BMJ Open

The Westlake BioBank for Chinese (WBBC) pilot cohort: a prospective study for the late adolescence

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-045564
Article Type:	Cohort profile
Date Submitted by the Author:	05-Oct-2020

Complete List of Authors:	Zhu, Xiao-Wei; Westlake University, Diseases & Population (DaP) Geninfo Lab; Fudan University, School of Life Sciences Liu, Ke-Qi ; Jiangxi Medical College Wang, Ping-Yu Liu, Jun-Quan Chen, Jin-Yang Xu, Xue-Jin; Jiading District Center for Disease Control and Prevention Xu, Jin-Jian; Jiangxi Medical College Qiu, Mo-Chang; Jiangxi Medical College Sun, Yi Liu, Chu Bai, Wei-Yang; Westlake University, Diseases & Population (DaP) Geninfo Lab; Fudan University, School of Life Sciences Zhao, Pian-Pian; Westlake University School of Life Sciences, Diseases & Population (DaP) Geninfo Lab Xia, Jiangwei; Westlake University, Diseases & Population (DaP) Geninfo Lab Gai, Si-Rui ; Westlake University, Diseases & Population (DaP) Geninfo Lab Guan, Peng-Lin; Westlake University, Diseases & Population (DaP) Geninfo Lab Qian, Yu; Westlake University, Diseases & Population (DaP) Geninfo Lab Cong, Pei-Kuan Cong; Westlake University, Diseases & Population (DaP) Geninfo Lab Xie, Shu-Yang ; Binzhou Medical University Zheng, Hou-Feng; Westlake University,

Keywords:	EPIDEMIOLOGY, OBSTETRICS, NUTRITION & DIETETICS, PAEDIATRICS

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

The Westlake BioBank for Chinese (WBBC) pilot cohort: a prospective study for the late adolescence

Xiao-Wei Zhu[1,2,3], Ke-Qi Liu[4], Ping-Yu Wang[5], Jun-Quan Liu[6,7], Jin-Yang Chen[7], Xue-Jin Xu[8], Jin-Jian Xu[4], Mo-Chang Qiu[4], Yi Sun[6], Chun Liu[9], Wei-Yang Bai[1,2,3], Pian-Pian Zhao[1,2], Jiangwei Xia[1,2], Si-Rui Gai[1,2], Peng-Lin Guan[1,2], Yu Qian[1,2], Pei-Kuan Cong[1,2], Shu-Yang Xie[5], Hou-Feng Zheng[1,2,3*]

[1] Disease & Population (DaP) Geninfo Lab, School of Life Sciences, Westlake University, Hangzhou, Zhejiang, 310024, China

[2] Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Westlake University, Hangzhou, Zhejiang, 310024, China

[3] School of Life Sciences, Fudan University, Shanghai 200433, China

[4] Jiangxi Medical College, Shangrao, Jiangxi, China

[5] Department of Epidemiology & Department of Biochemistry and Molecular Biology, Binzhou Medical University, Yantai, Shandong 264003, China

[6] Hangzhou Kingmed Diagnostics Co., Ltd., Hangzhou, Zhejiang, 310053, China

[7] Research and Development Department, Zhejiang Healthfuture Institute for Cell-Based Applied Technology, Hangzhou, Zhejiang, China

[8] Center for Disease Control and Prevention of Jiading District, Shanghai, China

[9] Department of Health Management, Hangzhou Normal University, Hangzhou, Zhejiang, 311121, China

*Corresponding author: Hou-Feng Zheng. E-mail: zhenghoufeng@westlake.edu.cn

ABSTRACT

Purpose: The Westlake BioBank for Chinese (WBBC) pilot cohort is a population-based prospective study with its major purpose to better understand the effect of genetic and environmental factors on growth and development from adolescents to adults. The dataset comprises a wide range of demographics, anthropometric measures, physical activity, sleep quality, age at menarche and bone mineral density.

Participants: A total of 14,726 participants (4,751 males and 9975 females) aged 14–25 years were recruited and the baseline survey was carried out from 2017 to 2019. The pilot cohort contains rich range of information regarding of demographics, anthropometric measurements, lifestyle and sleep patterns, clinical and health outcomes.

Findings to date: The mean age of study sample were 18.55 years for boys and 18.48 years for girls, respectively. The mean height and weight were 172.92 cm and 65.81 kg for boys, and 160.13 cm and 52.85 kg for girls. Results indicated that the prevalence of underweight in female was much higher than male, but the prevalence of overweight and obesity in female was lower than male. The mean of serum 25(OH)D level in the 14,726 young participants was 22.39 ± 5.33 ng/ml, and male had higher level of serum 25(OH)D than female, overall, 33.47% of the participants had vitamin D deficiency and even more participants suffered from vitamin D insufficiency (58.17%). The proportion of deficiency in women was much higher than that in men (41.83 vs. 16.35%). The issue of underweight and vitamin D deficiency in young people should be paid attention, especially in females. These results reflected the fact that thinness and paler skin are preferred in modern aesthetics of Chinese culture.

Future plans: WBBC is designed as a prospective cohort study and provides a unique and rich data set analyzing health trajectories from adolescents to young adults.

Key words: Cohort, Late adolescence, Nutrition, Obesity.

Strengths and limitations

The WBBC pilot study was designed to develop a longitudinal database on health-related factors for young people.

Firstly, this was a comprehensive cohort study to pay attention to the health of adolescence and young adults, including over ~14,000 participants.

Secondly, the WBBC pilot cohort is rich in longitudinal phenotypic data and archived biospecimens, including whole blood and serum as well as DNA samples. These resources will facilitate to digitize the information of the genomic, proteomic and metabolomic and microbiome of the participants, and further investigate the association of the information with the development of adolescence and young adults.

Thirdly, this cohort has now accrued sufficient events for reliable analyses, and the major findings on adiposity, BMD, sleep quality, nutrition and growth for adolescence are expected over the next few years.

There are several limitations to this cohort design.

First, although the participants covered all around the country, most of them were mainly from 3 provinces (Jiangxi, Shandong and Zhejiang), and the selection of participants was not random, but in a specific population (college students).

Second, the participants were collected in 3 phases at different time points, and the phenotypic information was updated in the later phases, for example, the items in the questionnaire of different phases were not always the same, therefore, the phenotypic data were not always uniform in 3 phases.

Introduction

The initial goal of setting up this cohort is trying to bring up more attentions to the skeletal health in adolescence and adults, and thus to provide strategy in prevention of osteoporosis in the elderly. Osteoporosis, a disease of increased bone fragility, is a systemic osteopathy characterized by a decrease in bone density/quality and the destruction of bone microstructure caused by genetic and environmental factors[1]. Bone mineral density (BMD), the bone mineral content (BMC) in bone tissue, is recognized as the most important predictor of osteoporosis. Currently, approximately 200 million people worldwide suffer from osteoporosis, and 83.9 million of which are in China [2]. As the trend of global aging is becoming more and more obvious, more attention should be paid to the bone health not only in the elderly, but also in adolescence and young adults. Population-based studies have shown that roughly half of the boys and one-third of the girls would undergo a fracture by age of 18 and 1/5 would have two or more fractures[3 4]. Epidemiologic studies have shown that a 10% increase in peak bone mass (PBM) at the population level reduces the risk of fracture later in life by 50%[5]. In fact, bone mass attained in early life was considered to be the most important modifiable determinant of lifelong skeletal health. A longitudinal data have shown that more than 94% of BMD was acquired at the age of sixteen[6] and approximately 40% to 60% of adult bone mass was accrued during the adolescent years in both women and men[7].

Any condition interfering with optimal peak bone mass accrual can, therefore, increase fracture risk later in life, adolescence is critical period for skeletal mineralization. The bone mass gain during adolescence is influenced by multiple factors, including genetic factors, ethnicity, and environmental factors, such as alcohol intake, cigarette smoking, physical activity, endocrine status (e.g. vitamin D), diet (calcium and protein intake) and other factors[5 8-15]. As environmental and behavioral factors account for 20% to 40% of adult peak bone mass[16 17], the early identification of the factors associated with poor bone health and the provision of reliable counseling might help children and teenagers take action to maximize BMD before their PBM was completed.

Overweight and obesity are a major public health problem[18]. Elevated body mass index (BMI) in adolescence had been associated with several obesity-related morbidities in adult life, such as diabetes, metabolic syndromes and some types of cancer[19]. And obesity in adolescence conferred very high risks for obesity in adults[20]; 70% of overweight adolescents had one or more concomitant conditions such as high blood pressures and fasting insulin, which were also risk factors for cardiovascular disease, and 23% of those accompanied with three or more concomitant conditions[21]. Following rapid economic development since the 1980s, China experiences a rapidly increasing of overweight and obesity among children and adolescents[22 23]. In 2019, a cross-sectional study[24] found that the prevalence of overweight in college students (aged 18-26 years) was 8.0%, and the prevalence of obesity was 3.5%. A recent study from 12 provinces in China showed that the prevalence of overweight and obesity were 14.0% and 10.5% in boys, and 9.7% and 7.1% in girls, respectively[25].

Vitamin D deficiency is becoming a public health problem in both developed and developing countries[26 27]. Besides its effect on musculoskeletal system, Vitamin D showed pleiotropic effect on human health, such as cardiovascular diseases[28], common infectious diseases[29] and autoimmune diseases[30]. Serum 25(OH)D is a good indicator of vitamin D storage and is an optimal method of assessing vitamin D levels[26]. According to the Endocrine Society clinical practice guidelines, vitamin D levels were defined as deficiency [25(OH)D<20 ng/mL], insufficiency [25(OH)D: 20-29 ng/mL)] and sufficiency [25(OH)D≥30 ng/mL] respectively[31]. Many people in central and western Europe had vitamin D concentration of 11–20 ng/mL in winter [32]. Studies from other countries, including Canada[33], Japan[34], Australia[35] and Iran[36], presented similar situations, with high prevalence of vitamin D insufficiency in different ethnicities. A study in north China found that more than 40% of adolescent girls had Vitamin D-deficiency in the winter[37]. Another study in Shanghai showed that more than one-third newborns had plasma 25(OH)D less than 20 ng/mL[38]. Even in Hong Kong (latitude 22° north), 72% of young adults were reported to have vitamin D deficiency[39].

The overall goal of the Westlake Biobank for Chinese (WBBC) pilot cohort is to recruit individuals at their late adolescence/young adulthood. The biological samples such as whole blood, serum, urine and faeces were collected, genomic DNA were extracted and the DNA sequence information were acquired through sequencing technique. A long questionnaire with questions concerning the environmental factors such as nutrition, sleep quality, physical activity, medication etc was provided. These data will help us to understand the association between the genetics, environmental factors, microbiome and health statue of adolescence population. With a broad range of phenotype collection on many aspects of participants' daily life, a wide range of scientific questions could be addressed. The main purpose of this particular paper is to profile the cohort, therefore, only limited findings were described, such as: (1) what are the prevalence of underweight, overweight, obesity and vitamin D deficiency in Chinese late adolescence? What is the reference value of serum vitamin D level in the young people? (2) what is the difference between male and female in term of height, weight, blood pressure, lifestyle and bone health in the young people.

Cohort description

Sampling Design

The Westlake Biobank for Chinese (WBBC) pilot study was collected in three main regions in China (Zhejiang province, Jiangxi province and Shandong province), but the participants covered all around the country (Table 1 and Figure 1). The baseline survey was carried out from 2017 to 2019. The target population was young people aged 14–25 years who were college students and available for follow-up studies. In the first phase of baseline (WBBC pilot 1), the participants were recruited from two colleges at Zhejiang province and Jiangxi province in Southeast China from September 2017 to March 2018 (Figure 2), and 1,258 and 2,769 participants were from Zhejiang and Jiangxi provinces, respectively, and 1,263 participants were from other 25 provinces of China (Table 1). From September 2018 to December 2018, the second phase of WBBC pilot project was initiated (WBBC pilot 2), the participants were recruited at the same college in Jiangxi province and a college in Shandong province in Northeast China (Figure 2). There were 2,920 participants from Shandong province, 2,032 participants from Jiangxi province and 1,306 participants from other 28 provinces of China in WBBC pilot 2 (Table 1). From September 2019, the WBBC pilot project phase 3 (WBBC pilot 3) recruited participants from the same college in Jiangxi province (Figure 2), most of the participants (2,504) of WBBC pilot 3 were from Jiangxi province and 6,74 participants were from other 26 provinces of China (Table 1). All participants provided their Chinese unique national identity (ID) number for unique reference at the health examination center in the campus. The inclusive criteria were: a). All study participants signed the informed consent form before taking part in the survey. b). Participants should complete the physical examination, bone mineral density scan, blood test and questionnaires. And the exclusion criteria were: a). No informed consent. b). Participants missed most of the items in data collection. In WBBC pilot 3, the urine and feces of the participants were collected, therefore, participants taking antibiotics should be excluded.

Data Collection Procedures

The WBBC pilot study is a multidisciplinary study and contains rich range of information regarding of demographics, anthropometric measurements, blood pressure and heart rate, lifestyle and sleep patterns, biological, clinical and health outcomes. Genomic data are available for 7,033 participants. Data and samples were collected via examinations, questionnaire and venipuncture (Table 2).

Measurements of anthropometric parameters

Anthropometric data included height, body weight, bust, waist, hip and thigh circumference, resting blood pressure, heart rate and hand grip. Height was measured to the nearest 0.1 cm with participants' light-weight clothes and shoes off; weight was measured to the nearest 0.01 kg with the weight scale (Ultrasonic surveying instrument, Beryl BYH01, China) calibrated daily before each series of measurements. Bust, waist, hip and thigh circumference were measured to the nearest 0.5 cm by using a measuring tape with the subject standing comfortably. Resting blood pressure and heart rate were measured on the left arm supported at heart-level sitting position using electronic sphygmomanometers (Yuwell YE660A, China). To ensure accurate data, the participants were asked to have rested for at least 5 minutes and have no excessive physical activity, tea or alcohol intake or smoking for at least one night. Using a handgrip dynamometer (CAMRY EH101, China), grip strength with both hands were tested for most of the participants in WBBC pilot 2 and WBBC pilot 3. In order to get more accurate results, the participants should make sure the arm that's being tested was at a 90 degree angle at the elbow[40] until the test was finished. Details of the methods and instruments used for measurements of anthropometric parameters were provided in Table 3.

