


BMJ Open 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 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,^{1,2,3} Shu-Yang Xie,⁵ Hou-Feng Zheng ^{1,2,3}

To cite: Zhu X-W, Liu K-Q, Wang P-Y, *et al.* Cohort profile: the Westlake BioBank for Chinese (WBBC) pilot project. *BMJ Open* 2021;**11**:e045564. doi:10.1136/bmjopen-2020-045564

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-045564>).

Received 05 October 2020
Accepted 01 June 2021

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 (4751 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 and anthropometric measurements, lifestyle and sleep patterns, clinical and health outcomes. Visit the WBBC website for more information (<https://wbcc.westlake.edu.cn/index.html>).

Findings to date The mean age of the study samples 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 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 analysing health trajectories from adolescents to young adults. WBBC will continue to collect samples with old age.

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 cohort with its aim to recruit at least 100 000 Chinese samples at different age. The pilot project of WBBC has focused

Strengths and limitations of this study

- This is a comprehensive cohort study to pay attention to the health of adolescence and young adults, including over ~14 000 participants.
- The Westlake BioBank for Chinese (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 digitise 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 three provinces (Jiangxi, Shandong and Zhejiang), and the selection of participants was not random.
- The participants were collected in three phases at different time points; the phenotypic data were not always uniform in three phases.

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, bone mineral density and so on. The main purpose of this particular paper is to profile the cohort; therefore, only limited findings were reported and few questions were asked, for example (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 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.

It is known that adolescence is a period of life with marked psychosocial, behavioural and biological changes¹; therefore, monitoring the trajectories of health-risk behaviours in adolescents is one of the important concerns of public health. Among health problems during adolescence, overweight and obesity



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Hou-Feng Zheng;
zhenghoufeng@westlake.edu.cn



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 study⁶ 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 adults²; 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 delay¹¹ and psychiatric disorders.¹² Therefore, in this study, it is also important to recognise 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 diseases¹⁵ 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 Australia²² 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 WBBC pilot cohort is to recruit individuals at their late adolescence/young adulthood.

Table 1 The participants of the cohort at baseline

Year	Phase	Area	Total	Sex	
				Male	Female
2017	WBBC pilot 1	Zhejiang province	1258	460	798
		Jiangxi province	2769	862	1907
		Other 25 provinces	1263	336	927
2018	WBBC pilot 2	Shandong province	2920	1115	1805
		Jiangxi province	2032	578	1454
		Other 28 provinces	1306	438	868
2019	WBBC pilot 3	Jiangxi province	2504	761	1743
		Other 26 provinces	674	201	473
Total			14726	4751	9975

WBBC, Westlake BioBank for Chinese.

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, and so on, was provided. These data will help us to understand the association between the genetics, environmental factors, microbiome and health status 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 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 1258 and 2769 participants were from Zhejiang and Jiangxi provinces, respectively, and 1263 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

