scientific reports

OPEN

# A reference equation for peak oxygen uptake for cycle ergometry in Chinese adult participants

Jinan Wang[1], Chuan Ren[1], Shunlin Xu[1], Yifang Yuan[2], Yanxin Song[1], Dingding Xie[1], Kexin Wang[1], Lei Yuan[1], Tao Shen[1], Ling Xu[1], Yida Tang[1], Wei Gao[1] & Wei Zhao[1,3]

**Accurately quantifying cardiorespiratory fitness (CRF) through cardiopulmonary exercise testing (CPET) is increasingly important for improving risk assessment and guiding clinical decisions. However, research on $\dot{V}O_2$peak reference values and predictive equations for the Chinese population remains limited. This study aimed to establish a $\dot{V}O_2$peak predictive equation for Chinese adults. This study analyzed healthy participants who underwent CPET at Peking University Third Hospital (PUTH) from September 1, 2017, to September 1, 2023. Data from September 1, 2017, and August 31, 2021 were used as the derivation cohort, and September 1, 2021, to September 1, 2023 were utilized as an external validation cohort for temporal validation. The derivation cohort underwent backward multivariate regression analysis to generate the $\dot{V}O_2$peak prediction equation, which was compared with the widely-used Wasserman, FRIEND and Xiangya equations. The PUTH derivation cohort (N = 4531, mean age: 50.7 years, 18–88 years) and validation cohort (N = 4624, mean age: 46.1 years, 18–89 years) included 48.8 and 48.5% men, respectively. With increasing age, both men and women $\dot{V}O_2$peak exhibited a general decline. The predictive equation for $\dot{V}O_2$peak was established based on the derivation cohort: $\dot{V}O_2$peak (mL·min$^{-1}$) = −24364.9 − 621.3 × Sex (Women = 1, Men = 2) −10.7 × Age + 0.2 × Height$^2$ (cm) + 6464.7 × Log(BMI) −24997.2 × Log(BSA) +12388.6 × Log(LBM) (adjusted $R^2$ = 0.624, $p < 0.001$). It demonstrated higher consistency between measured and predicted results compared to Xiangya, Wasserman, and FRIEND equations. This study presents the PUTH equation, a new $\dot{V}O_2$peak prediction equation for Chinese adults. Compared to existing equations, the PUTH equation shows reduced bias and improved accuracy, providing a more reliable tool for assessing CRF and guiding clinical interventions in the Chinese population.**

**Keywords** Cardiorespiratory fitness, Peak oxygen uptake, Cardiopulmonary exercise testing, $VO_2$peak predictive equation

In recent years, increasing evidence has demonstrated the critical role of cardiorespiratory fitness (CRF) in reducing both cardiovascular and all-cause mortality[1–3], with its prognostic value firmly established across various cardiovascular diseases[4,5]. CRF has been recognized as a vital indicator of cardiovascular health, even recommended by the American Heart Association as a novel vital sign[6]. Peak oxygen consumption ($VO_2$peak) is considered the “gold standard” for laboratory-based CRF measurements and represents the most significant indicator in cardiopulmonary exercise testing (CPET)[7]. $VO_2$peak reflects an individual’s maximal capacity to uptake, transport, and utilize oxygen, serving as both a prognostic marker for preoperative risks and an evaluative tool for assessing responses to exercise training or the effectiveness of patient treatments[7,8]. CPET is currently the sole assessment tool capable of providing a comprehensive, objective, and quantitative evaluation of overall cardiorespiratory function in a single measurement[9]. An individual’s $VO_2$peak is meaningful only when referenced against the normal values of healthy individuals. As the importance of accurately quantifying CRF through CPET has grown, especially in apparently healthy men and women, it has become essential for improving individualized risk assessments, guiding clinical decisions, and enhancing prognoses[10].

