for Excellence in Brain Science and Intelligence Technology, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China; The Clinical Hospital of Chengdu Brain Science Institute, MOE Key Lab for Neuroinformation, University of Electronic Science and Technology of China, Chengdu 625014, China; The Queensland Brain Institute, University of Queensland, Brisbane, QLD 4072, Australia.

Abstract: Adolescent obesity is one of the most important current public health concerns, owing to its increased prevalence and adverse effects on physical and mental health. Body mass index (BMI) is a measure of obesity, and relationships between brain and BMI have been found based on univariate association analyses. However, whether/how neuroanatomical features can be used to predict the BMI and its development at the individual level during adolescence are unclear. Here, we analyzed the large-scale longitudinal IMAGEN dataset, in which structural magnetic resonance imaging and BMI were acquired at both 14 and 19 years old in the same subjects. Using the voxel-wise gray matter volume (GMV) as features and the multivariate machine learning method, we constructed predictive models for individually predicting the BMI at both 14 and 19 years old, as well as the longitudinal development of BMI between the 2 ages. We found that, the whole-brain GMV could predict the individual BMI at both 14 and 19 years old, and the development of GMV in cerebellum could predict the individual development of BMI. The contributing brain regions for predicting 14- and 19-year-old BMIs did not differ at a coarse scale, but exhibited considerable differences at a fine scale. Our results highlight the importance of GMV in predicting the individual cross-sectional BMI and its longitudinal development during adolescence.

CCS CONCEPTS • Computing methodologies • Machine learning • Cross-validation

Additional Keywords and Phrases: BMI, Individualized prediction, Adolescence, Development, Gray matter volume

1 INTRODUCTION

The prevalence of overweight and obesity in adolescents has greatly increased in recent decades, especially in developing countries [1-3]. Adolescent obesity has both short-term and long-term adverse effects on health. Specifically, obese and overweight teens tend to have a lower level of life satisfaction, have more difficulty in making new friends, and also report more alcohol use [4]. Overweight in adolescents can also result in a variety of comorbidities, such as type 2 diabetes, obstructive sleep apnea, hypertension, dyslipidemia, and the metabolic syndrome [5]. Moreover, overweight and obesity during adolescence is related to subsequent overweight and obesity in adulthood [6] and increases the risk of adult morbidity and mortality [7, 8]. Body mass index (BMI) is a measure that is often used to study obesity. As such, a better understanding of neural mechanisms of BMI and its development during adolescence is needed.

Previous investigations of the neural mechanisms of obesity primarily focused on finding brain regions associated with obesity. Cross-sectional studies have found that obesity was associated with decreased gray matter volume (GMV) in both adolescents [9] and adults [10]. Longitudinal studies have revealed that increases in the volumes of the right hippocampus and parahippocampal gyrus were associated with lower BMI increases during early adolescence [11], and the change of BMI during old adulthood was related to the change in hippocampal volume [12]. However, these group-level association analyses obscured meaningful individual variations in adolescents and their development [13], whereas individualized prediction model can afford us the opportunity to identify novel treatment targets and help to individually tailor the course of interventions [14]. Structural magnetic resonance imaging (MRI) has been shown to predict individual BMI in adults [15]. However, it is unknown whether the structural MRI could also predict the individual BMI and its development during adolescence. Moreover, it is not clear whether there are differences in the neural mechanisms of obesity between mid- to late-adolescence.

In this study, the large-scale multi-site longitudinal IMAGEN dataset, in which both structural MRI and BMI were acquired at both 14 (baseline) and 19 years old (follow-up), were utilized to build predictive models to predict the BMI and its development during adolescence at the individual level. Specifically, elastic-net regression analyses were used to model the relationship between the whole-brain voxel-wise GMV and BMI in both baseline and follow-up, as well as the development of BMI. To ensure the generalizability of these predictive models, a nested leave-one-site-out cross-validation was applied. Moreover, voxel-wise GMV in each brain region were also separately used to conduct these predictions for testing each region's contribution in the prediction.