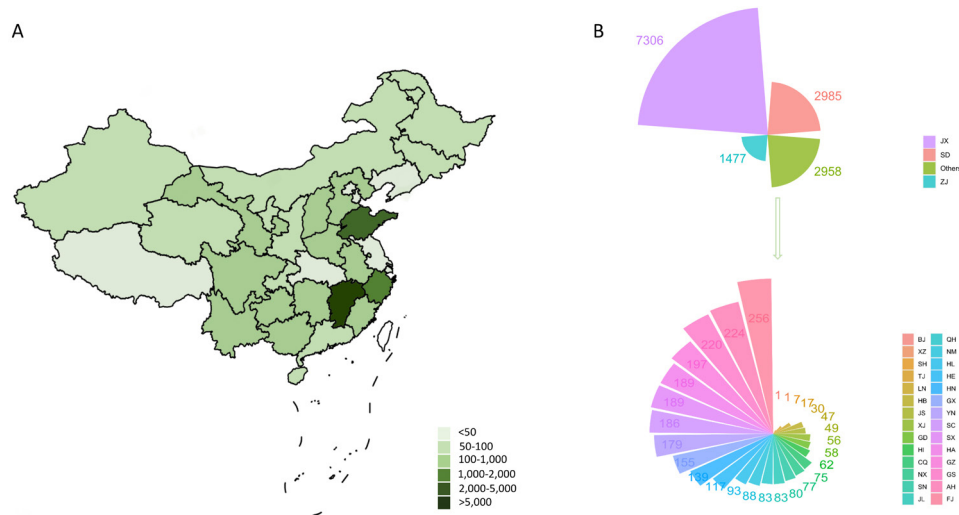


Figure 1 Maps showing the sources of the samples in the Westlake BioBank for Chinese (WBBC) pilot cohort. (A) provides a range of sample size in each province with different colour, and (B) provides the exact number of sample size in each province. AH, Anhui province; BJ, Beijing; CQ, Chongqing; FJ, 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, Sichuan 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.

in Shandong province in Northeast China (figure 2). There were 2920 participants from Shandong province, 2032 participants from Jiangxi province and 1306 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 centre in the campus. The inclusion criteria were: (1) All study participants signed the informed consent form before taking part in the survey; (2) 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:

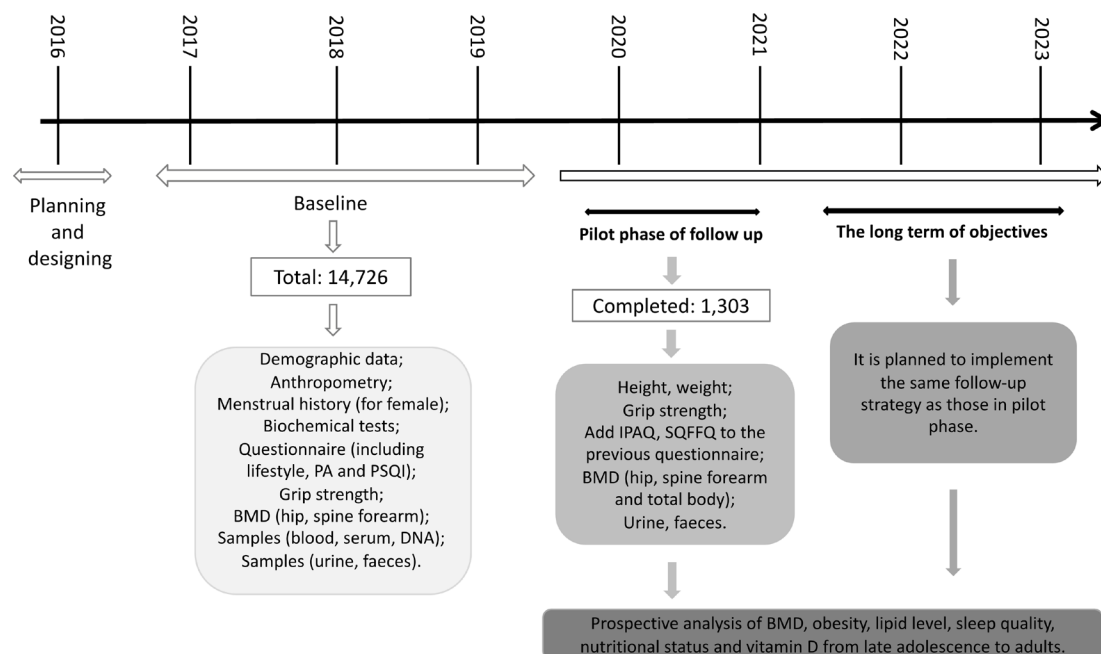


Figure 2 Data collection timeline. IPAQ, International Physical Activity Questionnaire; PA, Physical Activity; PSQI, Pittsburgh Sleep Quality Index; SQFFQ, Semi-Quantitative Food Frequency Questionnaire.

(1) age <14 year old or >25 year old; (2) participants have taken drugs which could affect bone metabolism (eg, glucocorticoids); (3) participants have illness which could cause secondary osteoporosis (eg, 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 (HR), lifestyle and sleep patterns, biological, clinical and health outcomes. Data and samples were collected via examinations, questionnaire and venipuncture (table 2).

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 and 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 haemoglobin (HbA1c)
Bone-related measures	Bone turnover markers (Osteocalcin, PINP, NBAP and β -CTX), serum 1,25(OH) ₂ D ₃ , serum 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 standardised protocols.