[1]Department of Cardiology and Institute of Vascular Medicine, National Health Commission, Key Laboratory of Cardiovascular Molecular Biology and Regulatory Peptides, Key Laboratory of Molecular Cardiovascular Science of Ministry of Education, Beijing Key Laboratory of Cardiovascular Receptors Research, Peking University Third Hospital, No. 49 Huayuanbei Road, Haidian District, Beijing 100191, China. [2]Clinical Research Institute, Institute of Advanced Clinical Medicine, Peking University, Beijing 100191, China. [3]Physical Examination Center of Peking University Third Hospital, Beijing 100191, China. email: beate_vv@126.com

It is well-known that reference values for $VO_2$peak vary due to influences such as region, race, sex, age, body type, lifestyle, and exercise habits[11]. As age increases, $VO_2$peak tends to decrease, with men generally showing higher values than women. Most commonly used equations provide inaccurate estimates of CRF in obese patients, especially those with severe obesity and low CRF[12]. Because athletes typically have higher $VO_2$peak values, comparing their measured $VO_2$peak with the predicted values for the general active population could lead to misdiagnosis[13]. Therefore, each exercise laboratory should select an appropriate set of reference values to best reflect the characteristics of the tested population. Despite numerous studies in this field, the majority have been concentrated in Europe and the Americas[14–16]. In recent years, there has been a rapid increase in CPET-related research among the Chinese population, although research on reference values and prediction equations for $VO_2$peak remains limited.

Therefore, the primary objective of this study is to establish a prediction equation for $VO_2$peak in Chinese adults using an electronically braked cycle ergometer for CPET and to assess the impact of sex, age, and body mass index (BMI) on reference values. Previous research has indicated that CRF in obese populations is often underestimated, and as such, we incorporate both lean body mass (LBM) and body surface area (BSA) to construct the equation[17].

## Method

### Participants

This retrospective study encompassed 22,943 CPET data sets conducted at the cardiovascular rehabilitation center of Peking University Third Hospital (PUTH) from September 1, 2017, to September 1, 2023. All participants underwent an evaluation by an experienced physician prior to exercise. The pre-exercise assessment included anamnesis, physical examination, resting electrocardiogram, exercise habit, and medication. Height and body weight were recorded to the nearest 0.5 cm and 0.1 kg, respectively, with participants dressed in light clothes and no shoes. The inclusion criteria for participants were: (1) age ≥ 18 years; (2) tested on electronically braked cycle ergometers; (3) in the case of multiple tests on the same participant, only the first test result was considered to avoid bias due to familiarity with CPET. Participants with a history of cardiovascular or pulmonary diseases, as well as musculoskeletal or neurological disorders, were excluded from the cohort. Participants who terminated the test due to abnormal clinical findings before reaching voluntary maximal effort or the occurrence of a positive electrocardiographic reaction were also excluded. A total of 9,155 participants were included, and additional participant information is illustrated in (Fig. 1).

This study adhered to the STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) guidelines[18] and was carried out in compliance with the amended Declaration of Helsinki[19]. Due to the retrospective nature of the study, Peking University Third Hospital Medical Science Research Ethics Committee (M2023790) waived the need of obtaining informed consent.

### Cardiopulmonary exercise testing

We used electronically braked cycle ergometer for symptom-limited exercise testing to evaluate $VO_2$peak. Gas exchange and ventilatory variables were continuously analyzed on a breath-by-breath basis using the computer-