2 MATERIALS AND METHODS

2.1 Participants

The IMAGEN dataset is a multi-site (8 sites) longitudinal dataset, which has recruited over 2000 adolescents of about 14 years old across France, the UK, Ireland and Germany [16]. Behavioral, cognitive, and neuroimaging data were acquired, and the participants were followed longitudinally. The study was approved by each of the local ethics committee. Verbal assent was obtained from each participant, and written consent was obtained from their legal guardians. For our analyses, participants with complete demographic information, BMI, and structural MRI data at both 14 and 19 years old were included, generating 970 subjects as shown in Table 1.

Table 1: Demographic information of the participants

	Baseline[a]	Follow-up[a]	Development[a]
Age (years)	14.49 (0.42)	19.42 (1.08)	4.93 (1.07)[c]
Gender	Female/male = 507/463		
Handedness	Left/right/both = 103/860/7		
Site	121/137/79/56/130/113/186/148		
vIQ[b]	81.81 (11.54)	-	-
pIQ[b]	108.96 (13.27)	-	-
TIV (ml)	1416 (125)	1420 (134)	4 (41)[c]
BMI (kg/m^2)	20.56 (3.1)	22.64 (3.8)	2.08 (2.68)[c]

- = Not applicable. [a]Mean (standard deviation). [b]vIQ, verbal IQ; pIQ, performance IQ. The verbal IQ and performance IQ were determined using the Wechsler Intelligence Scale for children, and they are raw values without age-normalization. [c]The change of an index from baseline to follow-up. TIV, total intracranial volume; BMI, body mass index.

2.2 MRI Acquisition

High-resolution anatomical MRI scanning was performed with 3 Tesla scanners from various manufactures (Siemens, Philips, General Electric and Bruker), using a standardized 3D T1-weighted magnetization prepared rapid acquisition gradient echo (MRPAGE) sequence based on the ADNI protocol (http://adni.loni.usc.edu/methods/mri-analysis/mri-acquisition/). The scanning parameters and sequence protocol were specifically chosen to be compatible with all scanners (sagittal slice plane, repetition time = 2,300 ms, echo time = 2.8 ms, flip angle = 8^o, 256 x 256 x 170 matrix, 1.1 x 1.1 x 1.1 mm voxel size).

2.3 MRI Processing

The structural data were first manually checked for quality and data with bad quality were removed. Here, we used the VBM8 toolbox in SPM8 (Statistical Parametric Mapping, https://www.fil.ion.ucl.ac.uk/spm/) with default parameters to preprocess the structural data at both baseline and follow-up. Briefly, T1-weighted images were segmented and spatially normalized with an already integrated Dartel template in MNI space. All images were then subjected to nonlinear modulations and corrected for each individual's head size. Finally, the resulting images were smoothed with a Gaussian kernel of 8 mm full width at half maximum, and were resliced to voxel size of 3 mm^3. Although the head size was corrected for, the total intracranial volume (TIV) was further used as a nuisance covariate. The gray matter images were masked with the Brainnetome atlas [17] to generate features in the following prediction processes.

2.4 Prediction Processes

2.4.1 Elastic-Net Penalized Linear Regression.

In our work, we tried to predict each individual's BMI utilizing the voxel-wise GMV as features. We performed 3 separate predictions: 1) using baseline voxel-wise GMV to predict the baseline BMI (baseline prediction); 2) using follow-up GMV to predict the follow-up BMI (follow-up prediction); 3) using the development of GMV ($\frac{follow\text{-}up\ GMV\ -\ baseline\ GMV}{baseline\ GMV}$) to predict the development of BMI ($\frac{follow\text{-}up\ BMI\ -\ baseline\ BMI}{baseline\ BMI}$) (development prediction). The baseline and follow-up predictions were cross-sectional predictions, whereas the development prediction was a longitudinal prediction. Given the high-dimensional features, we used the elastic-net penalized linear regression to model the relationship between the features and BMI. The elastic-

net regression is a combination of ridge and lasso regression. It allows for variable selection and assigns similar weights to the correlated features, which ensured its interpretability and stability [18] [19]. The lasso function (which implements lasso, ridge, as well as elastic-net regressions) in Matlab was used to conduct the elastic-net regression.

