

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

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

<sup>1</sup> Disease & Population (DaP) Geninfo Lab, School of Life Sciences, Westlake  
University, Hangzhou, Zhejiang, 310024, China

<sup>2</sup> Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Westlake  
University, Hangzhou, Zhejiang, 310024, China

<sup>3</sup> School of Life Sciences, Fudan University, Shanghai 200433, China

<sup>4</sup> Jiangxi Medical College, Shangrao, Jiangxi, China

<sup>5</sup> Department of Epidemiology & Department of Biochemistry and Molecular Biology,  
Binzhou Medical University, Yantai, Shandong 264003, China

<sup>6</sup> Hangzhou Kingmed Diagnostics Co., Ltd., Hangzhou, Zhejiang, 310053, China

<sup>7</sup> Research and Development Department, Zhejiang Healthfuture Institute for  
Cell-Based Applied Technology, Hangzhou, Zhejiang, China

<sup>8</sup> Center for Disease Control and Prevention of Jiading District, Shanghai, China

<sup>9</sup> Department of Health Management, Hangzhou Normal University, Hangzhou,  
Zhejiang, 311121, China

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

## 1 **ABSTRACT**

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

8 **Participants** A total of 14,726 participants (4,751 males and 9975 females) aged 14 to  
9 25 years were recruited and the baseline survey was carried out from 2017 to 2019.  
10 The pilot cohort contains rich range of information regarding of demographics,  
11 anthropometric measurements, blood pressure and heart rate, lifestyle and sleep  
12 patterns, biological, clinical and health outcomes.

13 **Findings to date** There were most of the participants were Chinese Han ethnic  
14 (97.78%), and more than 60% of them were originally from rural areas. For  
15 anthropometry measurements, the mean height and weight were 172.92 cm and 65.81  
16 kg for men, and 160.13 cm and 52.85 kg for women; the mean waist, hip and thigh  
17 circumference were 75.51cm, 90.75cm and 51.56 cm in men, and 71.67cm, 89.75cm  
18 and 51.68 among women. Results indicated that the prevalence of underweight in  
19 female was much higher than male, but the prevalence of overweight and obesity in  
20 female was lower than male. The mean of serum 25(OH)D level in the 14,726 young  
21 participants was 22.39 ng/ml, and male had higher level of serum 25(OH)D than  
22 female, overall, 33.47% of the participants had vitamin D deficiency and even more  
23 participants suffered from vitamin D insufficiency (58.17%). The proportion of  
24 deficiency in women was much higher than that in men (41.83 vs. 16.35%).

25 **Future plans** WBBC is designed as a prospective cohort study and provides a unique  
26 and rich data set analyzing health trajectories of Chinese adolescents to young adults.

## 27 **Introduction**

28 The initial goal of setting up this cohort is trying to bring up more attentions to the  
29 skeletal health in adolescence and adults, and thus to provide strategy in prevention of  
30 osteoporosis in the elderly. Osteoporosis, a disease of increased bone fragility, is a  
31 systemic osteopathy characterized by a decrease in bone density/quality and the  
32 destruction of bone microstructure caused by genetic and environmental factors<sup>1</sup>.  
33 Bone mineral density (BMD), the bone mineral content (BMC) in bone tissue, is  
34 recognized as the most important predictor of osteoporosis. Currently, approximately  
35 200 million people worldwide suffer from osteoporosis, and 83.9 million of which are  
36 in China<sup>2</sup>. As the trend of global aging is becoming more and more obvious, more  
37 attention should be paid to the bone health not only in the elderly, but also in  
38 adolescence and young adults. Population-based studies have shown that roughly half  
39 of the boys and one-third of the girls would undergo a fracture by age of 18 and 1/5  
40 would have two or more fractures<sup>3 4</sup>. Epidemiologic studies have shown that a 10%  
41 increase in peak bone mass (PBM) at the population level reduces the risk of fracture  
42 later in life by 50%<sup>5</sup>. In fact, bone mass attained in early life was considered to be the  
43 most important modifiable determinant of lifelong skeletal health. A longitudinal data  
44 have shown that more than 94% of BMD was acquired at the age of sixteen<sup>6</sup> and  
45 approximately 40% to 60% of adult bone mass was accrued during the adolescent  
46 years in both women and men<sup>7</sup>.

47  
48 Any condition interfering with optimal peak bone mass accrual can, therefore,  
49 increase fracture risk later in life, adolescence is critical period for skeletal  
50 mineralization. The bone mass gain during adolescence is influenced by multiple  
51 factors, including genetic factors, ethnicity, and environmental factors, such as alcohol  
52 intake, cigarette smoking, physical activity, endocrine status (e.g. vitamin D), diet  
53 (calcium and protein intake) and other factors<sup>5 8-13</sup>. As environmental and behavioral  
54 factors account for 20% to 40% of adult peak bone mass<sup>14 15</sup>, the early identification  
55 of the factors associated with poor bone health and the provision of reliable

56 counseling might help children and teenagers take action to maximize BMD before  
57 their PBM was completed.