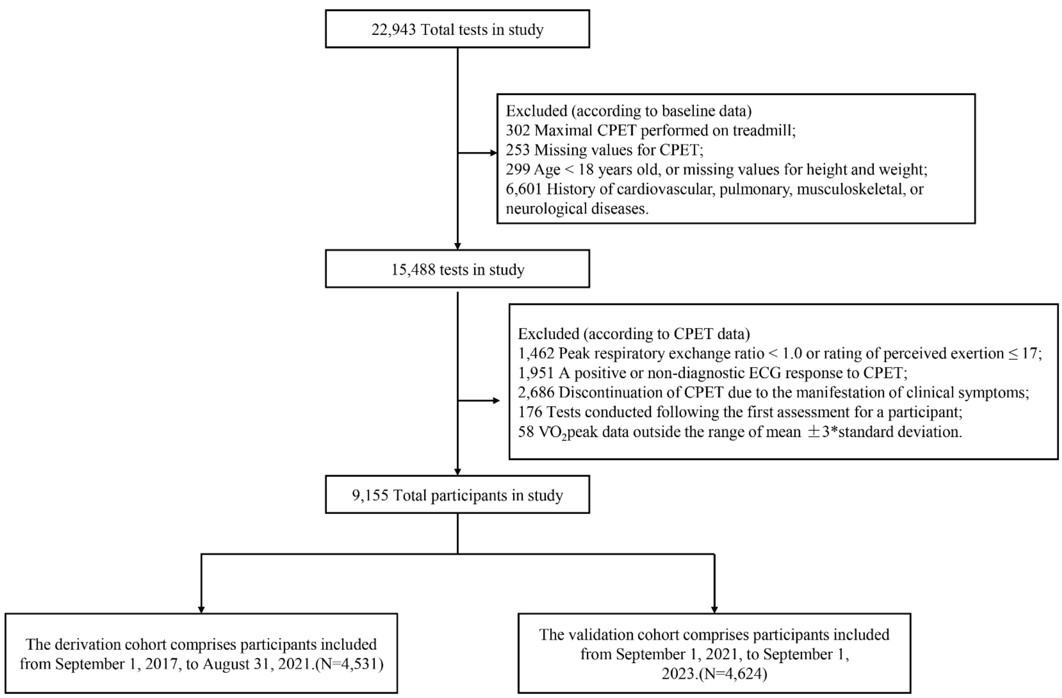


**Fig. 1**. Flowchart of the inclusion profile in the study. *CPET* cardiopulmonary exercise testing.

based MedGraphics system (Minnesota, USA) and Madecare system (Beijing, China). Despite differences in CPET equipment, all operators underwent uniform training to maintain consistent procedures, following the guidelines[20]. Calibration of the gas exchange measurement systems was performed daily before testing to ensure the accuracy of gas exchange determinations. $VO_2$peak was defined as the highest average value during a sequential 30-s period.

The certified and trained clinical exercise physiologists performed the CPET. Standardized procedures and instructions were followed to ensure consistent testing conditions for all participants, with a physician available at all times upon request. Before each test, the equipment was set up in a standardized manner, with reference gas calibration and volume calibration were performed. In the absence of chest pain and electrocardiogram abnormalities, participants were encouraged to exercise continuously until reaching maximal fatigue, which was defined as voluntary exhaustion, dyspnea, or fatigue.

The protocol included a 3 min resting period, followed by a 3 min unloaded warm-up, with a pedaling cadence of 60–70 revolutions per minute. Individualized continuous incremental load programs were devised based on the protocol, with a preset load increment rate of 10–40 W/min[16]. Participants were set to achieve symptom-limited CPET within 8–12 min. Upon self-perceived fatigue (such as shortness of breath and leg fatigue), participants entered the recovery period, during which data were continuously recorded for 6 min. 12-lead electrocardiogram was monitored continuously. Blood pressure was recorded every 3 min using an automatic digital sphygmomanometer. Oxygen saturation ($SPO_2$) was continuously recorded by a pulse oximeter. A CPET was deemed valid if at least two of the following criteria were met: (1) RER $\geq$ 1.0; (2) rating of perceived exertion > 17; (3) $VO_2$ reaches a plateau or slightly declines despite increased load.

### Development of the PUTH equation and comparison with other prediction equations

Data from participants between September 1, 2017, and August 31, 2021 (N = 4,531) constituted the derivation cohort, while data from September 1, 2021, to September 1, 2023 (N = 4,624) were used as a temporal validation cohort for external validation.

The derivation cohort underwent backward multivariate regression analysis to generate the $VO_2$peak prediction equation. Considering the correlation between BSA and LBM with CRF[21,22], we attempted to incorporate these indicators into the $VO_2$peak equation. These indicators were calculated using formulas derived from Chinese data[23,24], rather than being directly measured by the device. In the validation cohort, we compared the predicted $VO_2$peak and the percentage of achieved $VO_2$peak across the PUTH equation. For comparison with the measured $VO_2$peak, predicted $VO_2$peak for each participant were determined using the PUTH equation, Xiangya equation[25], Wasserman equation[16], and FRIEND equation[15]. The comparison was conducted both overall and stratified by sex, age groups, and BMI categories (underweight < 18.5 $kg{\cdot}m^{-2}$, normal weight 18.5–23.9 $kg{\cdot}m^{-2}$, overweight 24.0–27.9 $kg{\cdot}m^{-2}$, obese $\geq$ 28 $kg{\cdot}m^{-2}$)[26].