Biochemistry assessment

Participants came to the examination center in each college in the morning with at least 8 h of overnight fasting, about 25 ml of venous blood samples were collected for routine blood measurements, biochemical indexes, DNA extraction and so on. Venous blood samples were collected using ethylenediamine tetraacetic acid dipotassium (EDTAK2) anticoagulation tube (3×5.0 ml) and vacuum tube without anticoagulation (2×5.0 ml). Serum and plasma samples were separated from whole blood through centrifugation for 10 min at the relative centrifugal force 3000 g (Figure 3). Serum samples were forwarded to test biochemical indexes that included serum 25(OH)D level, serum calcium level, fasting blood glucose, kidney function test, hepatic function test, blood lipids [triglyceride, cholesterol, low-density lipoprotein (LDL), high-density lipoprotein (HDL)], triiodothyronine (T3), thyroid and parathyroid function and bone turnover markers (Table 2). Details of the platforms used for biochemical analysis are provided in Table 3. We also reserved serum samples (2×0.5ml) for each participant at - 80℃ for future use. Figure 3 displays the detail of the flow diagram of blood separation and detection of main blood biochemical indexes.

Questionnaire-based assessments

Baseline data collection for participants included a self-completion questionnaire. In Table 2, a list of core questions within the aforementioned domains, was provided. The questionnaire included social and demographic measures data (e.g. age, sex, ethnicity, family economic status and born place), menstrual history (for female), lifestyle (e.g. physical activity, smoking status, alcohol, tea and coffee intake), additional supplement (e.g. calcium and vitamin D), health status and other information. Sleep duration and sleep quality were assessed by the Pittsburgh Sleep Quality Index (PSQI)[41]. This is composed of 19 questions which reflect seven major components, all seven components are then summed up to create a scale from 0–21 points.

Bone mineral density assessment

Bone mineral density (BMD) is expressed in terms of bone mass per cm^2 (g/cm^2) and were assessed by dual-energy X-ray absorptiometry[42] (DXA, Discovery QDR 4500; Hologic Inc., Waltham, MA, USA). In WBBC pilot, the main sites for BMD evaluation were lumbar spine (L1–L4), femur (femoral neck and total hip) and distal third of the radius.

Whole Genome sequencing and genotyping

In the process of collecting fasting venous blood, 5-ml whole blood was used for isolation of genomic DNA. These genomic DNA were used for whole genome sequencing and genotyping (Figure 3). Whole genome sequencing was completed by NovaSeq 6000 system (Illumina Co., Ltd), and for now, 1,192 participants have been sequenced at mean depth of 14x, with highest depth of 65x. A Chinese specific reference panel will be constructed for imputation for Chinese population. Whole genome genotyping was completed with Infinium Asian Screening Array (ASA) (Illumina Co., Ltd), and 5,841 participants have been genotyped in approximately 700,000 SNPs.

Follow-up and outcome measures

We are seeking funding to follow the cohort to examine development and growth of the participants, and to investigate the effect of environmental factors on later outcomes. An important area of future research will focus on the development of bone mineral density and body weight from late adolescence to adulthood. Figure 2 shows the overall study plan. Follow-up surveys will be conducted according to the design of the subsequent research projects, the participants will be invited for re-survey with repeat interviews, including the questionnaire, anthropometric measurements, grip strength and bone mineral density collection as those used in the baseline stage and the data of nutritional status by food frequency questionnaire.

We have started a pilot follow-up study for WBBC pilot 2 since December 2019, and 1,303 participants had completed all examinations (Figure 2). The collected information included height, weight, grip strength and the updated questionnaire. Besides, we retested bone mineral density at spine (L1-L4), hip and distal third of the radius.

Statistical analysis

To test differences in means and proportions between male and female, we used T-test and Chi-square tests for continuous and categorical variables, respectively. All variables were presented by unadjusted proportions for categorical variables and unadjusted means with standard deviations (SD) for continuous variables. The variables demonstrating a p-value of less than 0.05 were considered statistically significant. All statistical analyses were analyzed using Stata 12.0 software.

Findings to date

This pilot cohort of WBBC is a large longitudinal survey conducted among adolescents and young adults in China. We surveyed 14,726 young people aged 14–25 years who were college students and available for completing follow-up studies. The baseline survey was carried out from 2017 to 2019, including WBBC pilot 1 (5,290 participants), WBBC pilot 2 (6,258 participants) and WBBC pilot 3 (3,178 participants).

We have several ongoing projects under WBBC pilot study. One of the most significant ongoing projects is the study of Chinese population structure. In WBBC pilot study, the data from 1,192 samples with whole genome sequencing and 5,841 samples with whole genome genotyping were available, collaborating with another research group in China, we finally achieved approximate ~10,000 Chinese samples with whole genome data, covering 30 provincial region of China. In addition, WBBC pilot study will provide a haplotype reference panel to improve the imputation accuracy of Chinse GWAS study, since our previous study[43] demonstrated that the existing reference panels, such as the 1000 Genome Phase3 panel[44] and the HRC (Haplotype Reference Consortium) panel[45], were not best fit for imputation for Chinese population, especially for the rare variant imputation.

Given the extensive range of data collection in the WBBC study, it is not feasible to present all the results, only limited findings were described in the present study. In summary, a total of 17,407 college students were invited, of whom, 14,983 (86.07 %) responded. After removing participants with missing data and invalid data, the final study included an effective sample size of 14,726 (84.60%) adolescents and young adults (with age from 14 to 25 years, and mean age at 18.5 years). Table 4 provides an overview of socio-demographic, anthropometry, cardiovascular system, lifestyle, grip strength and BMD characteristics of the WBBC pilot participants at baseline. Briefly, within the 14,726 sample, there were more females than males (67.74 vs. 32.26%), with mean age of 18.48 years for women and 18.55 years for men, respectively. Most of the participants were Chinese Han ethnic (97.78), and more than 60% of them were originally from rural areas (60.24% of males and 69.62% of females). For anthropometry measurements, the mean height and weight were 172.92 cm and 65.81 kg for men, and 160.13 cm and 52.85 kg for women; the mean waist, hip and thigh circumference were 75.51cm, 90.75cm and 51.56 cm in men, and 71.67cm, 89.75cm and 51.68 among women. The mean systolic blood pressure (SBP), diastolic blood pressure (DBP) and heart rate (HR) in participants were 113 mmHg, 71 mmHg and 86 beats/minute, respectively. In the cohort, only 5.49% of the participants were current smokers and 38.82% of them were regular drinkers. Regarding the current smoking status, there was significant difference between men and women (16.29 vs. 1.64%, $p < 0.001$). As for alcohol consumption, the proportion of current drinker in boys and girls were 62.03% and 29.19%, respectively, which is much higher in males ($p < 0.001$). The mean sleeping time estimated in women were higher than men (8.34 vs. 7.97 hours, $p < 0.001$). As for grip strength, the data collection was started from WBBC pilot 2, the mean of grip strength in boys were much higher than girls (grip-left: 36.66 vs. 27.38kg and grip-right: 39.90 vs. 29.76 kg, both $p < 0.001$).

Height and weight were measured using the standardized procedures. Body mass index (BMI) was calculated based on the formula: weight in kilograms divided by height in meters-squared (kg/m^2). According to the Working Group on Obesity in China (WGOC)[46], participants were defined as underweight (< 18.5 kg/m^2), normal weight (18.5-23.9 kg/m^2), overweight (24-27.9 kg/m^2) and obese (≥ 28 kg/m^2). Therefore, the WBBC pilot study provided an overall prevalence of underweight, overweight and obesity among young participants of 24.26%, 11.50% and 5.03%, respectively (Table 5). The prevalence of underweight in female was much higher than male (26.41% vs. 19.70%, $p < 0.0001$), but the prevalence of overweight in female was much lower than male (9.03% vs. 16.71%, $p < 0.0001$) (Table 5), similarly, the prevalence of obesity in female (3.21%) was lower than in male (8.88%) ($p < 0.0001$) (Table 5). Waist circumference (WC) is good indicator of abdominal visceral fat distribution and is a strong predictor of diabetes mellitus and cardiovascular disease [47]. It is meaningful to investigate the WC along with BMI among adolescents and young people. In WBBC pilot study, central obesity was defined as WC≥85 cm for men and as WC≥80 cm for women based on the recommendations of the WGOC [46]. In the cohort of 12,396 participants, the prevalence of central obesity was 14.62%, which was higher in male than in female (19.10% vs. 12.55%, $p < 0.0001$) (Table 5).

In WBBC pilot study, the prevalence of underweight were high in both male (19.70 %) and female (26.41%), though the prevalence of moderate and severe underweight decreased from 9.2% in 1975 to 8.4% in 2016 in girls and from 14.8% in 1975 to 12.4% in 2016 in boys in the world[48]. This phenomenon may due to the modern aesthetics of human stature that thinness is preferred, especially in China[49 50]. Recently, a study involving 2,023 young female participants (70.5% subjects aged 20-25 years) from eight Chinese universities[50] showed that 30.55% of the participants were underweight, and 57.39% of them would like to be much thinner, which would lead to more underweight individuals. Therefore, future studies should not only pay attention to the problem of obesity/overweight, but also to the underweight issue in young people.

Using simple anthropometric indices of body composition, such as BMI and WC, has been considered as a practical and valuable approach to the assessment of obesity for a long time. Waist-to-hip ratios (WHR), waist-to-height ratios (WHeR), a body shape index (ABSI)[51] and body roundness index (BRI)[52] were also as parameters of body fat and visceral adipose tissue volume. In WBBC pilot cohort, we had collected several anthropometric measures including height, weight, bust, waist, hip and thigh circumference and these data could help us examine the usefulness of these anthropometric parameters and identify the optimal cut-off of the parameters to evaluate overweight and obesity among adolescence and young people in future study.

In WBBC pilot study, the mean serum 25(OH)D level was 22.39 ± 5.33 ng/ml for all the participants (male: 25.15 ng/mL, and female: 21.05 ng/mL, $p < 0.0001$) (Table 6). Overall, 33.47% of the participants had vitamin D deficiency and even more participants suffered from vitamin D insufficiency (58.17%) (Table 6). In addition, the proportion of women with sufficient vitamin D was much lower than that of men (3.69 vs. 17.91%, $p < 0.0001$), while the proportion of deficiency in women was much higher than that in men (41.83 vs. 16.35%, $p < 0.0001$) (Table 6). Most of the participants (86.87%) preferred to stay indoors in spare time, the females were less willing to do exercise than males (53.68% vs 70.59%) (Table 7), and 44.58% of females hardly had outdoor activities, only 5.89% of females often had outdoor activities every week (Table 7). These results jointly suggested that the females had not enough sun exposure. Although food sources of vitamin D were not commonly recognized, only 10–20% of vitamin D in human bodies was obtained through food sources[53]. In WBBC pilot study, there was only 2.24% of the participants used vitamin D supplements (3.12% in male and 1.87% in female, $p = 0.00098$) (Table 6) and this was consistent with Zhou et al[54], which found that only 5.6 % of the students used vitamin D supplements in a university of Nanjing, China. It is noteworthy that vitamin D deficiency in females was significantly worse than in males. This may due to the modern aesthetics of Chinese culture that paler skin is preferred, especially in females. A questionnaire related to vitamin D and sun exposure was conducted at a university in Nanjing, China and found that 75.0 % of the students lacked sun exposure because they would like to avoid dark skin, and most of students (82.7 %) used sun protection, and sunscreen use were more popular in females[54], but it was reported that using the amount of sun cream recommended by World Health Organization exponentially suppressed vitamin D synthesis in the skin[55].

Collaboration

Participants have agreed to provide their pseudonymized data being made available to other approved researchers. The WBBC pilot study welcomes and offers global collaboration. The data is not freely available in the public domain, but specific proposals and ideas for future collaboration would be very welcome. Applicants for collaboration and more information are encouraged to contact Dr. Hou-Feng Zheng (Email address: zhenghoufeng@westlake.edu.cn), the person in charge of this project.

Funding

This study was supported by the National Natural Science Foundation of China (81871831), and by the Zhejiang Provincial Natural Science Foundation for Distinguished Young Scholars of China (LR17H070001), and by the Westlake Biobank for Chinese (WBBC) funds from Westlake University.

Acknowledgements

We gratefully acknowledge all the participants of this project, and thank all the people who helped us in the sample recruitment, including volunteers, laboratory technicians, nurses and clerical workers. We also thank the Westlake University Supercomputer Center for the facility support and technical assistance. In addition, we would like to thank Shanghai AvanTech BioSciences Co., Ltd for their support and assistance in sample storage and management.

Contributors

Hou-Feng Zheng gained funding and conceived of the study, and all authors were involved in the design and collection data of the study. Xiao-Wei Zhu and Hou-Feng Zheng analyzed the data and wrote the paper and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Conflict of interest

Jun-Quan Liu and Yi Sun are employees of Hangzhou Kingmed Diagnostics Co., Ltd. The other authors have no conflict of interest to declare.

Patient and public involvement

It was not appropriate or possible to involve patients or the public in the design, or conduct, or reporting, or dissemination plans of our research.

Ethics Approval and Consent to Participate

The study protocol and informed consent procedure were approved by the Ethics Committees at Westlake University. All study participants signed the informed consent form before taking part in the study.