58

59 Overweight and obesity are a major public health problem<sup>16</sup>. Elevated body mass  
60 index (BMI) in adolescence had been associated with several obesity-related  
61 morbidities in adult life, such as diabetes, metabolic syndromes and some types of  
62 cancer<sup>17</sup>. And obesity in adolescence conferred very high risks for obesity in adults<sup>18</sup>;  
63 70% of overweight adolescents had one or more concomitant conditions such as high  
64 blood pressures and fasting insulin, which were also risk factors for cardiovascular  
65 disease, and 23% of those accompanied with three or more concomitant conditions<sup>19</sup>.  
66 Following rapid economic development since the 1980s, China experiences a rapidly  
67 increasing of overweight and obesity among children and adolescents<sup>20 21</sup>. In 2019, a  
68 cross-sectional study<sup>22</sup> found that the prevalence of overweight in college students  
69 (aged 18-26 years) was 8.0%, and the prevalence of obesity was 3.5%. A recent study  
70 from 12 provinces in China showed that the prevalence of overweight and obesity  
71 were 14.0% and 10.5% in boys, and 9.7% and 7.1% in girls, respectively<sup>23</sup>.

72

73 Vitamin D deficiency is becoming a public health problem in both developed and  
74 developing countries<sup>24 25</sup>. Besides its effect on musculoskeletal system, Vitamin D  
75 showed pleiotropic effect on human health, such as cardiovascular diseases<sup>26</sup>,  
76 common infectious diseases<sup>27</sup> and autoimmune diseases<sup>28</sup>. Serum 25(OH)D is a good  
77 indicator of vitamin D storage and is an optimal method of assessing vitamin D  
78 levels<sup>24</sup>. According to the Endocrine Society clinical practice guidelines, vitamin D  
79 levels were defined as deficiency [25(OH)D<20 ng/mL], insufficiency [25(OH)D:  
80 20-29 ng/mL] and sufficiency [25(OH)D≥30 ng/mL] respectively<sup>29</sup>. Many people in  
81 central and western Europe had vitamin D concentration of 11–20 ng/mL in winter<sup>30</sup>.  
82 Studies from other countries, including Canada<sup>31</sup>, Japan<sup>32</sup>, Australia<sup>33</sup> and Iran<sup>34</sup>,  
83 presented similar situations, with high prevalence of vitamin D insufficiency in  
84 different ethnicities. A study in north China found that more than 40% of adolescent  
85 girls had Vitamin D-deficiency in the winter<sup>35</sup>. Another study in Shanghai showed that

86 more than one-third newborns had plasma 25(OH)D less than 20 ng/mL<sup>36</sup>. Even in  
87 Hong Kong (latitude 22° north), 72% of young adults were reported to have vitamin  
88 D deficiency<sup>37</sup>.

89  
90 The overall goal of the Westlake Biobank for Chinese (WBBC) pilot cohort is to  
91 recruit individuals at their late adolescence/young adulthood. The biological samples  
92 such as whole blood, serum, urine and faeces were collected, genomic DNA were  
93 extracted and the DNA sequence information were acquired through sequencing  
94 technique. A long questionnaire with questions concerning the environmental factors  
95 such as nutrition, sleep quality, physical activity, medication etc was provided. These  
96 data will help us to understand the association between the genetics, environmental  
97 factors, microbiome and health statue of adolescence population. With a broad range  
98 of phenotype collection on many aspects of participants' daily life, a wide range of  
99 scientific questions could be addressed. The main purpose of this particular paper is to  
100 profile the cohort, therefore, only limited findings were described, such as: (1) what  
101 are the prevalence of underweight, overweight, obesity and vitamin D deficiency in  
102 Chinese late adolescence? What is the reference value of serum vitamin D level in the  
103 young people? (2) what is the difference between male and female in term of height,  
104 weight, blood pressure, lifestyle and bone health in the young people.

105

## 106 **Cohort description**

### 107 **Sampling Design**

108 The Westlake Biobank for Chinese (WBBC) pilot study was collected in three main  
109 regions in China (Hangzhou city of Zhejiang province, Shangrao city of Jiangxi  
110 province and Yantai city of Shandong province), but the participants covered all  
111 around the country (Table 1 and Figure 1). The baseline survey was carried out from  
112 2017 to 2019. The target population was young people aged 14–25 years who were  
113 college students and available for follow-up studies. In the first phase of baseline  
114 (WBBC pilot 1), the participants were recruited from two colleges at Zhejiang

115 province and Jiangxi province in Southeast China from September 2017 to March  
116 2018 (Figure 2), and 1,258 and 2,769 participants were from Zhejiang and Jiangxi  
117 provinces, respectively, and 1,263 participants were from other 25 provinces of China  
118 (Table 1). From September 2018 to December 2018, the second phase of WBBC pilot  
119 project was initiated (WBBC pilot 2), the participants were recruited at the same  
120 college in Jiangxi province and a college in Shandong province in Northeast China  
121 (Figure 2). There were 2,920 participants from Shandong province, 2,032 participants  
122 from Jiangxi province and 1,306 participants from other 28 provinces of China in  
123 WBBC pilot 2 (Table 1). From September 2019, the WBBC pilot project phase 3  
124 (WBBC pilot 3) recruited participants from the same college in Jiangxi province  
125 (Figure 2), most of the participants (2,504) of WBBC pilot 3 were from Jiangxi  
126 province and 6,74 participants were from other 26 provinces of China (Table 1). All  
127 participants provided their Chinese unique national identity (ID) number for unique  
128 reference at the health examination center in the campus. The inclusive criteria were:  
129 a). All study participants signed the informed consent form before taking part in the  
130 survey. b). Participants should complete the physical examination, bone mineral  
131 density scan, blood test and questionnaires. And the exclusion criteria were: a). No  
132 informed consent. b). Participants missed most of the items in data collection. In  
133 WBBC pilot 3, the urine and feces of the participants were collected, therefore,  
134 participants taking antibiotics should be excluded.