### Statistical analysis

Data analysis was conducted using SPSS 26.0 (IBM, Armonk, New York). Continuous variables are reported as mean ± standard deviation, while categorical variables are reported as frequency (percentage). Additionally, one-way analysis of variance and chi-square tests were used for comparisons between derivation and validation cohorts. Box plots were generated to analyze trends in $VO_2$peak across different age groups. Bland–Altman plots and scatter plots were constructed to illustrate differences between the measured and predicted values from the PUTH equation, Xiangya equation, Wasserman equation, and FRIEND equation. P value < 0.05 was considered statistically significant.

## Results

### Baseline characteristics of PUTH derivation cohort and PUTH validation cohort

The baseline characteristics of PUTH derivation cohort and PUTH validation cohort are presented in Table 1. The PUTH derivation cohort included 4,531 participants with a mean age of 50.7 years, of whom 48.8% were men. The PUTH validation cohort included 4,624 participants with a mean age of 46.1 years, with 48.5% being men. Figure 2 illustrates the distribution of $VO_2$peak in derivation cohort by sex and age groups. $VO_2$peak in both men and women declined as a whole with age. At all ages, men had higher $VO_2$peak than women.

There were significant differences in age, height, weight, and BMI between the PUTH derivation and validation cohorts ($p < 0.05$). Overall, participants in the PUTH validation cohort were younger (46.1 ± 14.4 vs. 50.7 ± 14.6 years), taller (167.3 ± 8.4 vs. 166.6 ± 8.4 cm), lighter (69.0 ± 13.9 vs. 69.7 ± 13.5 kg), and had a lower BMI (24.5 ± 3.8 vs. 25.0 ± 3.8 $kg{\cdot}m^{-2}$) compared to the derivation cohort. However, there were no significant differences in sex, BSA, and LBM between the two cohorts. Additionally, in terms of CPET parameters, compared to the PUTH derivation cohort, the PUTH validation cohort had a lower peak RER (1.17 ± 0.10 vs. 1.18 ± 0.11, $p < 0.001$) and a higher $VO_2$peak (23.0 ± 6.6 vs. 21.6 ± 6.1 $mL{\cdot}kg^{-1}{\cdot}min^{-1}$, $p < 0.001$).

### PUTH prediction equation

The backward multivariate regression analysis revealed that sex, age, square of height, logarithmic transformations of BMI, logarithmic transformations of BSA, and logarithmic transformations of LBM were significant predictors of $VO_2$peak (adjusted $R^2 = 0.624$, $p < 0.001$). The equation is as follows:

$VO_2$peak ($mL{\cdot}min^{-1}$) = −24364.9−621.3×Sex (Women=1, Men=2)−10.7×Age+0.2×Height$^2$ (cm) + 6464.7×Log(BMI)−24997.2×Log(BSA) + 12388.6×Log(LBM).

| Variables | PUTH derivation cohort | | | PUTH validation cohort | | |
|---|---|---|---|---|---|---|
| | All (N = 4531) | Men (N = 2210) | Women (N = 2321) | All (N = 4624) | Men (N = 2244) | Women (N = 2380) |
| Age (yrs) | 50.7 ± 14.6 | 48.4 ± 14.8 | 53.0 ± 14.1 | 46.1 ± 14.4*** | 44.2 ± 14.0 | 48.0 ± 14.5 |
| Height (cm) | 166.6 ± 8.4 | 173.0 ± 6.1 | 160.6 ± 5.4 | 167.3 ± 8.4*** | 173.7 ± 6.0 | 161.2 ± 5.4 |
| Weight (kg) | 69.7 ± 13.5 | 77.0 ± 12.8 | 62.7 ± 10.1 | 69.0 ± 13.9** | 77.0 ± 12.9 | 61.3 ± 9.9 |
| BMI ($kg{\cdot}m^{-2}$) | 25.0 ± 3.8 | 25.7 ± 3.7 | 24.3 ± 3.7 | 24.5 ± 3.8*** | 25.5 ± 3.7 | 23.6 ± 3.6 |
| BSA ($m^2$) | 1.9 ± 0.2 | 2.0 ± 0.2 | 1.8 ± 0.1 | 1.9 ± 0.2 | 2.0 ± 0.2 | 1.7 ± 0.1 |
| LBM (kg) | 47.6 ± 9.4 | 55.8 ± 5.8 | 39.8 ± 3.9 | 47.6 ± 9.7 | 56.2 ± 5.8 | 39.5 ± 3.8 |
| Peak RER | 1.18 ± 0.11 | 1.19 ± 0.11 | 1.17 ± 0.10 | 1.17 ± 0.10*** | 1.17 ± 0.10 | 1.16 ± 0.10 |
| $VO_2$peak ($mL{\cdot}kg^{-1}{\cdot}min^{-1}$) | 21.6 ± 6.1 | 24.0 ± 6.4 | 19.3 ± 4.7 | 23.0 ± 6.6*** | 25.6 ± 7.0 | 20.6 ± 5.2 |
| $VO_2$peak ($mL{\cdot}min^{-1}$) | 1503.4 ± 511.7 | 1828.9 ± 497.5 | 1193.4 ± 280.2 | 1585.1 ± 542.1*** | 1942.6 ± 513.7 | 1248.0 ± 296.6 |