References

1. Armas LA, Recker RR. Pathophysiology of osteoporosis: new mechanistic insights. *Endocrinology and metabolism clinics of North America* 2012;41(3):475-86. doi: 10.1016/j.ecl.2012.04.006

2. Pisani P, Renna MD, Conversano F, et al. Major osteoporotic fragility fractures: Risk factor updates and societal impact. *World journal of orthopedics* 2016;7(3):171-81. doi: 10.5312/wjo.v7.i3.171

3. Mayranpaa MK, Makitie O, Kallio PE. Decreasing incidence and changing pattern of childhood fractures: A population-based study. *Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research* 2010;25(12):2752-9. doi: 10.1002/jbmr.155

4. Bishop N, Arundel P, Clark E, et al. Fracture prediction and the definition of osteoporosis in children and adolescents: the ISCD 2013 Pediatric Official Positions. *Journal of clinical densitometry : the official journal of the International Society for Clinical Densitometry* 2014;17(2):275-80. doi: 10.1016/j.jocd.2014.01.004

5. Gordon CM, Zemel BS, Wren TA, et al. The Determinants of Peak Bone Mass. *The Journal of pediatrics* 2017;180:261-69. doi: 10.1016/j.jpeds.2016.09.056

6. Berger C, Goltzman D, Langsetmo L, et al. Peak bone mass from longitudinal data: implications for the prevalence, pathophysiology, and diagnosis of osteoporosis. *Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research* 2010;25(9):1948-57. doi: 10.1002/jbmr.95

7. Golden NH, Abrams SA, Committee on N. Optimizing bone health in children and adolescents. *Pediatrics* 2014;134(4):e1229-43. doi: 10.1542/peds.2014-2173

8. Gaddini GW, Turner RT, Grant KA, et al. Alcohol: A Simple Nutrient with Complex Actions on Bone in the Adult Skeleton. *Alcoholism, clinical and experimental research* 2016;40(4):657-71. doi: 10.1111/acer.13000

9. Rudang R, Darelid A, Nilsson M, et al. Smoking is associated with impaired bone mass development in young adult men: a 5-year longitudinal study. *Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research* 2012;27(10):2189-97. doi: 10.1002/jbmr.1674

10. Troy KL, Mancuso ME, Butler TA, et al. Exercise Early and Often: Effects of Physical Activity and Exercise on Women's Bone Health. *International journal of environmental research and public health* 2018;15(5) doi: 10.3390/ijerph15050878

11. Reid IR. Vitamin D Effect on Bone Mineral Density and Fractures. *Endocrinology and metabolism clinics of North America* 2017;46(4):935-45. doi: 10.1016/j.ecl.2017.07.005

12. Monge MC. Optimizing bone health in adolescents. *Current opinion in obstetrics & gynecology* 2018;30(5):310-15. doi: 10.1097/GCO.0000000000000482

13. Zhu X, Zheng H. Factors influencing peak bone mass gain. *Front Med* 2020 doi: 10.1007/s11684-020-0748-y

14. Xia J, Xie SY, Liu KQ, et al. Systemic evaluation of the relationship between psoriasis, psoriatic arthritis and osteoporosis: observational and Mendelian randomisation study. *Ann Rheum Dis* 2020 doi: 10.1136/annrheumdis-2020-217892

15. Bai WY, Wang L, Ying ZM, et al. Identification of PIEZO1 polymorphisms for human bone mineral density. *Bone* 2020;133:115247. doi: 10.1016/j.bone.2020.115247

16. Heaney RP, Abrams S, Dawson-Hughes B, et al. Peak bone mass. *Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA* 2000;11(12):985-1009. doi: 10.1007/s001980070020

17. Bachrach LK. Acquisition of optimal bone mass in childhood and adolescence. *Trends in endocrinology and metabolism: TEM* 2001;12(1):22-8. doi: 10.1016/s1043-2760(00)00336-2

18. Collaborators GBDO, Afshin A, Forouzanfar MH, et al. Health Effects of Overweight and Obesity in 195 Countries over 25 Years. *The New England journal of medicine* 2017;377(1):13-27. doi: 10.1056/NEJMoa1614362

19. Llewellyn A, Simmonds M, Owen CG, et al. Childhood obesity as a predictor of morbidity in adulthood: a systematic review and meta-analysis. *Obesity reviews : an official journal of the International Association for the Study of Obesity* 2016;17(1):56-67. doi: 10.1111/obr.12316

20. Patton GC, Coffey C, Carlin JB, et al. Overweight and obesity between adolescence and young adulthood: a 10-year prospective cohort study. *The Journal of adolescent health : official publication of the Society for Adolescent Medicine* 2011;48(3):275-80. doi: 10.1016/j.jadohealth.2010.06.019

21. Freedman DS, Kahn HS, Mei Z, et al. Relation of body mass index and waist-to-height ratio to cardiovascular disease risk factors in children and adolescents: the Bogalusa Heart Study. *The American journal of clinical nutrition* 2007;86(1):33-40. doi: 10.1093/ajcn/86.1.33

22. Yu Z, Han S, Chu J, et al. Trends in overweight and obesity among children and adolescents in China from 1981 to 2010: a meta-analysis. *PloS one* 2012;7(12):e51949. doi: 10.1371/journal.pone.0051949

23. Ng M, Fleming T, Robinson M, et al. 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. *Lancet* 2014;384(9945):766-81. doi: 10.1016/S0140-6736(14)60460-8

24. Jiang Y, Wang J, Wu S, et al. Association between Take-Out Food Consumption and Obesity among Chinese University Students: A Cross-Sectional Study. *International journal of environmental research and public health* 2019;16(6) doi: 10.3390/ijerph16061071

25. Zhang J, Wang H, Wang Z, et al. Prevalence and stabilizing trends in overweight and obesity among children and adolescents in China, 2011-2015. *BMC public health* 2018;18(1):571. doi: 10.1186/s12889-018-5483-9

26. Holick MF. Vitamin D deficiency. *The New England journal of medicine* 2007;357(3):266-81. doi: 10.1056/NEJMra070553

27. van Schoor N, Lips P. Global Overview of Vitamin D Status. *Endocrinology and metabolism clinics of North America* 2017;46(4):845-70. doi: 10.1016/j.ecl.2017.07.002

28. Trehan N, Afonso L, Levine DL, et al. Vitamin D Deficiency, Supplementation, and Cardiovascular Health. *Critical pathways in cardiology* 2017;16(3):109-18. doi: 10.1097/HPC.0000000000000122

29. Watkins RR, Lemonovich TL, Salata RA. An update on the association of vitamin D deficiency with common infectious diseases. *Canadian journal of physiology and pharmacology* 2015;93(5):363-8. doi: 10.1139/cjpp-2014-0352

30. Illescas-Montes R, Melguizo-Rodriguez L, Ruiz C, et al. Vitamin D and autoimmune diseases. *Life sciences* 2019;233:116744. doi: 10.1016/j.lfs.2019.116744

31. Holick MF, Binkley NC, Bischoff-Ferrari HA, et al. Evaluation, treatment, and prevention of vitamin

D deficiency: an Endocrine Society clinical practice guideline. *The Journal of clinical endocrinology and metabolism* 2011;96(7):1911-30. doi: 10.1210/jc.2011-0385

32. Pludowski P, Grant WB, Bhattoa HP, et al. Vitamin d status in central europe. *International journal of endocrinology* 2014;2014:589587. doi: 10.1155/2014/589587
33. Newhook LA, Sloka S, Grant M, et al. Vitamin D insufficiency common in newborns, children and pregnant women living in Newfoundland and Labrador, Canada. *Maternal & child nutrition* 2009;5(2):186-91. doi: 10.1111/j.1740-8709.2008.00157.x
34. Yoshimura N, Muraki S, Oka H, et al. Profiles of vitamin D insufficiency and deficiency in Japanese men and women: association with biological, environmental, and nutritional factors and coexisting disorders: the ROAD study. *Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA* 2013;24(11):2775-87. doi: 10.1007/s00198-013-2372-z
35. Hirani V, Cumming RG, Blyth FM, et al. Vitamin D status among older community dwelling men living in a sunny country and associations with lifestyle factors: the Concord Health and Ageing in Men Project, Sydney, Australia. *The journal of nutrition, health & aging* 2013;17(7):587-93. doi: 10.1007/s12603-013-0013-z
36. Faghih S, Abdolahzadeh M, Mohammadi M, et al. Prevalence of vitamin d deficiency and its related factors among university students in shiraz, iran. *International journal of preventive medicine* 2014;5(6):796-9.
37. Fraser DR. Vitamin D-deficiency in Asia. *The Journal of steroid biochemistry and molecular biology* 2004;89-90(1-5):491-5. doi: 10.1016/j.jsbmb.2004.03.057
38. Yu X, Wang W, Wei Z, et al. Vitamin D status and related factors in newborns in Shanghai, China. *Nutrients* 2014;6(12):5600-10. doi: 10.3390/nu6125600
39. Wang EW, Pang MY, Siu PM, et al. Vitamin D status and cardiometabolic risk factors in young adults in Hong Kong: associations and implications. *Asia Pacific journal of clinical nutrition* 2018;27(1):231-37. doi: 10.6133/apjcn.022017.08
40. Roberts HC, Denison HJ, Martin HJ, et al. A review of the measurement of grip strength in clinical and epidemiological studies: towards a standardised approach. *Age and ageing* 2011;40(4):423-9. doi: 10.1093/ageing/afr051
41. Buysse DJ, Reynolds CF, 3rd, Monk TH, et al. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry research* 1989;28(2):193-213. doi: 10.1016/0165-1781(89)90047-4
42. Bazzocchi A, Ponti F, Albisinni U, et al. DXA: Technical aspects and application. *European journal of radiology* 2016;85(8):1481-92. doi: 10.1016/j.ejrad.2016.04.004
43. Bai WY, Zhu XW, Cong PK, et al. Genotype imputation and reference panel: a systematic evaluation on haplotype size and diversity. *Brief Bioinform* 2019 doi: 10.1093/bib/bbz108
44. Genomes Project C, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature* 2015;526(7571):68-74. doi: 10.1038/nature15393
45. McCarthy S, Das S, Kretzschmar W, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nature genetics* 2016;48(10):1279-83. doi: 10.1038/ng.3643
46. Zhou BF, Cooperative Meta-Analysis Group of the Working Group on Obesity in C. Predictive values of body mass index and waist circumference for risk factors of certain related diseases in Chinese adults--study on optimal cut-off points of body mass index and waist circumference in Chinese adults. *Biomedical and environmental sciences : BES* 2002;15(1):83-96.

47. Asia Pacific Cohort Studies C. Central obesity and risk of cardiovascular disease in the Asia Pacific Region. *Asia Pacific journal of clinical nutrition* 2006;15(3):287-92.

48. Collaboration NCDRF. 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. *Lancet* 2017;390(10113):2627-42. doi: 10.1016/S0140-6736(17)32129-3

49. Chen H, Jackson T. Predictors of changes in weight esteem among mainland Chinese adolescents: a longitudinal analysis. *Developmental psychology* 2009;45(6):1618-29. doi: 10.1037/a0016820

50. Zhang L, Qian H, Fu H. To be thin but not healthy - The body-image dilemma may affect health among female university students in China. *PloS one* 2018;13(10):e0205282. doi: 10.1371/journal.pone.0205282

51. Krakauer NY, Krakauer JC. A new body shape index predicts mortality hazard independently of body mass index. *PloS one* 2012;7(7):e39504. doi: 10.1371/journal.pone.0039504

52. Thomas DM, Bredlau C, Bosy-Westphal A, et al. Relationships between body roundness with body fat and visceral adipose tissue emerging from a new geometrical model. *Obesity* 2013;21(11):2264-71. doi: 10.1002/oby.20408

53. Mithal A, Wahl DA, Bonjour JP, et al. Global vitamin D status and determinants of hypovitaminosis D. *Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA* 2009;20(11):1807-20. doi: 10.1007/s00198-009-0954-6

54. Zhou M, Zhuang W, Yuan Y, et al. Investigation on vitamin D knowledge, attitude and practice of university students in Nanjing, China. *Public health nutrition* 2016;19(1):78-82. doi: 10.1017/S1368980015000373

55. Faurschou A, Beyer DM, Schmedes A, et al. The relation between sunscreen layer thickness and vitamin D production after ultraviolet B exposure: a randomized clinical trial. *The British journal of dermatology* 2012;167(2):391-5. doi: 10.1111/j.1365-2133.2012.11004.x

Tables

Table 1. The participants of the cohort at baseline.

Year	Phase	Area	Total	Sex	
				Male	Female
2017	WBBC pilot 1	Zhejiang province	1,258	460	798
		Jiangxi province	2,769	862	1,907
		Other 25 provinces	1,263	336	927
2018	WBBC pilot 2	Shandong province	2,920	1,115	1,805
		Jiangxi province	2,032	578	1,454
		Other 28 provinces	1,306	438	868
2019	WBBC pilot 3	Jiangxi province	2,504	761	1,743
		Other 26 provinces	674	201	473
Total			14,726	4,751	9,975

Table 2. Summary of data collected in the Cohort.

Measures	**Instruments**
Anthropometry*	
	Height and weight, bust, waist, hip and thigh circumference
Strength	
	Grip strength (right and left hand)
Cardiovascular system	
	Blood pressure (SBP & DBP), heart rate
Bone mineral density	
	Lumar spine, hip and forearm
Questionnaires	
Demographic data	Age, sex, ethnicity, family economic status, birthplace
Menstrual history (for female)	Age of menarche, Gynaecological disease history
Sleeping situation	Pittsburgh sleep quality index, PSQI
Lifestyle	Smoking status, alcohol, tea and coffee consumption, exercise
Supplementations	Calcium, vitamin D
Health status	Medical history and medications
Physical activity[#]	International Physical Activity Questionnaire, IPAQ
Dietary pattern[#]	Semi-Quantitative Food Frequency Questionnaire (SQFFQ)
Routine biochemistry and Haematology	
Lipid metabolism	Total cholesterol, triglycerides, HDL-cholesterol, LDL-cholesterol
Kidney-related measures	uric acid, creatinine and urea
Hepatic-related measures	Alanine aminotransferase, glutamic oxaloacetylase, total bilirubin, direct bilirubin, albumin, globulin
Thyroid and parathyroid function	TSH, PTH, T3, T4, FT3, FT4,
Routine diabetes test	Fasting plasma glucose, glycosylated hemoglobin (HbA1c)
Bone-related measures	Bone turnover markers (Osteocalcin, PINP, NBAP and β-CTX), serum $1,25(OH)_2D_3$, seruum IGF-1
Mineral elements	Calcium (Ca), phosphorus (P) and magnesium (Mg)
Samples	
Blood	25 ml
Serum	2 ml
Urine[#]	10 ml
Faeces[#]	15 g

* Direct measures carried out by trained interviewers according to standardized protocols; [#] measured only on WBBC pilot 3; BAP: Bone alkaline phosphatase; DBP: Diastolic blood pressure; FT3: free Triiodothyronine; FT4: free Tetraiodothyronine; HDL: High-density lipoprotein; IGF-1: Insulin-like growth factor 1; LDL: Low-density lipoprotein; PINP: N-terminal propeptide of procollagen type I; PTH: Parathyroid hormone; SBP: Systolic blood pressure; T3: Triiodothyronine; T4: Tetraiodothyronine; TSH: Thyroid stimulating hormone; β-CTx: β-isomerized C-terminal telopeptides

Table 3. List of anthropometric collected and platforms for biochemical tests at baseline in the cohort.

Variables	Analysis method/platform used
Height and weight	Ultrasonic surveying instrument, Beryl BYH01, China
Waist, hip and thigh circumference	Manufactured instrument (tape)
Resting blood pressure	Electronic sphygmomanometer, Yuwell YE660A, China
Heart rate	Electronic sphygmomanometer, Yuwell YE660A, China
Grip strength	Digital hand dynamometer, CAMRY EH101, China
Bone mineral density (BMD)	DXA, Discovery QDR 4500, Hologic Inc., Waltham, MA, USA
Blood routine (5 items)	SYSMEX 2100, Japan
Serum IGF-1	CLIA, DPC immulite 2000, Siemens, Germany
Serum calcium, phosphorus	ARCHITECT C16000, Abbott, USA
Fasting blood glucose	Cobas c501, Roche, Switzerland
Bone turnover markers	ECLIA, Cobas e602, Roche, Switzerland
Serum 1,25$(OH)_2D_3$	LC-MS/MS, AB Sciex API 4000™, USA
Thyroid and parathyroid function	Chemical luminescence, ARCHITECT System i2000, USA
Lipid metabolism	ARCHITECT C16000, Abbott, USA
Kidney-related measures	ARCHITECT C16000, Abbott, USA
Hepatic-related measures	ARCHITECT C16000, Abbott, USA

CLIA: chemiluminescent immunoassay; ECLIA: electrochemoluminescence immunoassay;
LC-MS/MS: Liquid chromatography-tandem mass spectrometry.