†Measured only on WBBC pilot 3.

BAP, bone alkaline phosphatase; β -CTX, β -isomerised C-terminal telopeptides; 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.

Measurements of anthropometric parameters

Anthropometric data included height, body weight, bust, waist, hip and thigh circumference, resting blood pressure, HR 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 HR 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 HR, the participants were asked to take a rest for at least 5 min 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° angle at the elbow²⁹ 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 centre in each college in the morning with at least 8 hours 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 (EDTA K₂) 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, high-density lipoprotein), 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 (2×0.5 mL) for each participant at -80°C 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

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, Waltham, Massachusetts, USA
Blood routine (five 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) ₂ D 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.

measures data (eg, age, sex, ethnicity, family economic status and born place), menstrual history (for female), life-style (eg, physical activity, smoking, alcohol, tea and coffee intake), additional supplement (eg, 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.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 to 21 points.

Bone mineral density assessment

Bone mineral density (BMD) is expressed in terms of bone mass per cm² (g/cm²) and were assessed by dual-energy

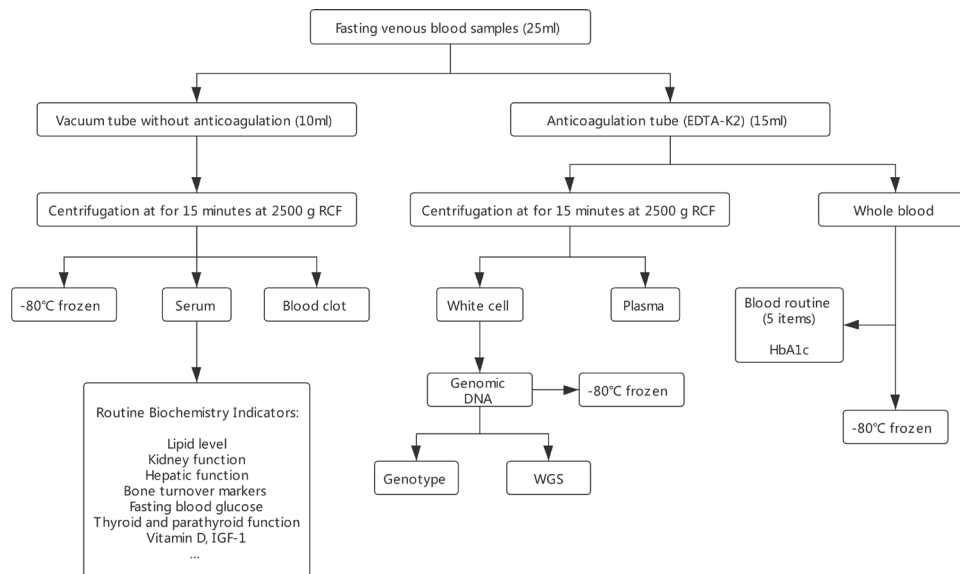


Figure 3 Flow diagram of main blood biochemical detection and blood conservation. RCF, relative centrifugal force; WGS, whole genome sequence.

X-ray absorptiometry³¹ (DXA, Discovery QDR 4500; Hologic, Waltham, Massachusetts, 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 (WGS) and genotyping (figure 3). WGS was completed by NovaSeq 6000 system (Illumina Co.), and for now, 4535 participants have been sequenced at a mean depth of 14×, with highest depth of 65×. A Chinese specific reference panel will be constructed for imputation for the Chinese population. Whole genome genotyping was completed with Infinium Asian Screening Array (Illumina Co., Ltd), and 5841 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 1303 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 χ^2 tests for continuous and categorical variables, respectively. All variables were presented by unadjusted proportions for categorical variables and unadjusted means with SD for continuous variables. The variables demonstrating a p value of less than 0.05 were considered statistically significant. All statistical analyses were analysed using Stata V.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 (5290 participants), WBBC pilot 2 (6258 participants) and WBBC pilot 3 (3178 participants).