**Table 1**. Baseline characteristics of the derivation cohort and validation cohort. Data are presented as arithmetic mean ± SD. *BMI* body mass index, *BSA* body surface area, *LBM* lean body mass, *PUTH* Peking University Third Hospital, *RER* respiratory exchange rate, $\dot{V}O_2$*peak* peak oxygen uptake. *Significantly different comparison between derivation cohort and validation cohort, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

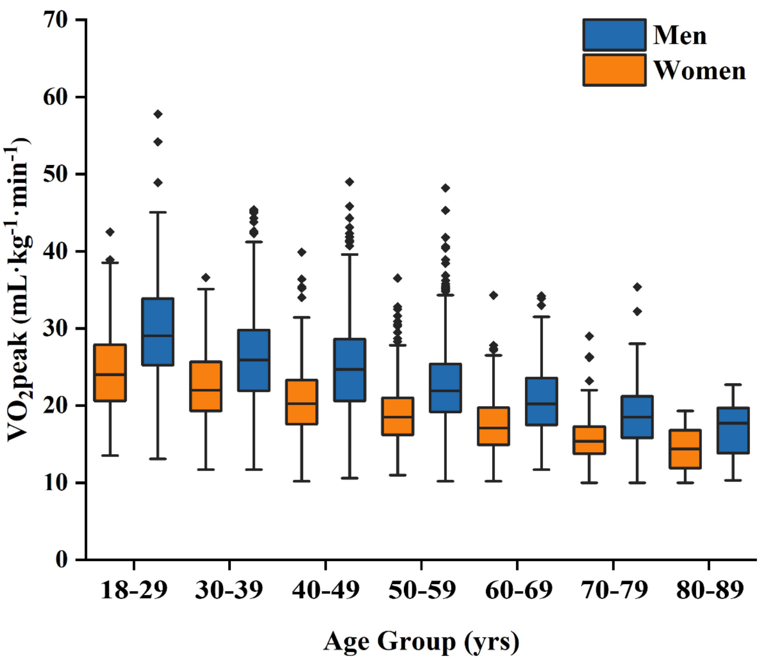


**Fig. 2**. The distribution of $VO_2$peak in derivation cohort by sex and age groups.

### Comparison of measured $\dot{V}O_2$peak and predicted $\dot{V}O_2$peak by the PUTH equation in the PUTH validation cohort

We calculated the predicted $VO_2$peak using the PUTH equation in the validation cohort, and compared it with measured $VO_2$peak by stratifying participants by sex, age groups, and BMI categories, as shown in Table 2. Measured and predicted $VO_2$peak differed significantly between men and women, participants aged 30–39, 40–49, and 50–59 years, and individuals with BMI values of 18.5–23.9 and 24–28 $kg{\cdot}m^{-2}$ ($p < 0.001$). The percentage of predicted $VO_2$peak was 102.5%, indicating minimal overall differences between measured and predicted values in the entire cohort. This finding highlights the high accuracy and broad applicability of the PUTH equation in predicting $VO_2$peak.