Specifically, the elastic-net regression model can be formulized as follows:

$$Y_{n\times 1} = X_{n\times m}\beta_{m\times 1} + \varepsilon_{n\times 1}$$

where $X_{n\times m}$ denotes the voxel-wise GMV features, *n* is the number of subjects, and *m* is the number of voxels used; $Y_{n\times 1}$ is the BMI of the *n* subjects; $\beta_{m\times 1}$denotes the regression coefficients; and $\varepsilon_{n\times 1}$ is the regression residual. And the estimates $\hat{\beta}$ from the elastic net method are defined as below:

$$\hat{\beta} = \arg\min_{\beta}\left\{\frac{1}{2n}\|Y_{n\times 1} - X_{n\times m}\beta_{m\times 1}\|_2^2 + \lambda\left(\alpha\|\beta_{m\times 1}\|_1 + \frac{1-\alpha}{2}\|\beta_{m\times 1}\|_2^2\right)\right\}$$

where α is a mixing parameter that controls the relative weighting of the L1-norm and L2-norm contributions, and λ is the regularization parameter that controls the amount of shrinkage on the regression coefficients.

2.4.2 Nested Leave-One-Site-Out Cross Validation.

To ensure the generalizability of the prediction model, we implemented the nested leave-one-site-out cross validation (as shown in the Figure 1). The outer loop was used to test the model in previously unseen subjects, and the inner loop was used to determine the optimal α and λ for the elastic-net regression model.

The outer loop: In each of the outer loop, data from one of the 8 sites iteratively served as the held-out testing set, whereas data from the other 7 sites served as the training set. Each voxel-wise GMV feature in both training and testing set was normalized using the mean and standard deviation values derived from the training set. The elastic-net regression model was trained based on the training set, and was then applied to the testing set. The partial correlation (r) between the actual and predicted values regressing out the effect of age (baseline age for the baseline prediction; follow-up age for the follow-up prediction; baseline age and Δage for the development prediction), gender, handedness, verbal IQ, performance IQ, and TIV (baseline TIV for the baseline prediction; follow-up TIV for the follow-up prediction; TIV and ΔTIV for the development prediction) on the testing set was used to quantify the prediction accuracy. The r values were then averaged across all the 8 testing sites to produce the final prediction performance.

The inner loop: The inner loop was used to determine the optimal hyper-parameters in each fold of the outer cross-validation. Specifically, in each of the inner loop, data from each of the 7 sites (the outer training set) were iteratively used as the inner testing set, and data from the 6 remaining sites were used to train the model under a certain (α, λ) set. For the choices of α and λ, we applied a grid search: the α was chosen from [0.2, 0.9] with a step of 0.1, and the λ was set as $\lambda = e^{\gamma}$, where γ was chosen from [-6, 0] with a step of 0.5, generating 104 groups of parameter sets. For each set of (α, λ), we can get the averaged prediction accuracy across the 7 sites. The (α, λ) set giving the best prediction performance was used on the outer training set.