We have several ongoing projects under WBBC pilot study. One of the most significant ongoing projects is the WGS study for Chinese population, the results were reported elsewhere.³² In brief, 4535 WGS individuals and 5481 high-density genotyping individuals were available in WBBC pilot project, covering 30 provincial regions of China. Since our previous study³³ demonstrated that the existing reference panels, such as the 1000 Genome Phase3 panel³⁴ 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

Table 4 Basic characteristics of participants in baseline of Westlake BioBank for Chinese pilot

Variables(unit)		Total	Sex		*P value
			Male	Female	
Sociodemographic					
Age (years)	N†	14 726	4751	9975	0.001
	M (SD)‡	18.5 (1.3)	18.6 (1.2)	18.5 (1.3)	
Gender	N (%)§	14 726 (100)	4751 (32.3)	9975 (67.7)	<0.001
Ethnicity					
					0.305
Han	N (%)	11 136 (97.8)	3642 (98.0)	7494 (97.7)	
Others	N (%)	253 (2.2)	75 (2.0)	178 (2.3)	
Hukou status					
					<0.001
Rural	N (%)	5951 (66.6)	1739 (60.2)	4212 (69.6)	
Urban	N (%)	2986 (33.4)	1148 (39.8)	1838 (30.4)	
Anthropometry					
Height (cm)	N	14277	4588	9689	<0.001
	M (SD)	164.2 (8.5)	172.9 (6.6)	160.1 (5.8)	
Weight (kg)	N	14279	4587	9692	<0.001
	M (SD)	57.01 (12.32)	65.81 (13.60)	52.85 (9.07)	
Waist (cm)	N	12 396	3905	8491	<0.001
	M (SD)	71.7 (9.8)	75.5 (11.1)	69.9 (8.7)	
Hip (cm)	N	12388	3902	8486	<0.001
	M (SD)	89.8 (7.3)	90.8 (8.3)	89.3 (6.7)	
Thigh (cm)	N	12351	3880	8471	0.102
	M (SD)	51.7 (5.7)	51.6 (6.5)	51.7 (5.3)	
Cardiovascular system					
SBP (mm Hg)	N	14277	4595	9682	<0.001
	M (SD)	113 (12)	121 (12)	110 (11)	
DBP (mm Hg)	N	14276	4595	9681	<0.001
	M (SD)	71 (9)	73 (9)	70 (8)	
Heart rate (beats/min)	N	14295	4599	9696	<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 (%)	2844 (38.8)	1333 (62.0)	1511 (29.2)	<0.001
Sleeping time (hours)	N	7247	2115	5132	<0.001
	M (SD)	8.2 (1.4)	8.0 (1.3)	8.3 (1.4)	
Grip strength (kg)					
Left hand	N	8932	2958	5974	<0.001
	M (SD)	27.38 (8.77)	36.66 (7.53)	22.79 (4.82)	
Right hand	N	8941	2967	5974	<0.001
	M (SD)	29.76 (9.66)	39.90 (8.32)	24.72 (5.35)	
BMD (g/cm²)					
Lumar spine	N	10154	3293	6861	<0.001
	M (SD)	0.910 (0.105)	0.926 (0.112)	0.903 (0.100)	
Total hip	N	10160	3296	6864	<0.001
	M (SD)	0.868 (0.127)	0.932 (0.139)	0.837 (0.108)	
Femoral neck	N	10160	3296	6864	<0.001
	M (SD)	0.778 (0.125)	0.846 (0.138)	0.746 (0.104)	

Continued

Table 4 Continued

Variables(unit)		Total	Sex		*P value
			Male	Female	
Forearm	N	9917	3238	6679	<0.001
	M (SD)	0.657 (0.059)	0.705 (0.057)	0.634 (0.453)	

*T-test and χ^2 tests for continuous and categorical variables, respectively, to refer the significant differences between males and females.

†N: sample size.

‡M (SD): mean (SD).

§N (%): sample size (percentage).

DBP, diastolic blood pressure; SBP, systolic blood pressure.

an online imputation server (<https://imputationserver.westlake.edu.cn/>) which could result in higher imputation accuracy compared with the existing panels, especially for lower frequency variants.³²

Given the extensive range of data collected in the WBBC study (<https://wbcc.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 sociodemographic, 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 75.5 cm, 90.8 cm and 51.6 cm in males, and 71.7 cm, 89.8 cm and 51.7 among females. The mean systolic blood pressure, diastolic blood pressure and HR in participants were 113 mm Hg, 71 mm Hg and 86 beats/min, 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 27.38 kg and grip-right: 39.90 vs 29.76 kg, both $p<0.001$).