### Comparison of the mean difference between measured and predicted $\dot{V}O_2$peak of different equations in the PUTH validation cohort

We compared the mean differences between measured and predicted $VO_2$peak using the PUTH, Xiangya, Wasserman, and FRIEND equations in the PUTH validation cohort. In the total population, the absolute mean difference between the measured and predicted $VO_2$peak using the PUTH equation was the smallest (37.0 ± 349.7 $mL{\cdot}min^{-1}$), followed by the mean differences for the Xiangya (209.6 ± 367.0 $mL{\cdot}min^{-1}$), Wasserman (456.5 ± 426.5 $mL{\cdot}min^{-1}$), and FRIEND equations (739.7 ± 487.6 $mL{\cdot}min^{-1}$), respectively. A statistical difference was found among the equations ($p < 0.001$).

| Variables | Measured $VO_2$peak ($mL \cdot min^{-1}$) | $VO_2$peak predicted by the PUTH equation ($mL \cdot min^{-1}$) | % predicted | *p* value |
|---|---|---|---|---|
| Total (N = 4624) | 1585.1 ± 542.1 | 1548.1 ± 410.4 | 102.5 ± 21.6 | < 0.001 |
| Sex groups | | | | |
| Men (N = 2244) | 1942.6 ± 513.7 | 1892.5 ± 285.6 | 102.8 ± 23.0 | < 0.001 |
| Women (N = 2380) | 1248.0 ± 296.6 | 1223.4 ± 182.0 | 102.2 ± 20.2 | < 0.001 |
| Age groups (yrs) | | | | |
| 18–29 (N = 675) | 1856.3 ± 595.2 | 1875.1 ± 411.4 | 98.6 ± 20.9 | 0.231 |
| 30–39 (N = 940) | 1785.1 ± 550.3 | 1755.3 ± 384.2 | 101.7 ± 21.5 | 0.020 |
| 40–49 (N = 1081) | 1671.1 ± 504.4 | 1587.2 ± 341.9 | 105.4 ± 21.9 | < 0.001 |
| 50–59 (N = 1004) | 1468.0 ± 447.4 | 1409.3 ± 302.7 | 104.4 ± 22.5 | < 0.001 |
| 60–69 (N = 677) | 1266.7 ± 368.3 | 1250.0 ± 263.3 | 101.4 ± 19.8 | 0.087 |
| 70–79 (N = 230) | 1064.2 ± 295.3 | 1086.8 ± 222.8 | 98.5 ± 20.7 | 0.137 |
| 80–89 (N = 17) | 930.2 ± 212.8 | 927.6 ± 184.5 | 101.7 ± 23.0 | 0.956 |
| BMI categories ($kg \cdot m^{-2}$) | | | | |
| < 18.5 (N = 142) | 1217.1 ± 391.8 | 1215.3 ± 334.8 | 101.6 ± 24.5 | 0.940 |
| 18.5–23.9 (N = 2063) | 1480.2 ± 510.0 | 1434.9 ± 384.9 | 103.5 ± 22.6 | < 0.001 |
| 24–27.9 (N = 1675) | 1663.6 ± 549.2 | 1624.0 ± 392.5 | 102.2 ± 20.8 | < 0.001 |
| ≥ 28 (N = 744) | 1769.4 ± 541.8 | 1754.8 ± 395.1 | 100.7 ± 19.5 | 0.268 |

**Table 2**. Measured and predicted $VO_2$peak values by the PUTH equation and percentage of predicted $VO_2$peak in the validation cohort were stratified by sex, age groups, and BMI categories. Data are presented as arithmetic mean ± SD. *BMI* body mass index, *FRIEND* fitness and the importance of exercise: a national data base, *PUTH* Peking University Third Hospital, *RER* respiratory exchange rate, $VO_2$*peak* peak oxygen uptake. *p* value from comparisons between measured and predicted $VO_2$peak.