Table 4. Basic characteristics of participants in baseline of WBBC pilot.

Variables(unit)		Total	Sex		*P value
			Male	Female	
Socio-demographic					
Age (years)	N[a]	14,726	4,751	9,975	0.001
	M (SD)[b]	18.50 (1.25)	18.55 (1.21)	18.48 (1.26)	
Gender	N (%)[c]	14,726 (100)	4,751 (32.26)	9,975 (67.74)	<0.001
Ethnicity					0.305
Han	N (%)	11,136 (97.78)	3,642 (97.98)	7,494 (97.68)	
Others	N (%)	253 (2.22)	75 (2.12)	178 (2.32)	
Hukou status					<0.001
Rural	N (%)	5,951 (66.59)	1,739 (60.24)	4,212 (69.62)	
Urban	N (%)	2,986(33.41)	1,148 (39.76)	1,838 (30.38)	
Anthropometry					
Height (cm)	N	14,277	4,588	9,689	<0.001
	M (SD)	164.24 (8.49)	172.92 (6.57)	160.13 (5.76)	
Weight (kg)	N	14,279	4,587	9,692	<0.001
	M (SD)	57.01 (12.32)	65.81 (13.60)	52.85 (9.07)	
Waist (cm)	N	12,396	3,905	8,491	<0.001
	M (SD)	71.67 (9.84)	75.51 (11.07)	69.90 (8.66)	
Hip (cm)	N	12,388	3,902	8,486	<0.001
	M (SD)	89.75 (7.28)	90.75 (8.28)	89.29 (6.73)	
Thigh (cm)	N	12,351	3,880	8,471	0.102
	M (SD)	51.68 (5.69)	51.56 (6.50)	51.74 (5.27)	
Cardiovascular system					
SBP (mmHg)	N	14,277	4,595	9,682	<0.001
	M (SD)	113 (12)	121 (12)	110 (11)	
DBP (mmHg)	N	14,276	4,595	9,681	<0.001
	M (SD)	71 (9)	73 (9)	70 (8)	
Heart rate (beats/minute)	N	14,295	4,599	9,696	<0.001
	M (SD)	86 (13)	83 (13)	87 (13)	
Lifestyle					
Smoking	N (%)	435 (5.94)	350 (16.29)	85 (1.64)	<0.001
Alcohol consumption	N (%)	2,844 (38.82)	1,333 (62.03)	1,511 (29.19)	<0.001
Sleeping time (hours)	N	7,247	2,115	5,132	
	M (SD)	8.23 (1.37)	7.97 (1.26)	8.34 (1.40)	<0.001
Grip strength (kg)					

Left hand	N	8,932	2,958	5,974	<0.001
	M (SD)	27.38 (8.77)	36.66 (7.53)	22.79 (4.82)	
Right hand	N	8,941	2,967	5,974	<0.001
	M (SD)	29.76 (9.66)	39.90 (8.32)	24.72 (5.35)	
BMD (g/cm^2)					
Lumar spine	N	10,154	3,293	6,861	<0.001
	M (SD)	0.910 (0.105)	0.926 (0.112)	0.903(0.100)	
Total hip	N	10,160	3,296	6,864	<0.001
	M (SD)	0.868 (0.127)	0.932 (0.139)	0.837 (0.108)	
Femoral neck	N	10,160	3,296	6,864	<0.001
	M (SD)	0.778 (0.125)	0.846 (0.138)	0.746 (0.104)	
Forearm	N	9,917	3,238	6,679	<0.001
	M (SD)	0.657 (0.059)	0.705 (0.057)	0.634 (0.453)	

[a] N: sample size; [b]M (SD): mean (standard deviation); [c]N (%): sample size (percentage); * T-test and Chi-square tests for continuous and categorical variables respectively to refer the significant differences between males and females; DBP: Diastolic blood pressure; SBP: Systolic blood pressure.

Table 5. Distribution of BMI or waist by sex in participants in WBBC pilot, 2017-2019. N=14,264.

Variables	**Total**	**Male**	**Female**	**#P**
Underweight, N (%)	3,460 (24.26%)	903 (19.70%)	2,557 (26.41%)	<0.0001
Normal, N (%)	8,446 (59.21%)	2,507 (54.70%)	5,939 (61.35%)	<0.0001
Overweight, N (%)	1,640 (11.50%)	766 (16.71%)	874 (9.03%)	<0.0001
Obesity, N (%)	718 (5.03%)	407 (8.88%)	311 (3.21%)	<0.0001
Central Obesity[δ], N (%)	1,812 (14.62%)	746 (19.10%)	1,066 (12.55%)	<0.0001

[δ] Total sample size for Central Obesity was 12,396. [#] Chi-square tests was calculated between male and female.

Table 6. Distribution of Vitamin D level by sex in subjects in WBBC pilot, 2017-2018. N=11,370.

Variables	**Total**	**Male**	**Female**	**P**
Serum 1,25(OH)D, mean (SD)	22.39 (5.33)	25.15 (5.43)	21.05 (4.73)	<0.0001$^{\eta}$
Vitamin D deficient, N (%)	3,806 (33.47%)	610 (16.35%)	3,196 (41.83%)	<0.0001$^{\#}$
Vitamin D insufficient, N (%)	6,614 (58.17%)	2,452 (65.74%)	4,162 (54.48%)	<0.0001$^{\#}$
Vitamin D sufficient, N (%)	950 (8.36%)	668 (17.91%)	282 (3.69%)	<0.0001$^{\#}$
Vitamin D supplementation$^{\delta}$, N (%)	164 (2.24%)	67 (3.12%)	97 (1.87%)	0.00098$^{\#}$

$^{\delta}$ The sample size for Vitamin D supplementation was 7,326. $^{\#}$ Chi-square tests was calculated between male and female. $^{\eta}$ T-test was calculated between male and female.

Table 7. Participants' general react about activity status in WBBC pilot, 2017-2018. N=7,326.[a]

	Total	**Male**	**Female**
What would you like to do in spare time?			
Stay indoors	6,364 (86.87%)	1,754 (81.62%)	4,610 (89.05%)
Take part in some activities	962 (13.13%)	395 (18.38%)	567 (10.95%)
Do you do exercise initiatively?			
Yes	4,296 (58.64%)	1,517 (70.59%)	2,779 (53.68%)
No	3,030 (41.36%)	632 (29.41%)	2,398 (46.32%)
How often do you have outdoor activities every week?			
Hardly	2,785 (38.02%)	477 (22.20%)	2,308 (44.58%)
Occasionally	3,865 (52.76%)	1,301 (60.54%)	2,564 (49.53%)
Often	676 (9.23%)	371 (17.26%)	305 (5.89%)

[a] Data are showed as n (%) of participants.

Figure legend

Figure 1. Maps showing the sources of the samples in the Westlake BioBank for Chinese (WBBC) pilot cohort. AH: Anhui province; BJ: Beijing; CQ: Chongqing; FJ: Fujian province; GD: Guangdong province; GS: Gansu province; GX: Guangxi Autonomous Region; GZ: Guizhou province; HA: Henan province; HB: Hubei province; HE: Hebei province; HI: Hainan province; HL: Heilongjiang province; HN: Hunan province; JL: Jilin province; JS: Jiangsu province; JX: Jiangxi province; LN: Liaoning province; NM: Neimeng Autonomous Region; NX: Ningxia Autonomous Region; QH: Qinghai province; SC: Sichuang province; SD: Shandong province; SH: Shanghai; SN: Shanxi province; SX: Shanxi province; TJ: Tianjin; XJ: Xinjiang Autonomous Region; XZ: Xizang Autonomous Region; YN: Yunnan province; ZJ: Zhejiang province.

Figure 2. Data collection timeline. IPAQ: International Physical Activity Questionnaire; PA: Physical Activity; PSQI: Pittsburgh sleep quality index; SQFFQ: Semi-Quantitative Food Frequency Questionnaire.

Figure 3. Flow diagram of main blood biochemical detection and blood conservation.
RCF: relative centrifugal force; WGS: whole genome sequence.

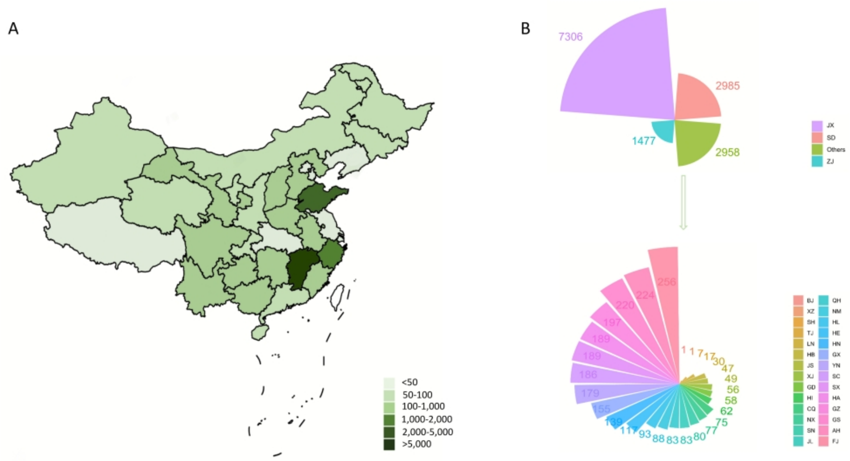

Figure 1. Maps showing the sources of the samples in the Westlake BioBank for Chinese (WBBC) pilot cohort. AH: Anhui province; BJ: Beijing; CQ: Chongqing; FJ: Fujian province; GD: Guangdong province; GS: Gansu province; GX: Guangxi Autonomous Region; GZ: Guizhou province; HA: Henan province; HB: Hubei province; HE: Hebei province; HI: Hainan province; HL: Heilongjiang province; HN: Hunan province; JL: Jilin province; JS: Jiangsu province; JX: Jiangxi province; LN: Liaoning province; NM: Neimeng Autonomous Region; NX: Ningxia Autonomous Region; QH: Qinghai province; SC: Sichuang province; SD: Shandong province; SH: Shanghai; SN: Shanxi province; SX: Shanxi province; TJ: Tianjin; XJ: Xinjiang Autonomous Region; XZ: Xizang Autonomous Region; YN: Yunnan province; ZJ: Zhejiang province.

338x190mm (300 x 300 DPI)

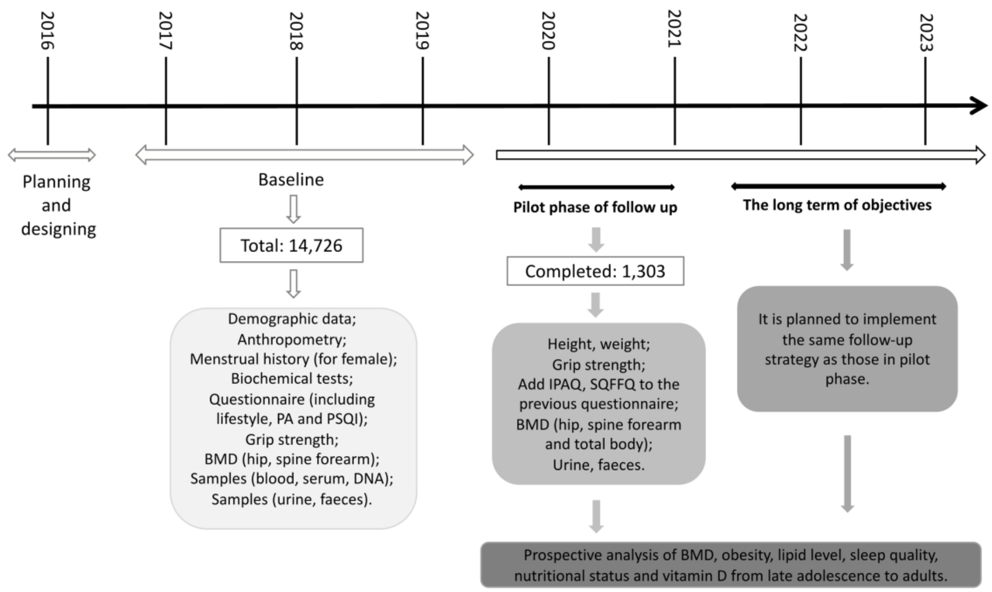

Figure 2. Data collection timeline. IPAQ: International Physical Activity Questionnaire; PA: Physical Activity; PSQI: Pittsburgh sleep quality index; SQFFQ: Semi-Quantitative Food Frequency Questionnaire.

254x190mm (300 x 300 DPI)

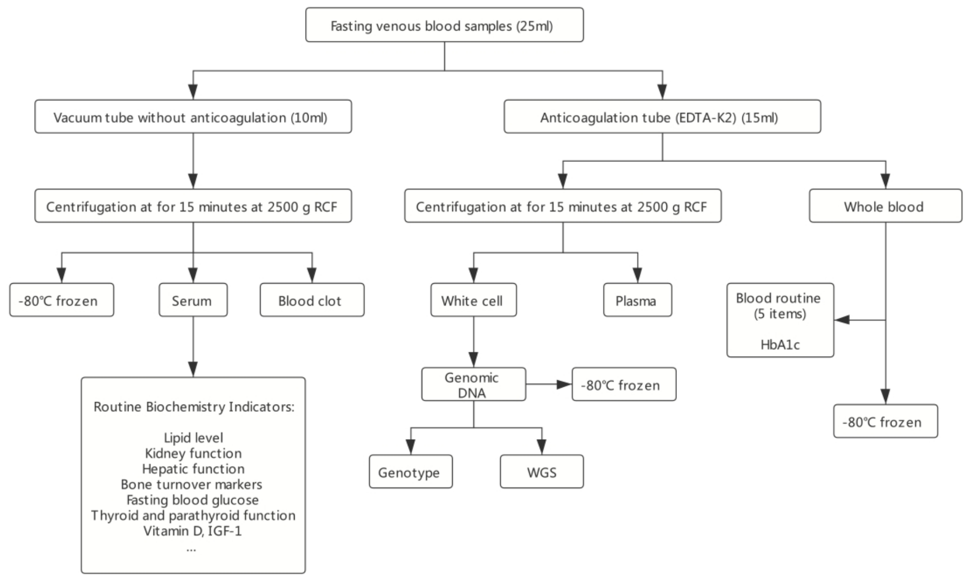

Figure 3. Flow diagram of main blood biochemical detection and blood conservation. RCF: relative centrifugal force; WGS: whole genome sequence.