135

### 136 **Data Collection Procedures**

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

142

#### 143 *Measurements of anthropometric parameters*

144 Anthropometric data included height, body weight, bust, waist, hip and thigh

145 circumference, resting blood pressure, heart rate and hand grip. Height was measured  
146 to the nearest 0.1 cm with participants' light-weight clothes and shoes off; weight was  
147 measured to the nearest 0.01 kg with the weight scale (Ultrasonic surveying  
148 instrument, Beryl BYH01, China) calibrated daily before each series of measurements.  
149 Bust, waist, hip and thigh circumference were measured to the nearest 0.5 cm by  
150 using a measuring tape with the subject standing comfortably. Resting blood pressure  
151 and heart rate were measured on the left arm supported at heart-level sitting position  
152 using electronic sphygmomanometers (Yuwell YE660A, China). To ensure accurate  
153 data, the participants were asked to have rested for at least 5 minutes and have no  
154 excessive physical activity, tea or alcohol intake or smoking for at least one night.  
155 Using a handgrip dynamometer (CAMRY EH101, China), grip strength with both  
156 hands were tested for most of the participants in WBBC pilot 2 and WBBC pilot 3. In  
157 order to get more accurate results, the participants should make sure the arm that's  
158 being tested was at a 90 degree angle at the elbow<sup>38</sup> until the test was finished. Details  
159 of the methods and instruments used for measurements of anthropometric parameters  
160 were provided in Table 3.

161

### 162 *Biochemistry assessment*

163 Participants came to the examination center in each college in the morning with at  
164 least 8 h of overnight fasting, about 25 ml of venous blood samples were collected for  
165 routine blood measurements, biochemical indexes, DNA extraction and so on. Venous  
166 blood samples were collected using ethylenediamine tetraacetic acid dipotassium  
167 (EDTAK2) anticoagulation tube (3×5.0 ml) and vacuum tube without anticoagulation  
168 (2×5.0 ml). Serum and plasma samples were separated from whole blood through  
169 centrifugation for 10 min at the relative centrifugal force 3000 g (Figure 3). Serum  
170 samples were forwarded to test biochemical indexes that included serum 25(OH)D  
171 level, serum calcium level, fasting blood glucose, kidney function test, hepatic  
172 function test, blood lipids [triglyceride, cholesterol, low-density lipoprotein (LDL),  
173 high-density lipoprotein (HDL)], triiodothyronine (T3), thyroid and parathyroid  
174 function and bone turnover markers (Table 2). Details of the platforms used for

175 biochemical analysis are provided in Table 3. We also reserved serum samples  
176 (2×0.5ml) for each participant at - 80□ for future use. Figure 3 displays the detail of  
177 the flow diagram of blood separation and detection of main blood biochemical  
178 indexes.

179

#### 180 *Questionnaire-based assessments*

181 Baseline data collection for participants included a self-completion questionnaire. In  
182 Table 2, a list of core questions within the aforementioned domains, was provided.  
183 The questionnaire included social and demographic measures data (e.g. age, sex,  
184 ethnicity, family economic status and born place), menstrual history (for female),  
185 lifestyle (e.g. physical activity, smoking status, alcohol, tea and coffee intake),  
186 additional supplement (e.g. calcium and vitamin D), health status and other  
187 information. Sleep duration and sleep quality were assessed by the Pittsburgh Sleep  
188 Quality Index (PSQI)<sup>39</sup>. This is composed of 19 questions which reflect seven major  
189 components, all seven components are then summed up to create a scale from 0–21  
190 points.

191

#### 192 *Bone mineral density assessment*

193 Bone mineral density (BMD) is expressed in terms of bone mass per cm<sup>2</sup> (g/cm<sup>2</sup>) and  
194 were assessed by dual-energy X-ray absorptiometry<sup>40</sup> (DXA, Discovery QDR 4500;  
195 Hologic Inc., Waltham, MA, USA). In WBBC pilot, the main sites for BMD  
196 evaluation were lumbar spine (L1–L4), femur (femoral neck and total hip) and distal  
197 third of the radius.

198

#### 199 *Whole Genome sequencing and genotyping*

200 In the process of collecting fasting venous blood, 5-ml whole blood was used for  
201 isolation of genomic DNA. These genomic DNA were used for whole genome  
202 sequencing and genotyping (Figure 3). Whole genome sequencing was completed by  
203 NovaSeq 6000 system (Illumina Co., Ltd), and for now, 1,192 participants have been  
204 sequenced at mean depth of 14x, with highest depth of 65x. A Chinese specific



205 reference panel will be constructed for imputation for Chinese population. Whole  
206 genome genotyping was completed with Infinium Asian Screening Array (ASA)  
207 (Illumina Co., Ltd), and 5,841 participants have been genotyped in approximately  
208 700,000 SNPs.