Height and weight were measured using the standardised procedures. 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 (WGO),³⁶ participants were defined as underweight ($<18.5 \text{ kg}/\text{m}^2$), normal weight ($18.5\text{--}23.9 \text{ kg}/\text{m}^2$), overweight ($24\text{--}27.9 \text{ kg}/\text{m}^2$) and obese ($\geq 28 \text{ kg}/\text{m}^2$). 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

Table 5 Distribution of body mass index or waist by sex in participants in Westlake BioBank for Chinese pilot, 2017–2019 (N=14264)

Variables	Total	Male	Female	P value*
Underweight, N (%)	3460 (24.3%)	903 (19.7%)	2557 (26.4%)	<0.0001
Normal, N (%)	8446 (59.2%)	2507 (54.7%)	5939 (61.4%)	<0.0001
Overweight, N (%)	1640 (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†, N (%)	1812 (14.6%)	746 (19.1%)	1066 (12.6%)	<0.0001

* χ^2 tests was calculated between male and female. Participants were defined as underweight ($<18.5 \text{ kg}/\text{m}^2$), normal weight ($18.5\text{--}23.9 \text{ kg}/\text{m}^2$), overweight ($24\text{--}27.9 \text{ kg}/\text{m}^2$) and obese ($\geq 28 \text{ kg}/\text{m}^2$). Central obesity was defined as WC ≥ 85 cm for males and as WC ≥ 80 cm for females.

†Total sample size for central obesity was 12396.

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 2023 young female participants (70.5% subjects aged 20–25 years) from eight Chinese universities⁴⁰ 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, waist-to-height ratios, a body shape index⁴² and body roundness index⁴³ 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 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 recognised, 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 WHO exponentially suppressed vitamin D synthesis in the skin.⁴⁶

COLLABORATION

Participants have agreed to provide their pseudonymised data being made available to other approved researchers. The WBBC pilot study welcomes and offers

Table 6 Distribution of vitamin D level by sex in subjects in Westlake BioBank for Chinese pilot, 2017–2018 (N=11 370)

Variables	Total	Male	Female	P value
Serum 1,25(OH)D, mean (SD)	22.4 (5.3)	25.2 (5.4)	21.1 (4.7)	<0.0001*
Vitamin D deficient, N (%)	3806 (33.5%)	610 (16.4%)	3196 (41.8%)	<0.0001†
Vitamin D insufficient, N (%)	6614 (58.2%)	2452 (65.7%)	4162 (54.5%)	<0.0001†
Vitamin D sufficient, N (%)	950 (8.4%)	668 (17.9%)	282 (3.7%)	<0.0001†
Vitamin D supplementation‡, N (%)	164 (2.2%)	67 (3.1%)	97 (1.9%)	0.00098†

*T-test was calculated between male and female.

† χ^2 tests was calculated between male and female.

‡The sample size for vitamin D supplementation was 7326.

**Table 7** Participants' general react about activity status in Westlake BioBank for Chinese pilot, 2017–2018 (N=7326*)

	Total	Male	Female
What would you like to do in spare time?			
Stay indoors	6364 (86.9%)	1754 (81.6%)	4610 (89.1%)
Take part in some activities	962 (13.1%)	395 (18.4%)	567 (11.0%)
Do you do exercise initiatively?			
Yes	4296 (58.6%)	1517 (70.6%)	2779 (53.7%)
No	3030 (41.4%)	632 (29.4%)	2398 (46.3%)
How often do you have outdoor activities every week?			
Hardly	2785 (38.0%)	477 (22.2%)	2308 (44.6%)
Occasionally	3865 (52.8%)	1301 (60.5%)	2564 (49.5%)
Often	676 (9.2%)	371 (17.3%)	305 (5.9%)

*Data are showed as n (%) of participants.

global collaboration. The data are 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.