363x223mm (300 x 300 DPI)

BMJ Open

Cohort profile: the Westlake BioBank for Chinese (WBBC) pilot project

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-045564.R1
Article Type:	Cohort profile
Date Submitted by the Author:	23-Mar-2021
Complete List of Authors:	Zhu, Xiao-Wei; Westlake University, Diseases & Population (DaP) Geninfo Lab; Fudan University, School of Life Sciences Liu, Ke-Qi ; Jiangxi Medical College Wang, Ping-Yu; Department of Epidemiology & Department of Biochemistry and Molecular Biology, Binzhou Medical University, Yantai, Shandong 264003, China Liu, Jun-Quan; Hangzhou Kingmed Diagnostics Co. Chen, Jin-Yang; Zhejiang Healthfuture Institute for Cell-Based Applied Technology Xu, Xue-Jin; Jiading District Center for Disease Control and Prevention Xu, Jin-Jian; Jiangxi Medical College Qiu, Mo-Chang; Jiangxi Medical College Sun, Yi; Hangzhou Kingmed Diagnostics Co. Liu, Chu; Hangzhou Normal University Bai, Wei-Yang; Westlake University, Diseases & Population (DaP) Geninfo Lab; Fudan University, School of Life Sciences Zhao, Pian-Pian; Westlake University School of Life Sciences, Diseases & Population (DaP) Geninfo Lab Xia, Jiangwei; Westlake University, Diseases & Population (DaP) Geninfo Lab Gai, Si-Rui ; Westlake University, Diseases & Population (DaP) Geninfo Lab Guan, Peng-Lin; Westlake University, Diseases & Population (DaP) Geninfo Lab Qian, Yu; Westlake University, Diseases & Population (DaP) Geninfo Lab Cong, Pei-Kuan Cong; Westlake University, Diseases & Population (DaP) Geninfo Lab Xie, Shu-Yang ; Binzhou Medical University Zheng, Hou-Feng; Westlake University,
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Epidemiology, Paediatrics, Public health
Keywords:	EPIDEMIOLOGY, NUTRITION & DIETETICS, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS

Cohort profile: the Westlake BioBank for Chinese (WBBC) pilot project

Xiao-Wei Zhu[1,2,3], Ke-Qi Liu[4], Ping-Yu Wang[5], Jun-Quan Liu[6,7], Jin-Yang Chen[7], Xue-Jin Xu[8], Jin-Jian Xu[4], Mo-Chang Qiu[4], Yi Sun[6], Chun Liu[9], Wei-Yang Bai[1,2,3], Pian-Pian Zhao[1,2], Jiangwei Xia[1,2], Si-Rui Gai[1,2], Peng-Lin Guan[1,2], Yu Qian[1,2], Pei-Kuan Cong[1,2], Shu-Yang Xie[5], Hou-Feng Zheng[1,2,3*]

[1] Disease & Population (DaP) Geninfo Lab, School of Life Sciences, Westlake University, Hangzhou, Zhejiang, 310024, China

[2] Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Westlake University, Hangzhou, Zhejiang, 310024, China

[3] Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, Zhejiang, 310024, China

[4] WBBC Jiangxi Center, Jiangxi Medical College, Shangrao, Jiangxi, China

[5] Department of Epidemiology & Department of Biochemistry and Molecular Biology, Binzhou Medical University, Yantai, Shandong 264003, China

[6] Hangzhou Kingmed Diagnostics Co., Ltd., Hangzhou, Zhejiang, 310053, China

[7] Research and Development Department, Zhejiang Healthfuture Institute for Cell-Based Applied Technology, Hangzhou, Zhejiang, China

[8] Center for Disease Control and Prevention of Jiading District, Shanghai, China

[9] Department of Health Management, Hangzhou Normal University, Hangzhou, Zhejiang, 311121, China

*Corresponding author: Hou-Feng Zheng. E-mail: zhenghoufeng@westlake.edu.cn

ABSTRACT

Purpose: The Westlake BioBank for Chinese (WBBC) pilot cohort is a population-based prospective study with its major purpose to better understand the effect of genetic and environmental factors on growth and development from adolescents to adults.

Participants: A total of 14,726 participants (4,751 males and 9,975 females) aged 14 - 25 years were recruited and the baseline survey was carried out from 2017 to 2019. The pilot cohort contains rich range of information regarding of demographics and anthropometric measurements, lifestyle and sleep patterns, clinical and health outcomes. Please visit the WBBC website for more information (https://wbbc.westlake.edu.cn/index.html).

Findings to date: The mean age of study sample were 18.6 years for males and 18.5 years for females, respectively. The mean height and weight were 172.9 cm and 65.81 kg for males, and 160.1 cm and 52.85 kg for females. Results indicated that the prevalence of underweight in female was much higher than male, but the prevalence of overweight and obesity in female was lower than male. The mean of serum 25(OH)D level in the 14,726 young participants was 22.4 ± 5.3 ng/ml, and male had a higher level of serum 25(OH)D than female, overall, 33.5% of the participants had vitamin D deficiency and even more participants suffered from vitamin D insufficiency (58.2%). The proportion of deficiency in females was much higher than that in males (41.8 vs. 16.4%). The issue of underweight and vitamin D deficiency in young people should be paid attention, especially in females. These results reflected the fact that thinness and paler skin are preferred in modern aesthetics of Chinese culture.

Future plans: WBBC pilot is designed as a prospective cohort study and provides a unique and rich data set analyzing health trajectories from adolescents to young adults. WBBC will continue to collect samples with old age.

Key words: Cohort, Late adolescence, Nutrition, Obesity.

Strengths and limitations

This is a comprehensive cohort study to pay attention to the health of adolescence and young adults, including over ~14,000 participants.

The WBBC pilot cohort is rich in longitudinal phenotypic data and archived biospecimens, including whole blood and serum as well as DNA samples. These resources will facilitate to digitize the information of the genomic, proteomic and metabolomic and microbiome of the participants.

Although the participants covered all around the country, most of them were mainly from 3 provinces (Jiangxi, Shandong and Zhejiang), and the selection of participants was not random.

The participants were collected in 3 phases at different time points, the phenotypic data were not always uniform in 3 phases.

Introduction

The Westlake BioBank for Chinese (WBBC) cohort is a population-based prospective study with its major purpose to better understand the effect of genetic and environmental factors on growth and development from adolescents to adults. WBBC is designed as a large-scale prospective cohort with its aim of recruiting at least 100,000 Chinese samples at different age. The pilot project of WBBC has focused on the study on the young population (Late adolescence), and has already collected a wide range of information including demographics and anthropometric measures, serological tests, physical activity, sleep quality, age at menarche and bone mineral density etc. The main purpose of this particular paper is to profile the cohort, therefore, only limited findings were reported and few questions were asked, e.g. (1) what is the prevalence of underweight, overweight, obesity and vitamin D deficiency in Chinese late adolescence? What is the reference value of serum vitamin D level with the young people? (2) what is the difference between male and female in term of height, weight, blood pressure, lifestyle and bone health in the young people.

It is known that adolescence is a period of life with marked psychosocial, behavioral and biological changes[1], therefore, monitoring the trajectories of health-risk behaviors in adolescents is one of the important concerns of public health. Among health problems during adolescence, overweight and obesity are highly prevalent[2 3]. Following the rapid economic development since the 1980s, China experiences a rapidly increasing of overweight and obesity among children and adolescents[4 5]. In 2019, a cross-sectional study[6] found that the prevalence of overweight in college students (aged 18-26 years) was 8.0%, and the prevalence of obesity was 3.5%. A recent study from 12 provinces in China showed that the prevalence of overweight and obesity were 14.0% and 10.5% in boys, and 9.7% and 7.1% in girls, respectively[7]. And obesity, in adolescence conferred very high risks for obesity in adults[2]; 70% of overweight adolescents had one or more concomitant conditions such as high blood pressures and fasting insulin, which were also risk factors for cardiovascular disease, and 23% of those accompanied with three or more concomitant conditions[8]. Further, elevated body mass index (BMI) in adolescence had been associated with several obesity-related morbidities in adult life, such as diabetes, metabolic syndromes and some types of cancer[9]. However, little is known about how underweight or severe low BMI could affect the health in adolescence, what are the long-term medical consequences. In adolescents, underweight was reported to be associated with scoliosis[10], pubertal delay[11], and psychiatric disorders[12]. Therefore, in this study, it is also important to recognize the issue of underweight in adolescents.

Hypovitaminosis D in children and adolescents is another re-emerging public health problem[13]. Besides its effect on the musculoskeletal system, Vitamin D showed a pleiotropic effect on human health, such as cardiovascular diseases[14], common infectious diseases[15] and autoimmune diseases[16]. Serum 25(OH)D is a good indicator of vitamin D storage and is an optimal method of assessing vitamin D levels[17]. According to the Endocrine Society clinical practice guidelines, vitamin D levels were defined as a deficiency [25(OH)D<20 ng/mL], insufficiency [25(OH)D: 20-29 ng/mL)] and sufficiency [25(OH)D≥30 ng/mL] respectively[18]. Many people in central and western Europe had vitamin D concentration of 11–20 ng/mL in winter [19]. Studies from other countries, including Canada[20], Japan[21], Australia[22] and Iran[23], presented similar situations, with high prevalence of vitamin D insufficiency in different ethnicities. A study in northern China found that more than 40% of adolescent girls had Vitamin D-deficient in the winter[24]. Another study in Shanghai showed that more than one-third newborns had plasma 25(OH)D less than 20 ng/mL[25]. Even in Hong Kong (latitude 22° north), 72% of young adults were reported to have vitamin D deficiency[26]. To the best of our knowledge, there was no large survey to assess levels of vitamin D deficiency in adolescents in China mainland; however, findings from some surveys suggest that vitamin D deficiency is prevalent in young people[26-28]. Moreover, the problem of vitamin D deficiency among young people in China has not been paid sufficient attention.

The overall goal of the Westlake Biobank for Chinese (WBBC) pilot cohort is to recruit individuals at their late adolescence/young adulthood. The biological samples such as whole blood, serum, urine and faeces were collected, genomic DNA was extracted and the DNA sequence information was acquired through sequencing technique. A long questionnaire with questions concerning the environmental factors such as nutrition, sleep quality, physical activity, medication etc was provided. These data will help us to understand the association between the genetics, environmental factors, microbiome and health statue of adolescence population. With a broad range of phenotype collection on many aspects of participants' daily life, a wide range of scientific questions could be addressed.

Cohort description

Sampling Design

The Westlake Biobank for Chinese (WBBC) pilot study was collected in three main regions in China (Zhejiang province, Jiangxi province and Shandong province), but the participants covered all around the country (Table 1 and Figure 1). The baseline survey was carried out from 2017 to 2019. The target population was young people aged 14–25 years who were college students and available for follow-up studies. In the first phase of baseline (WBBC pilot 1), the participants were recruited from two colleges at Zhejiang province and Jiangxi province in Southeast China from September 2017 to March 2018 (Figure 2), and 1,258 and 2,769 participants were from Zhejiang and Jiangxi provinces, respectively, and 1,263 participants were from other 25 provinces of China (Table 1). From September 2018 to December 2018, the second phase of WBBC pilot project was initiated (WBBC pilot 2), the participants were recruited at the same college in Jiangxi province and a college in Shandong province in Northeast China (Figure 2). There were 2,920 participants from Shandong province, 2,032 participants from Jiangxi province and 1,306 participants from other 28 provinces of China in WBBC pilot 2 (Table 1). From September 2019, the WBBC pilot project phase 3 (WBBC pilot 3) recruited participants from the same college in Jiangxi province (Figure 2), most of the participants (2,504) of WBBC pilot 3 were from Jiangxi province and 674 participants were from other 26 provinces of China (Table 1). All participants provided their Chinese unique national identity (ID) number for unique reference at the health examination center in the campus. The inclusion criteria were: a). All study participants signed the informed consent form before taking part in the survey. b). Participants should complete the physical examination, and should finish at least one of other items including bone mineral density scan, blood test and questionnaire. And the exclusion criteria were: a). Age < 14 year old or > 25 year old. b). Participants have taken drugs which could affect bone metabolism (e.g. glucocorticoids). c). Participants have illness which could cause secondary osteoporosis (e.g. Hyperparathyroidism). In WBBC pilot 3, the urine and faeces of the participants were collected, therefore, participants taking antibiotics should be excluded.

Data Collection Procedures

The WBBC pilot study is a multidisciplinary study and contains a rich range of information regarding of demographics, anthropometric measurements, blood pressure and heart rate, lifestyle and sleep patterns, biological, clinical and health outcomes. Genomic data are available for 7,033 participants. Data and samples were collected via examinations, questionnaire and venipuncture (Table 2).

Measurements of anthropometric parameters

Anthropometric data included height, body weight, bust, waist, hip and thigh circumference, resting blood pressure, heart rate and hand grip. Height was measured to the nearest 0.1 cm with participants' light-weight clothes and shoes off; weight was measured to the nearest 0.01 kg with the weight scale (Ultrasonic surveying instrument, Beryl BYH01, China) calibrated daily before each series of measurements. Bust, waist, hip and thigh circumference were measured to the nearest 0.5 cm by using a measuring tape with the subject standing comfortably. Resting blood pressure and heart rate were measured on the left arm supported at heart-level sitting position using electronic Sphygmomanometers (Yuwell YE660A, China). To ensure accurate data of resting blood pressure and heart rate, the participants were asked to take a rest for at least 5 minutes and have no excessive physical activity. Using a handgrip dynamometer (CAMRY EH101, China), grip strength with both hands were tested for most of the participants in WBBC pilot 2 and WBBC pilot 3. In order to get more accurate results, the participants should make sure the arm that's being tested was at a 90 degree angle at the elbow[29] until the test was finished. Details of the methods and instruments used for measurements of anthropometric parameters were provided in Table 3.

Biochemistry assessment

Participants came to the examination center in each college in the morning with at least 8 h of overnight fasting, about 25 ml of venous blood samples was collected for routine blood measurements, biochemical indexes, DNA extraction and so on. To ensure accurate data, the participants were asked not to have tea or alcohol intake or smoking for at least one night before blood sample collection. Venous blood samples were collected using ethylenediamine tetraacetic acid dipotassium (EDTAK2) anticoagulation tube (3×5.0 ml) and vacuum tube without anticoagulation (2×5.0 ml). Serum and plasma samples were separated from whole blood through centrifugation for 10 min at the relative centrifugal force 3000 g (Figure 3). Serum samples were forwarded to test biochemical indexes that included serum 25(OH)D level, serum calcium level, fasting blood glucose, kidney function test, hepatic function test, blood lipids [triglyceride, cholesterol, low-density lipoprotein (LDL), high-density lipoprotein (HDL)], triiodothyronine (T3), thyroid and parathyroid function and bone turnover markers (Table 2). Details of the platforms used for biochemical analysis are provided in Table 3. We also reserved serum samples (2×0.5ml) for each participant at - 80℃ for future use. Figure 3 displays the detail of the flow diagram of blood separation and detection of main blood biochemical indexes.