209

#### 210 *Follow-up and outcome measures*

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

220

221 We have started a pilot follow-up study for WBBC pilot 2 since December 2019, and  
222 1,303 participants had completed all examinations (Figure 2). The collected  
223 information included height, weight, grip strength and the updated questionnaire.  
224 Besides, we retested bone mineral density at spine (L1-L4), hip and distal third of the  
225 radius.

226

227

#### 228 **Statistical analysis**

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

235

## 236 **Findings to date**

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

243

244 We have several ongoing projects under WBBC pilot study. One of the most  
245 significant ongoing projects is the study of Chinese population structure. In WBBC  
246 pilot study, the data from 1,192 samples with whole genome sequencing and 5,841  
247 samples with whole genome genotyping were available, collaborating with another  
248 research group in China, we finally achieved approximate ~10,000 Chinese samples  
249 with whole genome data, covering 30 provincial region of China. In addition, WBBC  
250 pilot study will provide a haplotype reference panel to improve the imputation  
251 accuracy of Chinese GWAS study, since our previous study<sup>41</sup> demonstrated that the  
252 existing reference panels, such as the 1000 Genome Phase3 panel<sup>42</sup> and the HRC  
253 (Haplotype Reference Consortium) panel<sup>43</sup>, were not best fit for imputation for  
254 Chinese population, especially for the rare variant imputation.

255

256 Given the extensive range of data collection in the WBBC study, it is not feasible to  
257 present all the results, only limited findings were described in the present study. In  
258 summary, a total of 17,407 college students were invited, of whom, 14,983 (86.07 %)   
259 responded. After removing participants with missing data and invalid data, the final  
260 study included an effective sample size of 14,726 (84.60%) adolescents and young  
261 adults (with age from 14 to 25 years, and mean age at 18.5 years). Table 4 provides an  
262 overview of socio-demographic, anthropometry, cardiovascular system, lifestyle, grip  
263 strength and BMD characteristics of the WBBC pilot participants at baseline. Briefly,

264 within the 14,726 sample, there were more females than males (67.74 vs. 32.26%),  
265 with mean age of 18.48 years for women and 18.55 years for men, respectively. Most  
266 of the participants were Chinese Han ethnic (97.78), and more than 60% of them were  
267 originally from rural areas (60.24% of males and 69.62% of females). For  
268 anthropometry measurements, the mean height and weight were 172.92 cm and 65.81  
269 kg for men, and 160.13 cm and 52.85 kg for women; the mean waist, hip and thigh  
270 circumference were 75.51cm, 90.75cm and 51.56 cm in men, and 71.67cm, 89.75cm  
271 and 51.68 among women. The mean systolic blood pressure (SBP), diastolic blood  
272 pressure (DBP) and heart rate (HR) in participants were 113 mmHg, 71 mmHg and 86  
273 beats/minute, respectively. In the cohort, only 5.49% of the participants were current  
274 smokers and 38.82% of them were regular drinkers. Regarding the current smoking  
275 status, there was significant difference between men and women (16.29 vs. 1.64%,  $p$   
276  $< 0.001$ ). As for alcohol consumption, the proportion of current drinker in boys and  
277 girls were 62.03% and 29.19%, respectively, which is much higher in males ( $p$   
278  $< 0.001$ ). The mean sleeping time estimated in women were higher than men (8.34 vs.  
279 7.97 hours,  $p < 0.001$ ). As for grip strength, the data collection was started from  
280 WBBC pilot 2, the mean of grip strength in boys were much higher than girls  
281 (grip-left: 36.66 vs. 27.38kg and grip-right: 39.90 vs. 29.76 kg, both  $p < 0.001$ ).

282

283 Height and weight were measured using the standardized procedures. Body mass  
284 index (BMI) was calculated based on the formula: weight in kilograms divided by  
285 height in meters-squared ( $\text{kg}/\text{m}^2$ ). According to the Working Group on Obesity in  
286 China (WGOC)<sup>44</sup>, participants were defined as underweight ( $< 18.5 \text{ kg}/\text{m}^2$ ), normal  
287 weight ( $18.5\text{-}23.9 \text{ kg}/\text{m}^2$ ), overweight ( $24\text{-}27.9 \text{ kg}/\text{m}^2$ ) and obese ( $\geq 28 \text{ kg}/\text{m}^2$ ).  
288 Therefore, the WBBC pilot study provided an overall prevalence of underweight,  
289 overweight and obesity among young participants of 24.26%, 11.50% and 5.03%,  
290 respectively (Table 5). The prevalence of underweight in female was much higher  
291 than male (26.41% vs. 19.70%,  $p < 0.0001$ ), but the prevalence of overweight in  
292 female was much lower than male (9.03% vs. 16.71%,  $p < 0.0001$ ) (Table 5), similarly,  
293 the prevalence of obesity in female (3.21%) was lower than in male (8.88%) ( $p <$

294 0.0001) (Table 5). Waist circumference (WC) is good indicator of abdominal visceral  
295 fat distribution and is a strong predictor of diabetes mellitus and cardiovascular  
296 disease<sup>45</sup>. It is meaningful to investigate the WC along with BMI among adolescents  
297 and young people. In WBBC pilot study, central obesity was defined as  $WC \geq 85$  cm  
298 for men and as  $WC \geq 80$  cm for women based on the recommendations of the WGOC  
299<sup>44</sup>. In the cohort of 12,396 participants, the prevalence of central obesity was 14.62%,  
300 which was higher in male than in female (19.10% vs. 12.55%,  $p < 0.0001$ ) (Table 5).