| Variables | PUTH equation ($mL \cdot min^{-1}$) | Xiangya equation ($mL \cdot min^{-1}$) | Wasserman equation ($mL \cdot min^{-1}$) | FRIEND equation ($mL \cdot min^{-1}$) | *p* value |
|---|---|---|---|---|---|
| Total (N = 4624) | 37.0 ± 349.7 | 209.6 ± 367.0*** | −456.5 ± 426.5*** | −739.7 ± 487.6*** | < 0.001 |
| Sex groups | | | | | |
| Men (N = 2244) | 50.1 ± 432.9 | 284.9 ± 448.8*** | −645.0 ± 474.9*** | −1021.8 ± 477.9*** | < 0.001 |
| Women (N = 2380) | 24.6 ± 246.3 | 138.6 ± 248.0*** | −278.6 ± 275.0*** | −473.7 ± 317.8*** | < 0.001 |
| Age groups (yrs) | | | | | |
| 18–29 (N = 675) | −18.8 ± 406.5 | 246.0 ± 444.0*** | −762.9 ± 464.8*** | −1084.6 ± 483.4*** | < 0.001 |
| 30–39 (N = 940) | 29.7 ± 389.8 | 247.5 ± 403.3*** | −609.9 ± 443.5*** | −930.8 ± 487.2*** | < 0.001 |
| 40–49 (N = 1081) | 83.9 ± 358.6 | 258.1 ± 370.7*** | −425.6 ± 395.9*** | −724.3 ± 442.3*** | < 0.001 |
| 50–59 (N = 1004) | 58.7 ± 331.6 | 195.0 ± 338.4*** | −326.1 ± 362.8*** | −591.8 ± 423.7*** | < 0.001 |
| 60–69 (N = 677) | 16.7 ± 253.5 | 123.2 ± 255.2*** | −269.2 ± 275.3*** | −501.0 ± 359.6*** | < 0.001 |
| 70–79 (N = 230) | −22.6 ± 230.1 | 52.9 ± 238.0* | −216.2 ± 261.7*** | −391.8 ± 381.7*** | < 0.001 |
| 80–89 (N = 17) | 2.6 ± 189.5 | 0.5 ± 173.2 | −172.6 ± 170.8* | −398.2 ± 317.3*** | < 0.001 |
| BMI categories ($kg \cdot m^{-2}$) | | | | | |
| < 18.5 (N = 142) | 1.8 ± 283.7 | 101.3 ± 291.4* | −529.6 ± 341.0*** | −702.1 ± 345.1*** | < 0.001 |
| 18.5–23.9 (N = 2063) | 45.3 ± 341.9 | 227.3 ± 356.1*** | −417.3 ± 403.9*** | −686.5 ± 446.3*** | < 0.001 |
| 24–27.9 (N = 1675) | 39.6 ± 359.1 | 231.5 ± 378.2*** | −461.2 ± 435.1*** | −770.7 ± 498.4*** | < 0.001 |
| ≥ 28 (N = 744) | 14.6 ± 360.5 | 131.6 ± 370.9*** | −540.2 ± 467.1*** | −824.5 ± 571.8*** | < 0.001 |

**Table 3**. Difference in measured and predicted $VO_2$peak values by multiple equations in the validation cohort were stratified by sex, age groups, and BMI categories. Data are presented as arithmetic mean ± SD. *BMI* body mass index, *FRIEND* fitness and the importance of exercise: a national data base, *PUTH* Peking University Third Hospital, *RER* respiratory exchange rate, $VO_2$*peak* peak oxygen uptake. *Significantly different *vs.* PUTH equation, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

When the study population was further divided by sex, age, and BMI, the mean differences between measured and predicted $VO_2$peak were significantly smaller with the PUTH equation compared to the other three equations in most subgroups ($p < 0.001$). Except for the 80–89 age group, where the Xiangya equation showed the smallest difference (0.5 ± 173.2 $mL \cdot min^{-1}$), the PUTH equation generally had the smallest absolute differences in all other sex, age groups, and BMI categories followed by the Xiangya, Wasserman, and FRIEND equations (see Table 3).

### Comparison of people whose absolute value of (measured $\dot{V}O_2$peak-predicted $\dot{V}O_2$peak)/measured $\dot{V}O_2$peak is > 20% in different prediction equations

Table 4 presents the proportion of individuals in the PUTH validation cohort whose absolute value of (measured $VO_2$peak-predicted $VO_2$peak) / measured $VO_2$peak exceeded 20%, based on different equations and stratified by sex, age groups, and BMI categories. In the validation cohort, the proportion exceeding 20% were 33.0, 39.5, 62.9, and 81.4% for the PUTH, Xiangya, Wasserman, and FRIEND equations, respectively ($p < 0.001$). The PUTH equation had the fewest individuals exceeding 20% in both men and women. Among age groups, those under 70 had the fewest instances exceeding 20% with the PUTH equation, while those over 70 had the fewest with the Xiangya equation. In all BMI categories, the PUTH equation resulted in the lowest number of individuals exceeding the 20% threshold. These differences were statistically significant across all equations ($p < 0.001$).