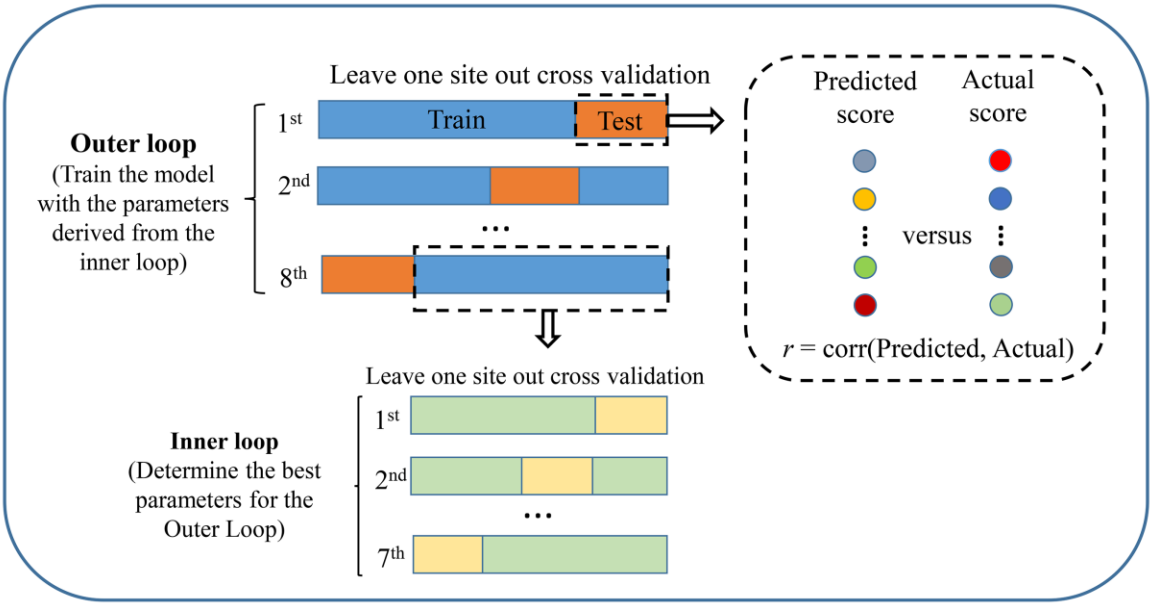

Figure 1: The nested leave-one-site-out cross validation.

2.4.3 Prediction Models.

The whole brain voxel-wise GMV features masked with the Brainnetome atlas [17] were used to separately conduct the baseline, follow-up, and development predictions. In addition, the whole brain can be divided into 8 brain regions, including frontal, temporal, parietal, insular, limbic, occipital lobes, subcortical nuclei, and cerebellum. To evaluate the contribution of each brain region in predicting the BMI or its development, we also separately used the voxel-wise GMV from each of the 8 regions as features to conduct the 3 predictions.

2.5 Contributing GM Voxels

The outer loop of the leave-one-site-out cross-validation generated 8 models, and we averaged the regression coefficients across the 8 models, producing a weight for each GM voxel. The GM voxels with a nonzero weight can be deemed as the contributing voxels for the prediction. Notably, the L1-norm penalization of elastic-net algorithm tends to select only a representative voxel from the highly correlated voxels. So, for voxels with a zero weight but showing a tight correlation ($r > 0.95$) with nonzero weighted voxels, we further weighted them according to its highly correlated nonzero weighted voxels [20]. Given the very discrete distributions of the final nonzero weighted voxels, we set a cluster-size threshold of 20 voxels.

3 RESULTS

3.1 The Development of BMI during Adolescence

From baseline to follow-up, the BMI increased significantly at the group level (t_{969} = -24.11, $P < 10^{-10}$) (Figure 2a). However, there were individual differences in the development of BMI (Figure 2b), with 82.68% of subjects showing a positive ΔBMI (follow-up BMI – baseline BMI), 16.8% of subjects showing a negative

ΔBMI, and 0.52% of subjects showing a zero ΔBMI. We then examined the demographic effects on the development of BMI. The BMI increased more in males than in females (t_{968} = 3.09, P = 0.002) from baseline to follow-up. Baseline verbal IQ showed a negative correlation with ΔBMI (r_{968} = -0.084, P = 0.009), whereas performance IQ was not significantly correlated with ΔBMI (r_{968} = -0.04, P = 0.22). Baseline TIV (r_{968} = 0.097, P = 0.003) and ΔTIV（r_{968} = 0.11, $P = 3.87\times10^{-4}$） were both significantly correlated with ΔBMI.

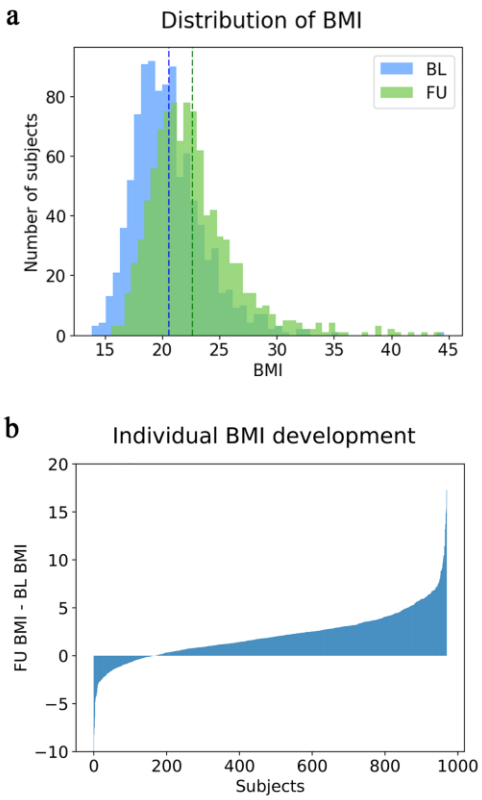

Figure 2: The development of BMI from baseline to follow-up. BL, baseline. FU, follow-up.