Questionnaire-based assessments

Baseline data collection for participants included a self-completion questionnaire. In Table 2, a list of core questions within the aforementioned domains, was provided. The questionnaire included social and demographic measures data (e.g. age, sex, ethnicity, family economic status and born place), menstrual history (for female), lifestyle (e.g. physical activity, smoking, alcohol, tea and coffee intake), additional supplement (e.g. calcium and vitamin D), health status and other information. In WBBC pilot 1, the questionnaire only had "yes" and "no" choices for current smoking and alcohol status, in WBBC pilot 2 and 3, we updated the questionnaire to include the frequency information for smoking and alcohol. Sleep duration and sleep quality were assessed by the Pittsburgh Sleep Quality Index (PSQI)[30]. This is composed of 19 questions which reflect seven major components, all seven components are then summed up to create a scale from 0–21 points.

Bone mineral density assessment

Bone mineral density (BMD) is expressed in terms of bone mass per cm^2 (g/cm^2) and were assessed by dual-energy X-ray absorptiometry[31] (DXA, Discovery QDR 4500; Hologic Inc., Waltham, MA, USA). In WBBC pilot, the main sites for BMD evaluation were lumbar spine (L1–L4), femur (femoral neck and total hip) and distal third of the radius.

Whole Genome sequencing and genotyping

In the process of collecting fasting venous blood, 5-ml whole blood was used for isolation of genomic DNA. These genomic DNA was used for whole genome sequencing and genotyping (Figure 3). Whole genome sequencing was completed by NovaSeq 6000 system (Illumina Co., Ltd), and for now, 1,192 participants have been sequenced at a mean depth of 14x, with highest depth of 65x. A Chinese specific reference panel will be constructed for imputation for the Chinese population. Whole genome genotyping was completed with Infinium Asian Screening Array (ASA) (Illumina Co., Ltd), and 5,841 participants have been genotyped in approximately

700,000 SNPs.

Follow-up and outcome measures

We are seeking funding to follow the cohort to examine the development and growth of the participants, and to investigate the effect of environmental factors on later outcomes. An important area of future research will focus on the development of bone mineral density and body weight from late adolescence to adulthood. Figure 2 shows the overall study plan. Follow-up surveys will be conducted according to the design of the subsequent research projects, the participants will be invited for survey with repeat interviews, including the questionnaire, anthropometric measurements, grip strength and bone mineral density collection as those used in the baseline stage and the data of nutritional status by food frequency questionnaire.

We have started a pilot follow-up study for WBBC pilot 2 since December 2019, and 1,303 participants had completed all examinations (Figure 2). The collected information included height, weight, grip strength and the updated questionnaire. Besides, we retested bone mineral density at the spine (L1-L4), hip and distal third of the radius.

Statistical analysis

To test the differences in means and proportions between male and female, we used T-test and Chi-square tests for continuous and categorical variables, respectively. All variables were presented by unadjusted proportions for categorical variables and unadjusted means with standard deviations (SD) for continuous variables. The variables demonstrating a p-value of less than 0.05 were considered statistically significant. All statistical analyses were analyzed using Stata 12.0 software.

Findings to date

This pilot cohort of WBBC is a large, longitudinal survey conducted among adolescents and young adults in China. We surveyed 14,726 young people aged 14–25 years who were college students and available for completing follow-up studies. The baseline survey was carried out from 2017 to 2019, including WBBC pilot 1 (5,290 participants), WBBC pilot 2 (6,258 participants) and WBBC pilot 3 (3,178 participants).

We have several ongoing projects under WBBC pilot study. One of the most significant ongoing projects is the whole genome sequencing (WGS) study for Chinese population, the results were reported elsewhere[32]. In brief, 4,535 whole-genome sequencing individuals and 5,481 high-density genotyping individuals were available in WBBC pilot project, covering 30 provincial regions of China. Since our previous study[33] demonstrated that the existing reference panels, such as the 1000 Genome Phase3 panel[34] and the HRC (Haplotype Reference Consortium) panel[35], were not the best fit for imputation for the Chinese population, our WGS data provided a population specific reference panel to improve the imputation accuracy of Chinese GWAS study. We also provided an online imputation server (https://imputationserver.westlake.edu.cn/) which could result in higher imputation accuracy compared to the existing panels, especially for lower frequency variants[32].

Given the extensive range of data collected in the WBBC study (https://wbbc.westlake.edu.cn/index.html), it is not feasible to present all the results, only limited findings were described in the present study. In summary, a total of 17,407 college students were invited, of whom, 14,983 (86.07 %) responded. After removing participants with missing data and invalid data, the final study included an effective sample size of 14,726 (84.60%) adolescents and young adults (with age from 14 to 25 years, and mean age 18.5 years). Table 4 provides an overview of social-demographic, anthropometry, cardiovascular system, lifestyle, grip strength and BMD characteristics of the WBBC pilot participants at baseline. Briefly, within the 14,726 samples, there were more females than males (67.7 vs. 32.3%), with a mean age of 18.5 years for female and 18.6 years for males, respectively. Most of the participants were Chinese Han ethnic (97.8%), and more than 60% of them were originally from rural areas (60.2% of males and 69.6% of females). For anthropometry measurements, the mean height and weight were 172.9 cm and 65.81 kg for males, and 160.1 cm and 52.9 kg for females; the mean waist, hip and thigh circumference were 75.5cm, 90.8cm and 51.6 cm in males, and 71.7cm, 89.8cm and 51.7 among females. The mean systolic blood pressure (SBP), diastolic blood pressure (DBP) and heart rate (HR) in participants were 113 mmHg, 71 mmHg and 86 beats/minute, respectively. In the cohort, only 5.5% of the participants were current smokers and 38.8% of them were regular drinkers. Regarding the current smoking status, there was a significant difference between males and females (16.3 vs. 1.6%, $p < 0.001$). As for alcohol, the proportion of current drinker in males and females was 62.0% and 29.2%, respectively, which is much higher in males ($p < 0.001$). The mean sleeping time estimated in females was higher than males (8.3 vs. 8.0 hours, $p < 0.001$). As for grip strength, the data collection was started from WBBC pilot 2, the mean of grip strength in males were much higher than females (grip-left: 36.66 vs. 27.38kg and grip-right: 39.90 vs. 29.76 kg, both $p < 0.001$).

Height and weight were measured using the standardized procedures. Body mass index (BMI) was calculated based on the formula: weight in kilograms divided by height in meters-squared (kg/m^2). According to the Working Group on Obesity in China (WGOC)[36], participants were defined as underweight (< 18.5 kg/m^2), normal weight (18.5-23.9 kg/m^2), overweight (24-27.9 kg/m^2) and obese (≥ 28 kg/m^2). Therefore, the WBBC pilot study provided an overall prevalence of underweight, overweight and obesity among young participants of 24.3%, 11.5% and 5.0%, respectively (Table 5). The prevalence of underweight in female was much higher than male (26.4% vs. 19.7%, $p < 0.0001$), but the prevalence of overweight in female was much lower than male (9.0% vs. 16.7%, $p < 0.0001$) (Table 5), similarly, the prevalence of obesity in female (3.2%) was lower than in male (8.9%) ($p < 0.0001$) (Table 5). Waist circumference (WC) is a good indicator of abdominal visceral fat distribution and is a strong predictor of diabetes mellitus and cardiovascular disease[37]. It is meaningful to investigate the WC along with BMI among adolescents and young people. In WBBC pilot study, central obesity was defined as WC≥85 cm for males and as WC≥80 cm for females based on the recommendations of the WGOC [36]. In the cohort of 12,396 participants, the prevalence of central obesity was 14.6%, which was higher in male than in female (19.1% vs. 12.6%, $p < 0.0001$) (Table 5).

In WBBC pilot study, the prevalence of underweight were high in both male (19.7 %) and female (26.4%), though the prevalence of moderate and severe underweight decreased from 9.2% in 1975 to 8.4% in 2016 in girls and from 14.8% in 1975 to 12.4% in 2016 in boys in the world[38]. These results might be due to the modern aesthetics of human stature that thinness is preferred, especially in China[39 40]. Recently, a study involving 2,023 young female participants (70.5% subjects aged 20-25 years) from eight Chinese universities[40] showed that 30.55% of the participants were underweight, and 57.39% of them would like to be much thinner, which would lead to more underweight individuals. A silhouette-matching test was administered in mainland China and found that the majority of the female participants indicated a preference to be more slender[41]. Their ideal figure was underweight and was far smaller than the most attractive female figure chosen by male participants[41]. Therefore, future studies should not only pay attention to the problem of obesity/overweight, but also to the underweight issue with young people.

Using simple anthropometric indices of body composition, such as BMI and WC, has been considered as a practical and valuable approach to the assessment of obesity for a long time. Waist-to-hip ratios (WHR), waist-to-height ratios (WHeR), a body shape index (ABSI)[42] and body roundness index (BRI)[43] were also as parameters of body fat and visceral adipose tissue volume. In WBBC pilotcohort, we had collected several anthropometric measures including height, weight, bust, waist, hip and thigh circumference and these data could help us examine the usefulness of these anthropometric parameters and identify the optimal cut-off of the parameters to evaluate overweight and obesity among adolescents and young people in future study.

In WBBC pilot study, the mean serum 25(OH)D level was 22.4 ± 5.3 ng/ml for all the participants (male: 25.2 ng/mL, and female: 21.1 ng/mL, $p < 0.0001$) (Table 6). Overall, 33.5% of the participants had vitamin D deficiency and even more participants suffered from vitamin D insufficiency (58.2%) (Table 6). In addition, the proportion of females with sufficient vitamin D was much lower than that of males (3.7 vs. 17.9%, $p < 0.0001$), while the proportion of deficiency in females was much higher than that in males (41.8 vs. 16.4%, $p < 0.0001$) (Table 6). Most of the participants (86.9%) preferred to stay indoors in spare time, the females were less willing to do exercise than males (53.7% vs 70.6%) (Table 7), and 44.6% of females hardly had outdoor activities, only 5.9% of females often had outdoor activities every week (Table 7). These results, jointly suggested that the females had not enough sun exposure. Although food sources of vitamin D were not commonly recognized, only 10–20% of vitamin D in human bodies was obtained through food sources[44]. In WBBC pilot study, there was only 2.2% of the participants used vitamin D supplements (3.1% in male and 1.9% in female, $p = 0.00098$) (Table 6) and this was consistent with Zhou et al[45], which found that only 5.6 % of the students used vitamin D supplements in a university of Nanjing, China. It is noteworthy that vitamin D deficiency in females was significantly worse than in males. This may due to the modern aesthetics of Chinese culture that paler skin is preferred, especially in females. A questionnaire related to vitamin D and sun exposure was conducted at a university in Nanjing, China and found that 75.0 % of the students lacked sun exposure because they would like to avoid dark skin[45]. In addition, most of the students (82.7 %) used sun protection, and sunscreen use was more popular in females[45]. However, it was reported that using the amount of sun cream recommended by World Health Organization exponentially suppressed vitamin D synthesis in the skin[46].

Collaboration

Participants have agreed to provide their pseudonymized data being made available to other approved researchers. The WBBC pilot study welcomes and offers global collaboration. The data is not freely available in the public domain, but specific proposals and ideas for future collaboration would be very welcome. Applicants for collaboration and more information are encouraged to contact Dr. Hou-Feng Zheng (Email address: zhenghoufeng@westlake.edu.cn), the person in charge of this project.

Funding

This study was supported by the National Natural Science Foundation of China (32061143019 and 81871831), and by the Westlake Biobank for Chinese (WBBC) funds from Westlake University.

Acknowledgements

We gratefully acknowledge all the participants of this project, and thank all the people who helped us in the sample recruitment, including volunteers, laboratory technicians, nurses and clerical workers. We also thank the High-Performance Computing Center at Westlake University for the facility support and technical assistance. In addition, we would like to thank Shanghai AvanTech BioSciences Co., Ltd for their support and assistance in sample storage and management.

Contributors

H.-F.Z. gained funds and conceived of the study, and K.-Q.L., P.-Y.W., J.-Y.C., X.-J.X, J.-J.X., M.-C.Q., C.L. and S.-Y.X. were involved in the sample collection. J.-Q.L. and Y.S. were involved in the blood biochemistry testing. W.-Y.B., P.-P.Z., J.X., S.-R.G., P.-L.G., Y.Q. and P.-K.C. were involved in the DNA extraction. X.-W.Z. and H.-F.Z. analyzed the data and wrote the paper. All authors read and approved the final manuscript.

Conflict of interest

Jun-Quan Liu and Yi Sun are employees of Hangzhou Kingmed Diagnostics Co., Ltd. The other authors have no conflict of interest to declare.

Patient and public involvement

It was not appropriate or possible to involve patients or the public in the design, or conduct, or reporting, or dissemination plans of our research.

Ethics Approval and Consent to Participate

The study protocol and informed consent procedure were approved by the Ethics Committees at Westlake University. All study participants signed the informed consent form before taking part in the study.

References

1. Steinbeck K, Hazell P, Cumming RG, et al. The study design and methodology for the ARCHER study--adolescent rural cohort study of hormones, health, education, environments and relationships. BMC pediatrics 2012;**12**:143 doi: 10.1186/1471-2431-12-143[published Online First: Epub Date]|.

2. Patton GC, Coffey C, Carlin JB, et al. Overweight and obesity between adolescence and young adulthood: a 10-year prospective cohort study. The Journal of adolescent health : official publication of the Society for Adolescent Medicine 2011;**48**(3):275-80 doi: 10.1016/j.jadohealth.2010.06.019[published Online First: Epub Date]|.

3. Sawyer SM, Afifi RA, Bearinger LH, et al. Adolescence: a foundation for future health. Lancet 2012;**379**(9826):1630-40 doi: 10.1016/S0140-6736(12)60072-5[published Online First: Epub Date]|.

4. Yu Z, Han S, Chu J, et al. Trends in overweight and obesity among children and adolescents in China from 1981 to 2010: a meta-analysis. PloS one 2012;**7**(12):e51949 doi: 10.1371/journal.pone.0051949[published Online First: Epub Date]|.

5. Ng M, Fleming T, Robinson M, et al. 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. Lancet 2014;**384**(9945):766-81 doi: 10.1016/S0140-6736(14)60460-8[published Online First: Epub Date]|.

6. Jiang Y, Wang J, Wu S, et al. Association between Take-Out Food Consumption and Obesity among Chinese University Students: A Cross-Sectional Study. International journal of environmental research and public health 2019;**16**(6) doi: 10.3390/ijerph16061071[published Online First: Epub Date]|.

7. Zhang J, Wang H, Wang Z, et al. Prevalence and stabilizing trends in overweight and obesity among children and adolescents in China, 2011-2015. BMC public health 2018;**18**(1):571 doi: 10.1186/s12889-018-5483-9[published Online First: Epub Date]|.