301

302 In WBBC pilot study, the prevalence of underweight were high in both male (19.70 %)  
303 and female (26.41%), though the prevalence of moderate and severe underweight  
304 decreased from 9.2% in 1975 to 8.4% in 2016 in girls and from 14.8% in 1975 to 12.4%  
305 in 2016 in boys in the world<sup>46</sup>. This phenomenon may due to the modern aesthetics of  
306 human stature that thinness is preferred, especially in China<sup>47 48</sup>. Recently, a study  
307 involving 2,023 young female participants (70.5% subjects aged 20-25 years) from  
308 eight Chinese universities<sup>48</sup> showed that 30.55% of the participants were underweight,  
309 and 57.39% of them would like to be much thinner, which would lead to more  
310 underweight individuals. Therefore, future studies should not only pay attention to the  
311 problem of obesity/overweight, but also to the underweight issue in young people.

312

313 Using simple anthropometric indices of body composition, such as BMI and WC, has  
314 been considered as a practical and valuable approach to the assessment of obesity for  
315 a long time. Waist-to-hip ratios (WHR), waist-to-height ratios (WHeR), a body shape  
316 index (ABSI)<sup>49</sup> and body roundness index (BRI)<sup>50</sup> were also as parameters of body fat  
317 and visceral adipose tissue volume. In WBBC pilot cohort, we had collected several  
318 anthropometric measures including height, weight, bust, waist, hip and thigh  
319 circumference and these data could help us examine the usefulness of these  
320 anthropometric parameters and identify the optimal cut-off of the parameters to  
321 evaluate overweight and obesity among adolescence and young people in future study.  
322 It is noteworthy that vitamin D deficiency in females was significantly worse than in  
323 males. This may due to the modern aesthetics of Chinese culture that paler skin is

324 preferred, especially in females. A questionnaire related to vitamin D and sun  
325 exposure was conducted at a university in Nanjing, China and found that 75.0 % of  
326 the students lacked sun exposure because they would like to avoid dark skin, and most  
327 of students (82.7 %) used sun protection, and sunscreen use were more popular in  
328 females<sup>51</sup>, but it was reported that using the amount of sun cream recommended by  
329 World Health Organization exponentially suppressed vitamin D synthesis in the skin<sup>52</sup>.  
330 In WBBC pilot study, the mean serum 25(OH)D level was  $22.39 \pm 5.33$  ng/ml for all  
331 the participants (male: 25.15 ng/mL and female: 21.05 ng/mL,  $p < 0.0001$ ) (Table 6).  
332 Overall, 33.47% of the participants had vitamin D deficiency and even more  
333 participants suffered from vitamin D insufficiency (58.17%) (Table 6). In addition, the  
334 proportion of women with sufficient vitamin D was much lower than that of men  
335 (3.69 vs. 17.91%,  $p < 0.0001$ ), while the proportion of deficiency in women was  
336 much higher than that in men (41.83 vs. 16.35%,  $p < 0.0001$ ) (Table 6). Most of the  
337 participants (86.87%) preferred to stay indoors in spare time, the females were less  
338 willing to do exercise than males (53.68% vs 70.59%) (Table 7), and 44.58% of  
339 females hardly had outdoor activities, only 5.89% of females often had outdoor  
340 activities every week (Table 7). These results jointly suggested that the females had  
341 not enough sun exposure. Although food sources of vitamin D were not commonly  
342 recognized, only 10–20% of vitamin D in human bodies was obtained through food  
343 sources<sup>53</sup>. In WBBC pilot study, there was only 2.24% of the participants used  
344 vitamin D supplements (3.12% in male and 1.87% in female,  $p = 0.00098$ ) (Table 6)  
345 and this was consistent with Zhou et al<sup>51</sup>, which found that only 5.6 % of the students  
346 used vitamin D supplements in a university of Nanjing, China.

347

## 348 **Strengths and limitations**

349 The WBBC pilot study was designed to develop a longitudinal database on health-  
350 related factors for young people. Firstly, this was a comprehensive cohort study to pay  
351 attention to the health of adolescence and young adults, including over ~14,000  
352 participants (Table 1). Secondly, the WBBC pilot cohort is rich in longitudinal

353 phenotypic data and archived biospecimens, including whole blood and serum as well  
354 as DNA samples. These resources will facilitate to digitize the information of the  
355 genomic, proteomic and metabolomic and microbiome of the participants, and further  
356 investigate the association of the information with the development of adolescence  
357 and young adults. Thirdly, this cohort has now accrued sufficient events for reliable  
358 analyses, and the major findings on adiposity, BMD, sleep quality, nutrition and  
359 growth for adolescence are expected over the next few years.

360

361 There are several limitations to this cohort design. First, although the participants  
362 covered all around the country, most of them were mainly from 3 provinces (Jiangxi,  
363 Shandong and Zhejiang), and the selection of participants was not random, but in a  
364 specific population (college students). Second, the participants were collected in 3  
365 phases at different time points, and the phenotypic information was updated in the  
366 later phases, for example, the items in the questionnaire of different phases were not  
367 always the same, therefore, the phenotypic data were not always uniform in 3 phases.