3.2 Prediction Results Using Whole Brain GMV

Using the whole brain voxel-wise GMV as features, we separately conducted the baseline prediction (using baseline GMV to predict the baseline BMI), the follow-up prediction (using follow-up GMV to predict the follow-up BMI), and the development prediction (using the development of GMV to predict the development of BMI). The mean prediction results across all sites were shown in Figure 3a. For the baseline prediction, the mean prediction accuracy was r = 0.42±0.09, with statistically significant partial correlations (P values < 0.05) between predicted and actual scores achieved in all of the 8 sites (Figure 3b). For the follow-up prediction, the mean performance was r = 0.38±0.12，also with statistically significant performance (P values < 0.05) in all of the 8 sites (Figure 3b). We compared the prediction performance between the baseline and follow-up predictions across all the 8 sites using a paired t-test, and did not find a significant difference between them (t_7= 1, P = 0.35). For the development prediction, the mean performance was r = 0.15±0.1, with 4 of the 8 sites showing significant prediction performance.

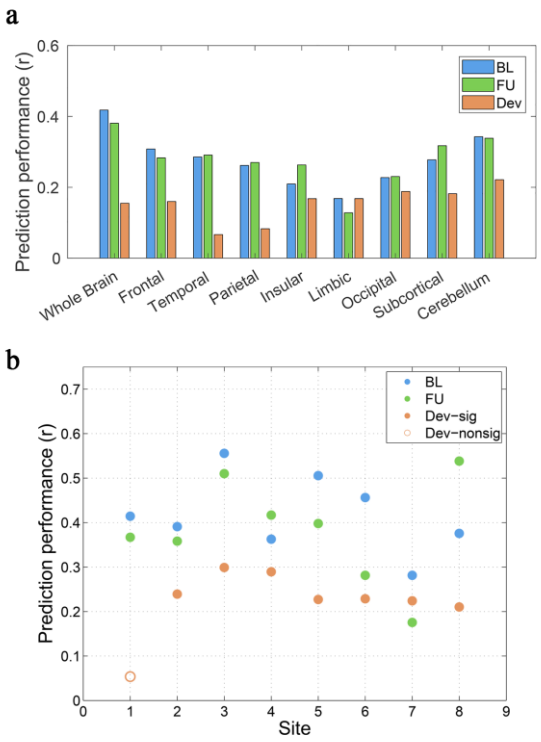

Figure 3: Prediction results with voxel-wise GMV as features. a. The mean prediction results across all sites. b. Prediction results for each acquisition site. The baseline and follow-up predictions show prediction results with whole brain GMV as features, while the development prediction shows prediction results with cerebellum GMV as features. BL, baseline; FU, follow-up; Dev, development; Dev-sig, sites with significant prediction performance; Dev-nonsig, sites with non-significant prediction performance.

3.3 Prediction Results of Regional GMV

In addition to the whole brain voxel-wise GMV, we separately used the voxel-wise GMV from each of the 8 brain regions as features to conduct the same prediction processes. For the baseline and follow-up predictions, the regional voxel-wise GMV performed worse than the whole brain voxel-wise GMV (Figure 3a). There were no significant difference between the baseline and follow-up prediction performance for each brain region (paired *t*-test, all *P*s > 0.18). The first 4 regions showing the best performance were the cerebellum, frontal lobe, temporal lobe and subcortical nuclei in both baseline and follow-up predictions. For the development prediction, GMV in cerebellum showed better performance (r = 0.22±0.07) than the whole-brain GMV, with 7 of the 8 sites had significant partial correlations (*P* values < 0.05) between the actual and predicted scores (Figure 3b). The first 4 regions showing the best performance in the development prediction were the cerebellum, occipital lobe, subcortical nuclei, and limbic region.