8. Freedman DS, Kahn HS, Mei Z, et al. Relation of body mass index and waist-to-height ratio to cardiovascular disease risk factors in children and adolescents: the Bogalusa Heart Study. The American journal of clinical nutrition 2007;**86**(1):33-40 doi: 10.1093/ajcn/86.1.33[published Online First: Epub Date]|.

9. Llewellyn A, Simmonds M, Owen CG, et al. Childhood obesity as a predictor of morbidity in adulthood: a systematic review and meta-analysis. Obesity reviews : an official journal of the International Association for the Study of Obesity 2016;**17**(1):56-67 doi: 10.1111/obr.12316[published Online First: Epub Date]|.

10. Jeon K, Kim DI. The Association between Low Body Weight and Scoliosis among Korean Elementary School Students. International journal of environmental research and public health 2018;**15**(12) doi: 10.3390/ijerph15122613[published Online First: Epub Date]|.

11. He Q, Karlberg J. Bmi in childhood and its association with height gain, timing of puberty, and final height. Pediatric research 2001;**49**(2):244-51 doi: 10.1203/00006450-200102000-00019[published Online First: Epub Date]|.

12. Hudson LD, Nicholls DE, Lynn RM, et al. Medical instability and growth of children and adolescents with early onset eating disorders. Archives of disease in childhood 2012;**97**(9):779-84 doi: 10.1136/archdischild-2011-301055[published Online First: Epub Date]|.

13. Antonucci R, Locci C, Clemente MG, et al. Vitamin D deficiency in childhood: old lessons and current challenges. Journal of pediatric endocrinology & metabolism : JPEM 2018;**31**(3):247-60 doi: 10.1515/jpem-2017-0391[published Online First: Epub Date]|.
14. Trehan N, Afonso L, Levine DL, et al. Vitamin D Deficiency, Supplementation, and Cardiovascular Health. Critical pathways in cardiology 2017;**16**(3):109-18 doi: 10.1097/HPC.0000000000000122[published Online First: Epub Date]|.
15. Watkins RR, Lemonovich TL, Salata RA. An update on the association of vitamin D deficiency with common infectious diseases. Canadian journal of physiology and pharmacology 2015;**93**(5):363-8 doi: 10.1139/cjpp-2014-0352[published Online First: Epub Date]|.
16. Illescas-Montes R, Melguizo-Rodriguez L, Ruiz C, et al. Vitamin D and autoimmune diseases. Life sciences 2019;**233**:116744 doi: 10.1016/j.lfs.2019.116744[published Online First: Epub Date]|.
17. Holick MF. Vitamin D deficiency. The New England journal of medicine 2007;**357**(3):266-81 doi: 10.1056/NEJMra070553[published Online First: Epub Date]|.
18. Holick MF, Binkley NC, Bischoff-Ferrari HA, et al. Evaluation, treatment, and prevention of vitamin D deficiency: an Endocrine Society clinical practice guideline. The Journal of clinical endocrinology and metabolism 2011;**96**(7):1911-30 doi: 10.1210/jc.2011-0385[published Online First: Epub Date]|.
19. Pludowski P, Grant WB, Bhattoa HP, et al. Vitamin d status in central europe. International journal of endocrinology 2014;**2014**:589587 doi: 10.1155/2014/589587[published Online First: Epub Date]|.
20. Newhook LA, Sloka S, Grant M, et al. Vitamin D insufficiency common in newborns, children and pregnant women living in Newfoundland and Labrador, Canada. Maternal & child nutrition 2009;**5**(2):186-91 doi: 10.1111/j.1740-8709.2008.00157.x[published Online First: Epub Date]|.
21. Yoshimura N, Muraki S, Oka H, et al. Profiles of vitamin D insufficiency and deficiency in Japanese men and women: association with biological, environmental, and nutritional factors and coexisting disorders: the ROAD study. Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA 2013;**24**(11):2775-87 doi: 10.1007/s00198-013-2372-z[published Online First: Epub Date]|.
22. Hirani V, Cumming RG, Blyth FM, et al. Vitamin D status among older community dwelling men living in a sunny country and associations with lifestyle factors: the Concord Health and Ageing in Men Project, Sydney, Australia. The journal of nutrition, health & aging 2013;**17**(7):587-93 doi: 10.1007/s12603-013-0013-z[published Online First: Epub Date]|.
23. Faghih S, Abdolahzadeh M, Mohammadi M, et al. Prevalence of vitamin d deficiency and its related factors among university students in shiraz, iran. International journal of preventive medicine 2014;**5**(6):796-9
24. Fraser DR. Vitamin D-deficiency in Asia. The Journal of steroid biochemistry and molecular biology 2004;**89-90**(1-5):491-5 doi: 10.1016/j.jsbmb.2004.03.057[published Online First: Epub Date]|.
25. Yu X, Wang W, Wei Z, et al. Vitamin D status and related factors in newborns in Shanghai, China. Nutrients 2014;**6**(12):5600-10 doi: 10.3390/nu6125600[published Online First: Epub Date]|.
26. Wang EW, Pang MY, Siu PM, et al. Vitamin D status and cardiometabolic risk factors in young

adults in Hong Kong: associations and implications. Asia Pacific journal of clinical nutrition 2018;**27**(1):231-37 doi: 10.6133/apjcn.022017.08[published Online First: Epub Date]|.

27. Guo Y, Ke HJ, Liu Y, et al. Prevalence of vitamin D insufficiency among children in southern china: A cross-sectional survey. Medicine 2018;**97**(25):e11030 doi: 10.1097/MD.0000000000011030[published Online First: Epub Date]|.

28. Yan X, Zhang N, Cheng S, et al. Gender Differences in Vitamin D Status in China. Medical science monitor : international medical journal of experimental and clinical research 2019;**25**:7094-99 doi: 10.12659/MSM.916326[published Online First: Epub Date]|.

29. Roberts HC, Denison HJ, Martin HJ, et al. A review of the measurement of grip strength in clinical and epidemiological studies: towards a standardised approach. Age and ageing 2011;**40**(4):423-9 doi: 10.1093/ageing/afr051[published Online First: Epub Date]|.

30. Buysse DJ, Reynolds CF, 3rd, Monk TH, et al. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry research 1989;**28**(2):193-213 doi: 10.1016/0165-1781(89)90047-4[published Online First: Epub Date]|.

31. Bazzocchi A, Ponti F, Albisinni U, et al. DXA: Technical aspects and application. European journal of radiology 2016;**85**(8):1481-92 doi: 10.1016/j.ejrad.2016.04.004[published Online First: Epub Date]|.

32. Cong P, Bai W-Y, Li J, et al. Genomic analyses of 10,376 individuals provides comprehensive map of genetic variations, structure and reference haplotypes for Chinese population. bioRxiv 2021:2021.02.06.430086 doi: 10.1101/2021.02.06.430086[published Online First: Epub Date]|.

33. Bai WY, Zhu XW, Cong PK, et al. Genotype imputation and reference panel: a systematic evaluation on haplotype size and diversity. Briefings in bioinformatics 2019 doi: 10.1093/bib/bbz108[published Online First: Epub Date]|.

34. Genomes Project C, Auton A, Brooks LD, et al. A global reference for human genetic variation. Nature 2015;**526**(7571):68-74 doi: 10.1038/nature15393[published Online First: Epub Date]|.

35. McCarthy S, Das S, Kretzschmar W, et al. A reference panel of 64,976 haplotypes for genotype imputation. Nature genetics 2016;**48**(10):1279-83 doi: 10.1038/ng.3643[published Online First: Epub Date]|.

36. Zhou BF, Cooperative Meta-Analysis Group of the Working Group on Obesity in C. Predictive values of body mass index and waist circumference for risk factors of certain related diseases in Chinese adults--study on optimal cut-off points of body mass index and waist circumference in Chinese adults. Biomedical and environmental sciences : BES 2002;**15**(1):83-96

37. Asia Pacific Cohort Studies C. Central obesity and risk of cardiovascular disease in the Asia Pacific Region. Asia Pacific journal of clinical nutrition 2006;**15**(3):287-92

38. Collaboration NCDRF. 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. Lancet 2017;**390**(10113):2627-42 doi: 10.1016/S0140-6736(17)32129-3[published Online First: Epub Date]|.

39. Chen H, Jackson T. Predictors of changes in weight esteem among mainland Chinese adolescents: a longitudinal analysis. Developmental psychology 2009;**45**(6):1618-29 doi: 10.1037/a0016820[published Online First: Epub Date]|.

40. Zhang L, Qian H, Fu H. To be thin but not healthy - The body-image dilemma may affect health among female university students in China. PloS one 2018;**13**(10):e0205282 doi: 10.1371/journal.pone.0205282[published Online First: Epub Date]|.

41. Wang K, Liang R, Ma ZL, et al. Body image attitude among Chinese college students. PsyCh journal 2018;**7**(1):31-40 doi: 10.1002/pchj.200[published Online First: Epub Date]|.

42. Krakauer NY, Krakauer JC. A new body shape index predicts mortality hazard independently of body mass index. PloS one 2012;**7**(7):e39504 doi: 10.1371/journal.pone.0039504[published Online First: Epub Date]|.

43. Thomas DM, Bredlau C, Bosy-Westphal A, et al. Relationships between body roundness with body fat and visceral adipose tissue emerging from a new geometrical model. Obesity 2013;**21**(11):2264-71 doi: 10.1002/oby.20408[published Online First: Epub Date]|.

44. Mithal A, Wahl DA, Bonjour JP, et al. Global vitamin D status and determinants of hypovitaminosis D. Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA 2009;**20**(11):1807-20 doi: 10.1007/s00198-009-0954-6[published Online First: Epub Date]|.

45. Zhou M, Zhuang W, Yuan Y, et al. Investigation on vitamin D knowledge, attitude and practice of university students in Nanjing, China. Public health nutrition 2016;**19**(1):78-82 doi: 10.1017/S1368980015000373[published Online First: Epub Date]|.

46. Faurschou A, Beyer DM, Schmedes A, et al. The relation between sunscreen layer thickness and vitamin D production after ultraviolet B exposure: a randomized clinical trial. The British journal of dermatology 2012;**167**(2):391-5 doi: 10.1111/j.1365-2133.2012.11004.x[published Online First: Epub Date]|.

Tables

Table 1. The participants of the cohort at baseline.

Year	Phase	Area	Total	Sex	
				Male	Female
2017	WBBC pilot 1	Zhejiang province	1,258	460	798
		Jiangxi province	2,769	862	1,907
		Other 25 provinces	1,263	336	927
2018	WBBC pilot 2	Shandong province	2,920	1,115	1,805
		Jiangxi province	2,032	578	1,454
		Other 28 provinces	1,306	438	868
2019	WBBC pilot 3	Jiangxi province	2,504	761	1,743
		Other 26 provinces	674	201	473
Total			14,726	4,751	9,975

Table 2. Summary of data collected in the Cohort.

Measures	Instruments
Anthropometry*	
	Height and weight, bust, waist, hip and thigh circumference
Strength	
	Grip strength (right and left hand)
Cardiovascular system	
	Blood pressure (SBP & DBP), heart rate
Bone mineral density	
	Lumar spine, hip and forearm
Questionnaires	
Demographic data	Age, sex, ethnicity, family economic status, birthplace
Menstrual history (for female)	Age of menarche, Gynaecological disease history
Sleeping situation	Pittsburgh sleep quality index, PSQI
Lifestyle	Smoking status, alcohol, tea and coffee consumption, exercise
Supplementations	Calcium, vitamin D
Health status	Medical history and medications
Physical activity[#]	International Physical Activity Questionnaire, IPAQ
Dietary pattern[#]	Semi-Quantitative Food Frequency Questionnaire (SQFFQ)
Routine biochemistry and Haematology	
Lipid metabolism	Total cholesterol, triglycerides, HDL-cholesterol, LDL-cholesterol
Kidney-related measures	uric acid, creatinine and urea
Hepatic-related measures	Alanine aminotransferase, glutamic oxaloacetylase, total bilirubin, direct bilirubin, albumin, globulin
Thyroid and parathyroid function	TSH, PTH, T3, T4, FT3, FT4,
Routine diabetes test	Fasting plasma glucose, glycosylated hemoglobin (HbA1c)
Bone-related measures	Bone turnover markers (Osteocalcin, PINP, NBAP and β-CTX), serum $1,25(OH)_2D_3$, seruum IGF-1
Mineral elements	Calcium (Ca), phosphorus (P) and magnesium (Mg)
Samples	
Blood	25 ml
Serum	2 ml
Urine[#]	10 ml
Faeces[#]	15 g

* Direct measures carried out by trained interviewers according to standardized protocols; [#] measured only on WBBC pilot 3; BAP: Bone alkaline phosphatase; DBP: Diastolic blood pressure; FT3: free Triiodothyronine; FT4: free Tetraiodothyronine; HDL: High-density lipoprotein; IGF-1: Insulin-like growth factor 1; LDL: Low-density lipoprotein; PINP: N-terminal propeptide of procollagen type I; PTH: Parathyroid hormone; SBP: Systolic blood pressure; T3: Triiodothyronine; T4: Tetraiodothyronine; TSH: Thyroid stimulating hormone; β-CTx: β-isomerized C-terminal telopeptides

Table 3. List of anthropometric collected and platforms for biochemical tests at baseline in the cohort.

Variables	Analysis method/platform used
Height and weight	Ultrasonic surveying instrument, Beryl BYH01, China
Waist, hip and thigh circumference	Manufactured instrument (tape)
Resting blood pressure	Electronic sphygmomanometer, Yuwell YE660A, China
Heart rate	Electronic sphygmomanometer, Yuwell YE660A, China
Grip strength	Digital hand dynamometer, CAMRY EH101, China
Bone mineral density (BMD)	DXA, Discovery QDR 4500, Hologic Inc., Waltham, MA, USA
Blood routine (5 items)	SYSMEX 2100, Japan
Serum IGF-1	CLIA, DPC immulite 2000, Siemens, Germany
Serum calcium, phosphorus	ARCHITECT C16000, Abbott, USA
Fasting blood glucose	Cobas c501, Roche, Switzerland
Bone turnover markers	ECLIA, Cobas e602, Roche, Switzerland
Serum $1,25(OH)_2D_3$	LC-MS/MS, AB Sciex API 4000™, USA
Thyroid and parathyroid function	Chemical luminescence, ARCHITECT System i2000, USA
Lipid metabolism	ARCHITECT C16000, Abbott, USA
Kidney-related measures	ARCHITECT C16000, Abbott, USA
Hepatic-related measures	ARCHITECT C16000, Abbott, USA

CLIA: chemiluminescent immunoassay; ECLIA: electrochemoluminescence immunoassay; LC-MS/MS: Liquid chromatography-tandem mass spectrometry.