368

## 369 **Collaboration**

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

376

## 377 **Funding**

378 This study was supported by the National Natural Science Foundation of China  
379 (81871831), and by the Zhejiang Provincial Natural Science Foundation for  
380 Distinguished Young Scholars of China (LR17H070001), and by the Westlake  
381 Biobank for Chinese (WBBC) funds from Westlake University.

382

### 383 **Acknowledgements**

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

390

### 391 **Contributors**

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

396

### 397 **Conflict of interest**

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

400

### 401 **Patient and public involvement**

402 No patients were involved in the described linkage between existing registries  
403 providing an anonymous data set.

404

### 405 **Ethics Approval and Consent to Participate**

406 The study protocol and informed consent procedure were approved by the Ethics  
407 Committees at Westlake University. All study participants signed the informed  
408 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. *Frontiers of medicine* 2020 doi: 10.1007/s11684-020-0748-y
14. 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
15. 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
16. 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
  17. 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
  18. 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
  19. 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
  20. 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
  21. 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
  22. 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
  23. 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
  24. Holick MF. Vitamin D deficiency. *The New England journal of medicine* 2007;357(3):266-81. doi: 10.1056/NEJMr070553
  25. 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
  26. 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.000000000000122
  27. 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
  28. 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
  29. 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
  30. 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
  31. 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
32. 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
  33. 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
  34. Faghih S, Abdolazadeh 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.
  35. 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
  36. 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
  37. 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
  38. 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
  39. 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
  40. 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
  41. 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
  42. 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
  43. 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
  44. 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.
  45. 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.
  46. 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

47. 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
48. 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
49. 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
50. 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
51. 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
52. Faurischou 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
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

## 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 <sup>#</sup>	International Physical Activity Questionnaire, IPAQ
Dietary pattern <sup>#</sup>	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 $\beta$ -CTX), serum 1,25(OH) <sub>2</sub> D <sub>3</sub> , serum IGF-1
Mineral elements	Calcium (Ca), phosphorus (P) and magnesium (Mg)
Samples	
Blood	25 ml
Serum	2 ml
Urine <sup>#</sup>	10 ml
Faeces <sup>#</sup>	15 g

\* Direct measures carried out by trained interviewers according to standardized protocols; <sup>#</sup> 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;  $\beta$ -CTx:  $\beta$ -isomerized C-terminal telopeptides

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

<b>Variables</b>	<b>Analysis method/platform used</b>
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)	SYSTEMEX 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) <sub>2</sub> D <sub>3</sub>	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	
<b>Socio-demographic</b>					
Age (years)	N <sup>a</sup>	14,726	4,751	9,975	0.001
	M (SD) <sup>b</sup>	18.50 (1.25)	18.55 (1.21)	18.48 (1.26)	
Gender	N (%) <sup>c</sup>	14,726 (100)	4,751 (32.26)	9,975 (67.74)	<0.001
Ethnicity					0.305
<i>Han</i>	N (%)	11,136 (97.78)	3,642 (97.98)	7,494 (97.68)	
<i>Others</i>	N (%)	253 (2.22)	75 (2.12)	178 (2.32)	
Hukou status					<0.001
<i>Rural</i>	N (%)	5,951 (66.59)	1,739 (60.24)	4,212 (69.62)	
<i>Urban</i>	N (%)	2,986(33.41)	1,148 (39.76)	1,838 (30.38)	
<b>Anthropometry</b>					
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)	
<b>Cardiovascular system</b>					
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)	
<b>Lifestyle</b>					
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
<b>Grip strength (kg)</b>					

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)	
<b>BMD (g/cm<sup>2</sup>)</b>					
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)	

<sup>a</sup> N: sample size; <sup>b</sup>M (SD): mean (standard deviation); <sup>c</sup>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 <sup>δ</sup> , N (%)	1,812 (14.62%)	746 (19.10%)	1,066 (12.55%)	<0.0001

<sup>δ</sup> Total sample size for Central Obesity was 12,396. <sup>#</sup> 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 <sup>†</sup>
Vitamin D deficient, N (%)	3,806 (33.47%)	610 (16.35%)	3,196 (41.83%)	<0.0001 <sup>#</sup>
Vitamin D insufficient, N (%)	6,614 (58.17%)	2,452 (65.74%)	4,162 (54.48%)	<0.0001 <sup>#</sup>
Vitamin D sufficient, N (%)	950 (8.36%)	668 (17.91%)	282 (3.69%)	<0.0001 <sup>#</sup>
Vitamin D supplementation <sup>§</sup> , N (%)	164 (2.24%)	67 (3.12%)	97 (1.87%)	0.00098 <sup>#</sup>

<sup>§</sup> The sample size for Vitamin D supplementation was 7,326. <sup>#</sup> Chi-square tests was calculated between male and female. <sup>†</sup>T-test was calculated between male and female.

Table 7. Participants' general react about activity status in WBBC pilot, 2017-2018. N=7,326.<sup>a</sup>

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

<sup>a</sup> Data are showed as n (%) of participants.