3.4 Contributing GM Voxels

The contributing GM voxels for predicting baseline and follow-up BMI, as well as the development of BMI involved widespread regions, as shown in Figure 4. We sorted clusters according to the absolute regression

coefficient value of the cluster peak voxel. For the baseline prediction, the most important 10 clusters located in lateral occipital gyrus, cerebellum, globus pallidus, middle frontal gyrus, cingulate gyrus, precentral gyrus, inferior parietal lobule, and postcentral gyrus. For follow-up prediction, the most important 10 clusters located in cerebellum, inferior temporal gyrus, thalamus, globus pallidus, medioventral occipital cortex, inferior frontal gyrus, inferior parietal lobule, and precentral gyrus. For the development prediction, the most important 10 clusters located in cerebellum, hippocampus, lateral occipital gyrus, superior temporal gyrus, medioventral occipital cortex, orbital gyrus, and inferior parietal lobule.

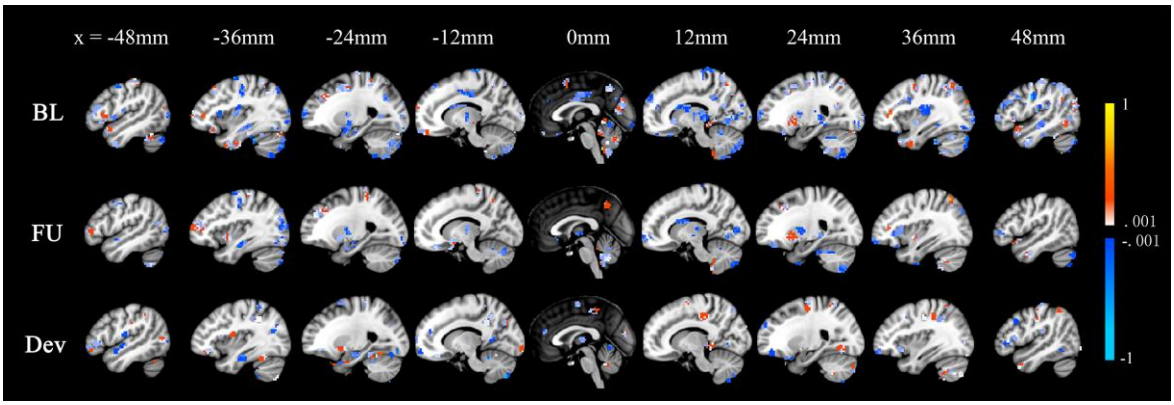

Figure 4: Contributing voxels in each prediction model using the whole brain voxel-wise GMV as features. BL, baseline; FU, follow-up; Dev, development.

4 DISCUSSION AND CONCLUSION

In this work, we found that the BMI increased from mid- to late-adolescence at the group level, and there were individual differences in the development of BMI during this period. GMV could predict the individual BMI at both mid- and late-adolescence, and the development of GMV could also predict the development of BMI at the individual level. These prediction models generalized well to previously unseen subjects from different acquisition sites. The identified GMV consisted of a wide range of regions, and the contributions of each brain region did not differ at a coarse scale, but had considerable differences at a finer scale between the baseline and follow-up predictions. In the cross-sectional predictions, the prediction based on whole-brain GMV performed better than that based on regional GMV, whereas in the longitudinal prediction the cerebellum GMV gave better prediction accuracy than that based on whole-brain GMV. These illustrated a more localized neural mechanism in the development of BMI than in the BMI .Our results may help guide early interventions for obesity in adolescents.

ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China (Grant No. 2017YFB1002502). We thank the IMAGEN consortium for providing the IMAGEN dataset.

REFERENCES

[1] Abarca-Gómez L, Abdeen ZA, Hamid ZA, Abu-Rmeileh NM, Acosta-Cazares B, Acuin C*, et al.* (2017) Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a

pooled analysis of 2416 population-based measurement studies in 128· 9 million children, adolescents, and adults. *The Lancet* 390(10113):2627-2642.