Table 4. Basic characteristics of participants in baseline of WBBC pilot.

Variables(unit)		Total	Sex		*P value
			Male	Female	
Socio-demographic					
Age (years)	N[a]	14,726	4,751	9,975	0.001
	M (SD)[b]	18.5 (1.3)	18.6 (1.2)	18.5 (1.3)	
Gender	N (%)[c]	14,726 (100)	4,751 (32.3)	9,975 (67.7)	<0.001
Ethnicity					0.305
Han	N (%)	11,136 (97.8)	3,642 (98.0)	7,494 (97.7)	
Others	N (%)	253 (2.2)	75 (2.0)	178 (2.3)	
Hukou status					<0.001
Rural	N (%)	5,951 (66.6)	1,739 (60.2)	4,212 (69.6)	
Urban	N (%)	2,986(33.4)	1,148 (39.8)	1,838 (30.4)	
Anthropometry					
Height (cm)	N	14,277	4,588	9,689	<0.001
	M (SD)	164.2 (8.5)	172.9 (6.6)	160.1 (5.8)	
Weight (kg)	N	14,279	4,587	9,692	<0.001
	M (SD)	57.01 (12.32)	65.81 (13.60)	52.85 (9.07)	
Waist (cm)	N	12,396	3,905	8,491	<0.001
	M (SD)	71.7 (9.8)	75.5 (11.1)	69.9 (8.7)	
Hip (cm)	N	12,388	3,902	8,486	<0.001
	M (SD)	89.8 (7.3)	90.8 (8.3)	89.3 (6.7)	
Thigh (cm)	N	12,351	3,880	8,471	0.102
	M (SD)	51.7 (5.7)	51.6 (6.5)	51.7 (5.3)	
Cardiovascular system					
SBP (mmHg)	N	14,277	4,595	9,682	<0.001
	M (SD)	113 (12)	121 (12)	110 (11)	
DBP (mmHg)	N	14,276	4,595	9,681	<0.001
	M (SD)	71 (9)	73 (9)	70 (8)	
Heart rate (beats/minute)	N	14,295	4,599	9,696	<0.001
	M (SD)	86 (13)	83 (13)	87 (13)	
Lifestyle					
Smoking	N (%)	435 (5.9)	350 (16.3)	85 (1.6)	<0.001
Alcohol status	N (%)	2,844 (38.8)	1,333 (62.0)	1,511 (29.2	<0.001
Sleeping time (hours)	N	7,247	2,115	5,132	
	M (SD)	8.2 (1.4)	8.0 (1.3)	8.3 (1.4)	<0.001
Grip strength (kg)					

Left hand	N	8,932	2,958	5,974	<0.001
	M (SD)	27.38 (8.77)	36.66 (7.53)	22.79 (4.82)	
Right hand	N	8,941	2,967	5,974	<0.001
	M (SD)	29.76 (9.66)	39.90 (8.32)	24.72 (5.35)	
BMD (g/cm^2)					
Lumar spine	N	10,154	3,293	6,861	<0.001
	M (SD)	0.910 (0.105)	0.926 (0.112)	0.903(0.100)	
Total hip	N	10,160	3,296	6,864	<0.001
	M (SD)	0.868 (0.127)	0.932 (0.139)	0.837 (0.108)	
Femoral neck	N	10,160	3,296	6,864	<0.001
	M (SD)	0.778 (0.125)	0.846 (0.138)	0.746 (0.104)	
Forearm	N	9,917	3,238	6,679	<0.001
	M (SD)	0.657 (0.059)	0.705 (0.057)	0.634 (0.453)	

[a] N: sample size; [b]M (SD): mean (standard deviation); [c]N (%): sample size (percentage); * T-test and Chi-square tests for continuous and categorical variables respectively to refer the significant differences between males and females; DBP: Diastolic blood pressure; SBP: Systolic blood pressure.

Table 5. Distribution of BMI or waist by sex in participants in WBBC pilot, 2017-2019. N=14,264.

Variables	**Total**	**Male**	**Female**	**#P**
Underweight, N (%)	3,460 (24.3%)	903 (19.7%)	2,557 (26.4%)	<0.0001
Normal, N (%)	8,446 (59.2%)	2,507 (54.7%)	5,939 (61.4%)	<0.0001
Overweight, N (%)	1,640 (11.5%)	766 (16.7%)	874 (9.0%)	<0.0001
Obesity, N (%)	718 (5.0%)	407 (8.9%)	311 (3.2%)	<0.0001
Central Obesity[δ], N (%)	1,812 (14.6%)	746 (19.1%)	1,066 (12.6%)	<0.0001

[δ] Total sample size for Central Obesity was 12,396. [#] Chi-square tests was calculated between male and female. Participants were defined as underweight (< 18.5 kg/m2), normal weight (18.5-23.9 kg/m2), overweight (24-27.9 kg/m2) and obese (⩾ 28 kg/m2). Central obesity was defined as WC⩾85 cm for males and as WC⩾80 cm for females.

Table 6. Distribution of Vitamin D level by sex in subjects in WBBC pilot, 2017-2018. N=11,370.

Variables	**Total**	**Male**	**Female**	**P**
Serum 1,25(OH)D, mean (SD)	22.4 (5.3)	25.2 (5.4)	21.1 (4.7)	<0.0001[η]
Vitamin D deficient, N (%)	3,806 (33.5%)	610 (16.4%)	3,196 (41.8%)	<0.0001[#]
Vitamin D insufficient, N (%)	6,614 (58.2%)	2,452 (65.7%)	4,162 (54.5%)	<0.0001[#]
Vitamin D sufficient, N (%)	950 (8.4%)	668 (17.9%)	282 (3.7%)	<0.0001[#]
Vitamin D supplementation[δ], N (%)	164 (2.2%)	67 (3.1%)	97 (1.9%)	0.00098[#]

[δ] The sample size for Vitamin D supplementation was 7,326. [#] Chi-square tests was calculated between male and female. [η] T-test was calculated between male and female.

Table 7. Participants' general react about activity status in WBBC pilot, 2017-2018. N=7,326.[a]

	Total	**Male**	**Female**
What would you like to do in spare time?			
Stay indoors	6,364 (86.9%)	1,754 (81.6%)	4,610 (89.1%)
Take part in some activities	962 (13.1%)	395 (18.4%)	567 (11.0. %)
Do you do exercise initiatively?			
Yes	4,296 (58.6%)	1,517 (70.6%)	2,779 (53.7%)
No	3,030 (41.4%)	632 (29.4%)	2,398 (46.3%)
How often do you have outdoor activities every week?			
Hardly	2,785 (38.0%)	477 (22.2%)	2,308 (44.6%)
Occasionally	3,865 (52.8%)	1,301 (60.5%)	2,564 (49.5%)
Often	676 (9.2%)	371 (17.3%)	305 (5.9%)

[a] Data are showed as n (%) of participants.

Figure legend

Figure 1. Maps showing the sources of the samples in the Westlake BioBank for Chinese (WBBC) pilot cohort. A) provides a range of sample size in each province with different color, and B) provides the exact number of sample size in each province. AH: Anhui province; BJ: Beijing; CQ: Chongqing; FJ: Fujian province; GD: Guangdong province; GS: Gansu province; GX: Guangxi Autonomous Region; GZ: Guizhou province; HA: Henan province; HB: Hubei province; HE: Hebei province; HI: Hainan province; HL: Heilongjiang province; HN: Hunan province; JL: Jilin province; JS: Jiangsu province; JX: Jiangxi province; LN: Liaoning province; NM: Neimeng Autonomous Region; NX: Ningxia Autonomous Region; QH: Qinghai province; SC: Sichuang province; SD: Shandong province; SH: Shanghai; SN: Shanxi province; SX: Shanxi province; TJ: Tianjin; XJ: Xinjiang Autonomous Region; XZ: Xizang Autonomous Region; YN: Yunnan province; ZJ: Zhejiang province.

Figure 2. Data collection timeline. IPAQ: International Physical Activity Questionnaire; PA: Physical Activity; PSQI: Pittsburgh sleep quality index; SQFFQ: Semi-Quantitative Food Frequency Questionnaire.

Figure 3. Flow diagram of main blood biochemical detection and blood conservation.
RCF: relative centrifugal force; WGS: whole genome sequence.

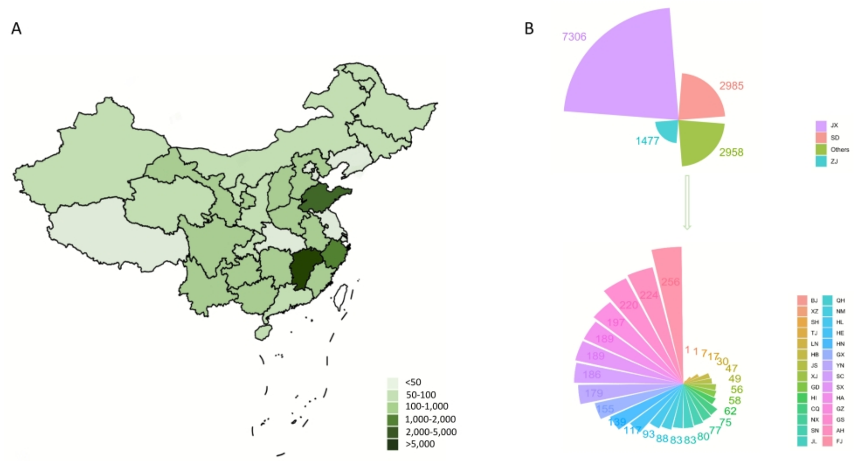

Figure 1. Maps showing the sources of the samples in the Westlake BioBank for Chinese (WBBC) pilot cohort. A) provides a range of sample size in each province with different color, and B) provides the exact number of sample size in each province. AH: Anhui province; BJ: Beijing; CQ: Chongqing; FJ: Fujian province; GD: Guangdong province; GS: Gansu province; GX: Guangxi Autonomous Region; GZ: Guizhou province; HA: Henan province; HB: Hubei province; HE: Hebei province; HI: Hainan province; HL: Heilongjiang province; HN: Hunan province; JL: Jilin province; JS: Jiangsu province; JX: Jiangxi province; LN: Liaoning province; NM: Neimeng Autonomous Region; NX: Ningxia Autonomous Region; QH: Qinghai province; SC: Sichuang province; SD: Shandong province; SH: Shanghai; SN: Shanxi province; SX: Shanxi province; TJ: Tianjin; XJ: Xinjiang Autonomous Region; XZ: Xizang Autonomous Region; YN: Yunnan province; ZJ: Zhejiang province.

338x190mm (400 x 400 DPI)

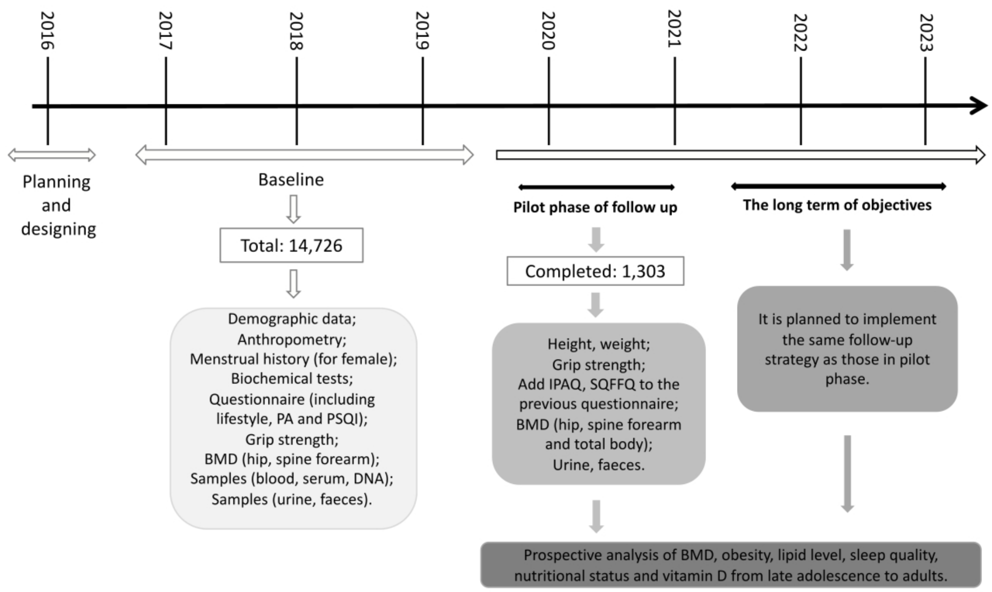

Figure 2. Data collection timeline. IPAQ: International Physical Activity Questionnaire; PA: Physical Activity; PSQI: Pittsburgh sleep quality index; SQFFQ: Semi-Quantitative Food Frequency Questionnaire.

254x190mm (400 x 400 DPI)

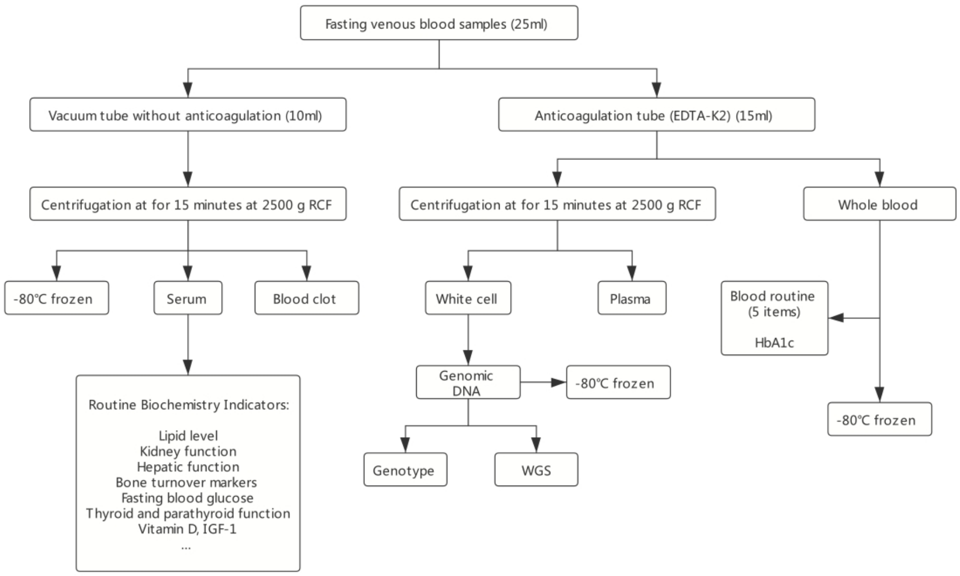

Figure 3. Flow diagram of main blood biochemical detection and blood conservation.
RCF: relative centrifugal force; WGS: whole genome sequence.

363x223mm (400 x 400 DPI)