## Figure legend

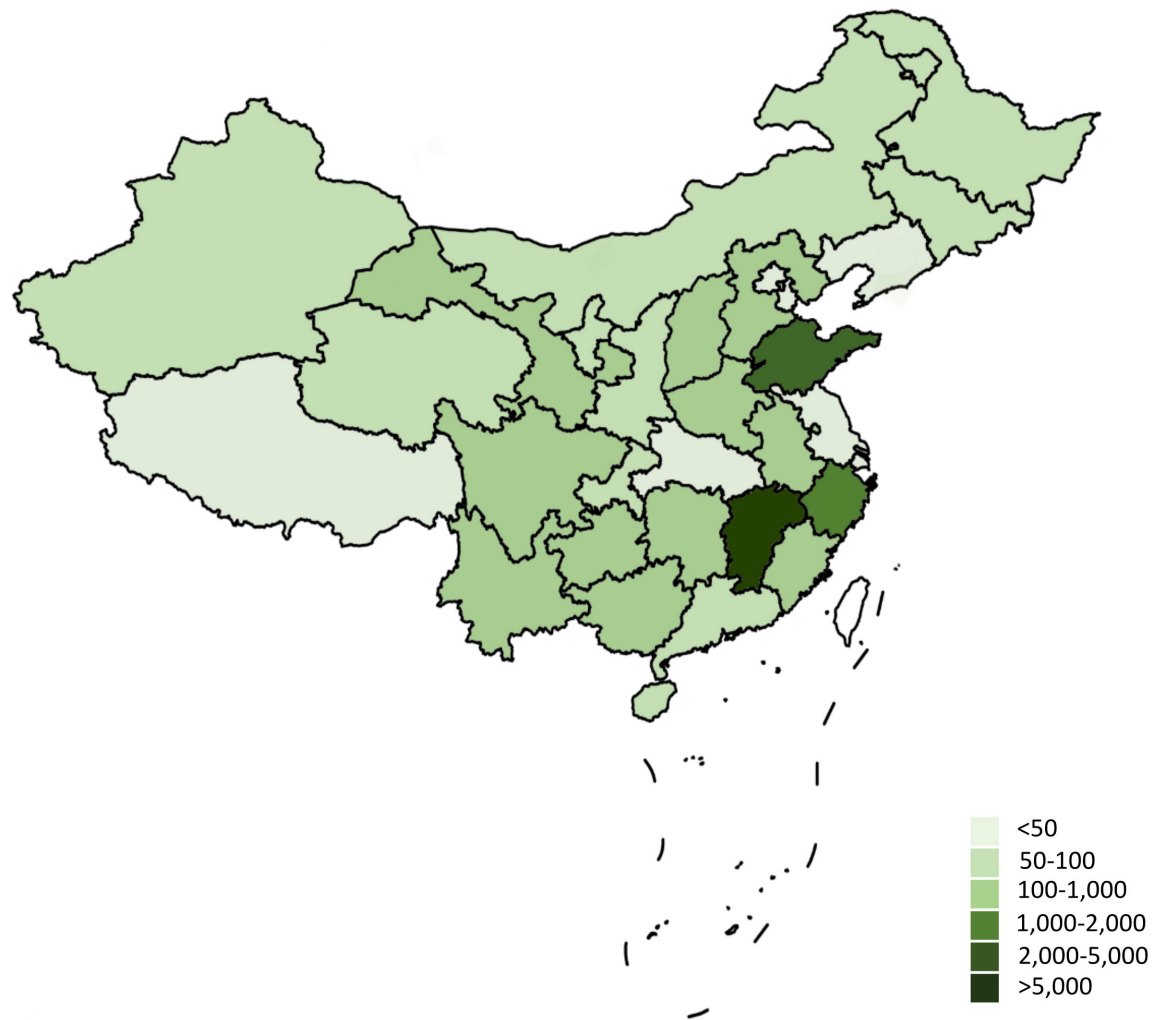
Figure 1. Maps showing the sources of the samples in the Westlake BioBank for Chinese (WBBC) pilot cohort. AH: Anhui province; BJ: Beijing; CQ: Chongqing; FJ: Fujian province; GD: Guangdong province; GS: Gansu province; GX: Guangxi Autonomous Region; GZ: Guizhou province; HA: Henan province; HB: Hubei 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: 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.