Author affiliations

¹Diseases & Population (DaP) Geninfo Lab, School of Lifesciences, Westlake University, Hangzhou, Zhejiang, China

²Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, Zhejiang, China

³Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Westlake University, Hangzhou, Zhejiang, China

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

⁵WBBC Shandong Center, Binzhou Medical University, Yantai, Shandong, China

⁶Hangzhou Kingmed Diagnostics Co., Ltd, Hangzhou, Zhejiang, China

⁷Zhejiang Healthfuture Institute for Cell-Based Applied Technology, Hangzhou, Zhejiang, China

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

⁹Department of Health Management, Hangzhou Normal University, Hangzhou, Zhejiang, China

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 AvantaBioBioSciences Co., Ltd for their support and assistance in sample storage and management.

Contributors H-FZ gained funds and conceived of the study, and K-QL, P-YW, J-YC, X-JX, J-JX, M-CQ, CL and S-YX were involved in the sample collection. J-QL and YS were involved in the blood biochemistry testing. W-YB, P-PZ, JX, S-RG, P-LG, YQ and P-KC were involved in the DNA extraction. X-WZ and H-FZ analysed the data and wrote the paper. All authors read and approved the final manuscript.

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

Map disclaimer The depiction of boundaries on the map(s) in this article does not imply the expression of any opinion whatsoever on the part of BMJ (or any member of its group) concerning the legal status of any country, territory, jurisdiction or area or of its authorities. The map(s) are provided without any warranty of any kind, either express or implied.

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

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not required.

Ethics approval 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.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. The data are 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.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Hou-Feng Zheng <http://orcid.org/0000-0001-5681-8598>