[2] Ma S, Hou D, Zhang Y, Yang L, Sun J, Zhao M*, et al.* (2020) Trends in abdominal obesity among Chinese children and adolescents, 1993–2015. *J. Pediatr. Endocrinol. Metab.* 1(ahead-of-print).

[3] Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C*, et al.* (2014) Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The lancet* 384(9945):766-781.

[4] Fonseca H, Matos MG, Guerra A, & Pedro JG (2009) Are overweight and obese adolescents different from their peers? *Int. J. Pediatr. Obes.* 4(3):166-174.

[5] Daniels SR, Arnett DK, Eckel RH, Gidding SS, Hayman LL, Kumanyika S*, et al.* (2005) Overweight in children and adolescents: pathophysiology, consequences, prevention, and treatment. *Circulation* 111(15):1999-2012.

[6] Guo SS, Wu W, Chumlea WC, & Roche AF (2002) Predicting overweight and obesity in adulthood from body mass index values in childhood and adolescence. *The American journal of clinical nutrition* 76(3):653-658.

[7] Must A, Jacques PF, Dallal GE, Bajema CJ, & Dietz WH (1992) Long-term morbidity and mortality of overweight adolescents: a follow-up of the Harvard Growth Study of 1922 to 1935. *N. Engl. J. Med.* 327(19):1350-1355.

[8] Reilly JJ & Kelly J (2011) Long-term impact of overweight and obesity in childhood and adolescence on morbidity and premature mortality in adulthood: systematic review. *Int. J. Obes.* 35(7):891-898.

[9] Alosco ML, Stanek KM, Galioto R, Korgaonkar MS, Grieve SM, Brickman AM*, et al.* (2014) Body mass index and brain structure in healthy children and adolescents. *Int. J. Neurosci.* 124(1):49-55.

[10] Dekkers IA, Jansen PR, & Lamb HJ (2019) Obesity, brain volume, and white matter microstructure at MRI: a cross-sectional UK biobank study. *Radiology* 291(3):763-771.

[11] Hashimoto T, Takeuchi H, Taki Y, Yokota S, Hashizume H, Asano K*, et al.* (2015) Increased posterior hippocampal volumes in children with lower increase in body mass index: a 3-year longitudinal MRI study. *Dev. Neurosci.* 37(2):153-160.

[12] Bobb JF, Schwartz BS, Davatzikos C, & Caffo B (2014) Cross‐sectional and longitudinal association of body mass index and brain volume. *Hum. Brain Mapp.* 35(1):75-88.

[13] Foulkes L & Blakemore SJ (2018) Studying individual differences in human adolescent brain development. *Nat. Neurosci.* 21(3):315-323.

[14] Rosenberg MD, Casey BJ, & Holmes AJ (2018) Prediction complements explanation in understanding the developing brain. *Nature communications* 9(1):1-13.

[15] Vakli P, Deák-Meszlényi RJ, Auer T, & Vidnyánszky Z (2020) Predicting Body Mass Index From Structural MRI Brain Images Using a Deep Convolutional Neural Network. *Front. Neuroinform.* 14:10.

[16] Schumann G, Loth E, Banaschewski T, Barbot A, Barker G, Büchel C*, et al.* (2010) The IMAGEN study: reinforcement-related behaviour in normal brain function and psychopathology. *Mol. Psychiatry* 15(12):1128-1139.

[17] Fan L, Li H, Zhuo J, Zhang Y, Wang J, Chen L*, et al.* (2016) The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture. *Cereb. Cortex* 26(8):3508-3526.

[18] Moradi E, Khundrakpam B, Lewis JD, Evans AC, & Tohka J (2017) Predicting symptom severity in autism spectrum disorder based on cortical thickness measures in agglomerative data. *Neuroimage* 144:128-141.

[19] Sui J, Jiang R, Bustillo J, & Calhoun V (2020) Neuroimaging-based individualized prediction of cognition and behavior for mental disorders and health: Methods and promises. *Biol. Psychiatry*.

[20] Cui Z, Su M, Li L, Shu H, & Gong G (2017) Individualized Prediction of Reading Comprehension Ability Using Gray Matter Volume. *Cereb. Cortex*:1-17.