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

A



B

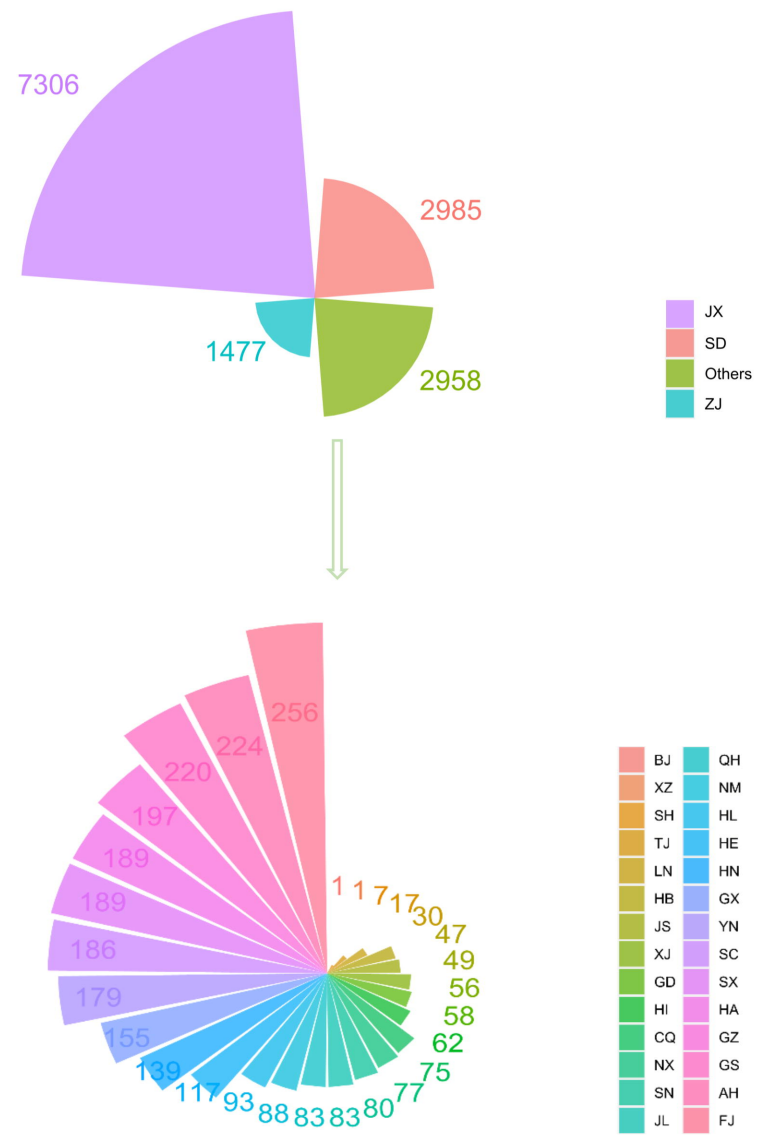


Figure 2

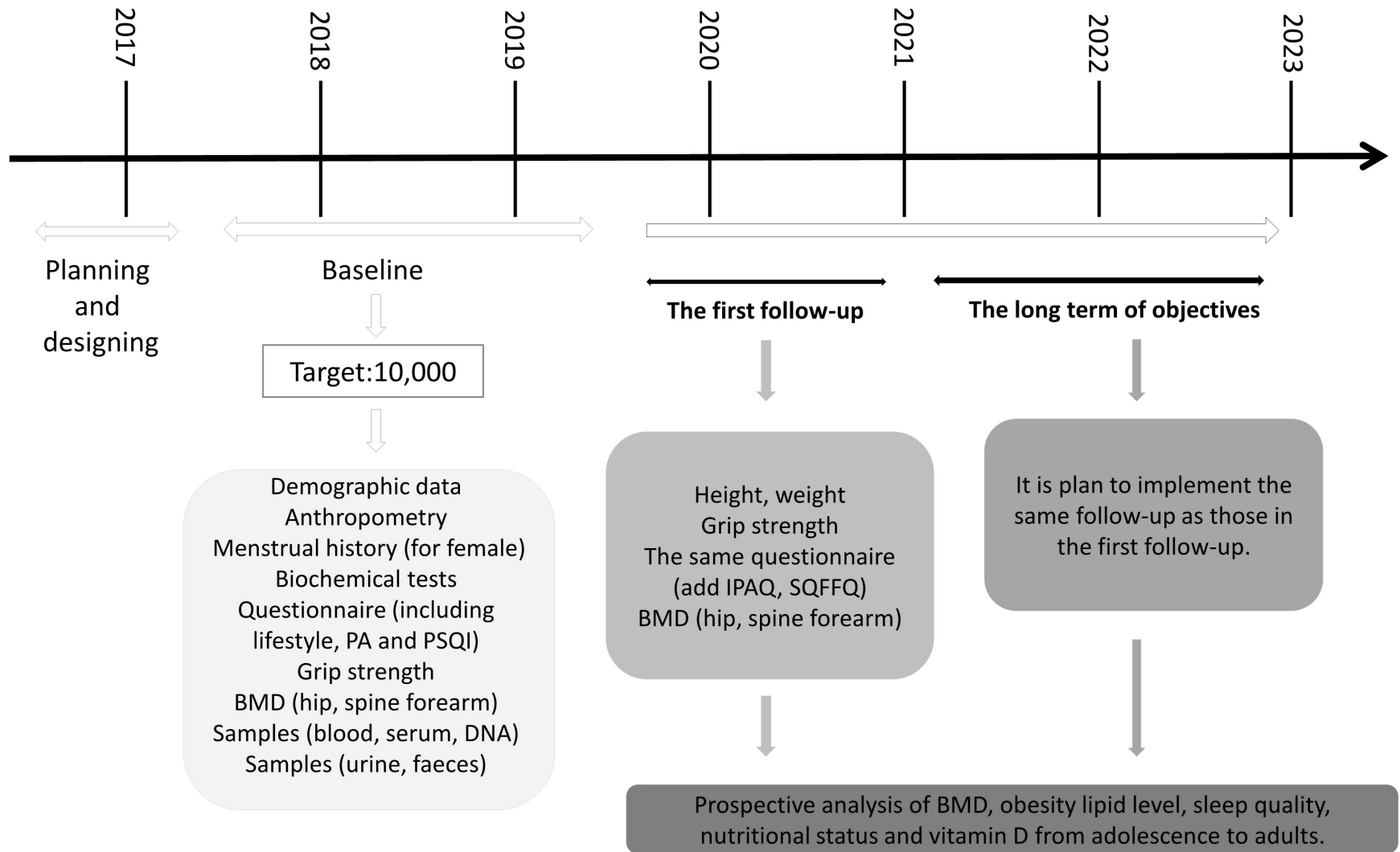


Figure 3