REFERENCES

- 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 Pediatr* 2012;12:143.
- Patton GC, Coffey C, Carlin JB, *et al.* Overweight and obesity between adolescence and young adulthood: a 10-year prospective cohort study. *J Adolesc Health* 2011;48:275–80.
- Sawyer SM, Afifi RA, Bearinger LH, *et al.* Adolescence: a foundation for future health. *Lancet* 2012;379:1630–40.
- 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:e51949.
- 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;384:766–81.
- Jiang Y, Wang J, Wu S, *et al.* Association between Take-Out food consumption and obesity among Chinese university students: a cross-sectional study. *Int J Environ Res Public Health* 2019;16. doi:10.3390/ijerph16061071. [Epub ahead of print: 25 03 2019].
- 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;18:571.
- 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. *Am J Clin Nutr* 2007;86:33–40.
- 9 Llewellyn A, Simmonds M, Owen CG, *et al.* Childhood obesity as a predictor of morbidity in adulthood: a systematic review and meta-analysis. *Obes Rev* 2016;17:56–67.
 - 10 Jeon K, Kim D-I. The association between low body weight and scoliosis among Korean elementary school students. *Int J Environ Res Public Health* 2018;15. doi:10.3390/ijerph15122613. [Epub ahead of print: 22 11 2018].
 - 11 He Q, Karlberg J. Bmi in childhood and its association with height gain, timing of puberty, and final height. *Pediatr Res* 2001;49:244–51.
 - 12 Hudson LD, Nicholls DE, Lynn RM, *et al.* Medical instability and growth of children and adolescents with early onset eating disorders. *Arch Dis Child* 2012;97:779–84.
 - 13 Antonucci R, Locci C, Clemente MG, *et al.* Vitamin D deficiency in childhood: old lessons and current challenges. *J Pediatr Endocrinol Metab* 2018;31:247–60.
 - 14 Trehan N, Afonso L, Levine DL, *et al.* Vitamin D deficiency, supplementation, and cardiovascular health. *Crit Pathw Cardiol* 2017;16:109–18.
 - 15 Watkins RR, Lemonovich TL, Salata RA. An update on the association of vitamin D deficiency with common infectious diseases. *Can J Physiol Pharmacol* 2015;93:363–8.
 - 16 Illescas-Montes R, Melguizo-Rodríguez L, Ruiz C, *et al.* Vitamin D and autoimmune diseases. *Life Sci* 2019;233:116744.
 - 17 Holick MF. Vitamin D deficiency. *N Engl J Med* 2007;357:266–81.
 - 18 Holick MF, Binkley NC, Bischoff-Ferrari HA, *et al.* Evaluation, treatment, and prevention of vitamin D deficiency: an endocrine Society clinical practice guideline. *J Clin Endocrinol Metab* 2011;96:1911–30.
 - 19 Pludowski P, Grant WB, Bhattoa HP, *et al.* Vitamin D status in central Europe. *Int J Endocrinol* 2014;2014:589587.
 - 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. *Matern Child Nutr* 2009;5:186–91.
 - 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. *Osteoporos Int* 2013;24:2775–87.
 - 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. *J Nutr Health Aging* 2013;17:587–93.
 - 23 Faghieh S, Abdolazadeh M, Mohammadi M, *et al.* Prevalence of vitamin D deficiency and its related factors among university students in Shiraz, Iran. *Int J Prev Med* 2014;5:796–9.
 - 24 Fraser DR. Vitamin D-deficiency in Asia. *J Steroid Biochem Mol Biol* 2004;89-90:491–5.
 - 25 Yu X, Wang W, Wei Z, *et al.* Vitamin D status and related factors in newborns in Shanghai, China. *Nutrients* 2014;6:5600–10.
 - 26 Wang EW-L, Pang MY-C, Siu PM-F, *et al.* Vitamin D status and cardiometabolic risk factors in young adults in Hong Kong: associations and implications. *Asia Pac J Clin Nutr* 2018;27:231–7.
 - 27 Guo Y, Ke H-J, Liu Y, *et al.* Prevalence of vitamin D insufficiency among children in southern China: a cross-sectional survey. *Medicine* 2018;97:e11030.
 - 28 Yan X, Zhang N, Cheng S, *et al.* Gender differences in vitamin D status in China. *Med Sci Monit* 2019;25:7094–9.
 - 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 Ageing* 2011;40:423–9.
 - 30 Buysse DJ, Reynolds CF, Monk TH, *et al.* The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry Res* 1989;28:193–213.
 - 31 Bazzocchi A, Ponti F, Albisinni U, *et al.* DXA: technical aspects and application. *Eur J Radiol* 2016;85:1481–92.
 - 32 Cong P, Bai W-Y, Li J. Genomic analyses of 10,376 individuals provides comprehensive map of genetic variations, structure and reference haplotypes for Chinese population. *bioRxiv* 2021.
 - 33 Bai W-Y, Zhu X-W, Cong P-K, *et al.* Genotype imputation and reference panel: a systematic evaluation on haplotype size and diversity. *Brief Bioinform* 2019. doi:10.1093/bib/bbz108. [Epub ahead of print: 06 Nov 2019].
 - 34 , Auton A, Brooks LD, *et al.* 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* 2015;526:68–74.
 - 35 McCarthy S, Das S, Kretzschmar W, *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* 2016;48:1279–83.
 - 36 Zhou B-F, Cooperative Meta-Analysis Group of the Working Group on Obesity in China. 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. *Biomed Environ Sci* 2002;15:83–96.
 - 37 Asia Pacific Cohort Studies Collaboration. Central obesity and risk of cardiovascular disease in the Asia Pacific region. *Asia Pac J Clin Nutr* 2006;15:287–92.
 - 38 NCD Risk Factor Collaboration (NCD-RisC). 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;390:2627–42.
 - 39 Chen H, Jackson T. Predictors of changes in weight Esteem among mainland Chinese adolescents: a longitudinal analysis. *Dev Psychol* 2009;45:1618–29.
 - 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:e0205282.
 - 41 Wang K, Liang R, Ma Z-L, *et al.* Body image attitude among Chinese college students. *Psych J* 2018;7:31–40.
 - 42 Krakauer NY, Krakauer JC. A new body shape index predicts mortality hazard independently of body mass index. *PLoS One* 2012;7:e39504.
 - 43 Thomas DM, Bredlau C, Bosity-Westphal A, *et al.* Relationships between body roundness with body fat and visceral adipose tissue emerging from a new geometrical model. *Obesity* 2013;21:2264–71.
 - 44 Mithal A, Wahl DA, Bonjour J-P, *et al.* Global vitamin D status and determinants of hypovitaminosis D. *Osteoporos Int* 2009;20:1807–20.
 - 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 Nutr* 2016;19:78–82.
 - 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. *Br J Dermatol* 2012;167:391–5.