### Consistency evaluation and accuracy between prediction equations

The consistency evaluation between measured and predicted results using Bland–Altman analysis for different equations is shown in Fig. 3. Among sex-specific evaluations, the PUTH equation exhibited the least bias for both men (50.2 mL·min$^{-1}$) and women (24.6 mL·min$^{-1}$). The Xiangya equation showed a slightly higher bias (Men = 284.9 mL·min$^{-1}$; Women = 138.6 mL·min$^{-1}$), while the FRIEND equation exhibited the highest bias (Men = -1021.6 mL·min$^{-1}$; Women = −473.7 mL·min$^{-1}$).

Figure 4 presents scatter plots comparing the measured $VO_2$peak with the predicted values from four equations for men and women. The PUTH equation aligns most closely with the 45-degree reference line, demonstrating the highest accuracy. In contrast, the Xiangya equation systematically underestimates $VO_2$peak across the entire range, while the Wasserman and FRIEND equations overestimate $VO_2$peak, particularly at higher values. These findings indicate that the PUTH equation provides the most accurate predictions across different sex groups and $VO_2$peak levels.

## Discussion

This study presents a new reference equation for $VO_2$peak in Chinese adults, derived from a large cohort tested with electronically braked cycle ergometers. The PUTH equation, which incorporates sex, age, height, BMI, LBM, and BSA, outperforms existing equations in predicting $VO_2$peak, providing more accurate results across different BMI categories. Notably, it highlights the importance of body composition factors, especially LBM and BSA, in estimating CRF. The PUTH equation demonstrates minimal bias and better accuracy, making it highly applicable in clinical settings for diverse age and BMI groups.

### Factors influencing $\dot{V}O_2$peak

Previous research has indicated that $VO_2$peak is associated with genetics, age, sex, body type, exercise habits, lifestyle, and cardiovascular condition[14,25,27–30]. A systematic review suggested that $VO_2$peak is higher in men than women and decreases with age[30]. Consistent with these findings, our study based on a Chinese cohort also revealed higher measured $VO_2$peak in men, gradually decreasing with age (see Fig. 2). However, our study provides some unique insights, especially regarding the consideration of body composition and BMI.

| Variables | PUTH equation | Xiangya equation | Wasserman equation | FRIEND equation | *p* value |
|---|---|---|---|---|---|
| Total (N = 4624) | 1526 (33.0) | 1827 (39.5)* | 2909 (62.9)* | 3762 (81.4)* | < 0.001 |
| Sex groups | | | | | |
| Men (N = 2244) | 816 (36.4) | 1006 (44.8)* | 1583 (70.5)* | 1986 (88.5)* | < 0.001 |
| Women (N = 2380) | 710 (29.8) | 821 (34.5)* | 1326 (55.7)* | 1776 (74.6)* | < 0.001 |
| Age groups (yrs) | | | | | |
| 18–29 (N = 675) | 232 (34.4) | 282 (41.8)* | 536 (79.4)* | 629 (93.2)* | < 0.001 |
| 30–39 (N = 940) | 306 (32.6) | 378 (40.2)* | 676 (71.9)* | 839 (89.3)* | < 0.001 |
| 40–49 (N = 1081) | 365 (33.8) | 446 (41.3)* | 662 (61.2)* | 870 (80.5)* | < 0.001 |
| 50–59 (N = 1004) | 346 (34.5) | 418 (41.6)* | 561 (55.9)* | 757 (75.4)* | < 0.001 |
| 60–69 (N = 677) | 194 (28.7) | 224 (33.1) | 347 (51.3)* | 508 (75.0)* | < 0.001 |
| 70–79 (N = 230) | 77 (33.5) | 75 (32.6) | 120 (52.2)* | 146 (63.5)* | < 0.001 |
| 80–89 (N = 17) | 6 (35.3) | 4 (23.5) | 7 (41.2) | 13 (76.5) | 0.013 |
| BMI categories (kg·m$^{-2}$) | | | | | |
| < 18.5 (N = 142) | 55 (38.7) | 55 (38.7) | 118 (83.1)* | 126 (88.7)* | < 0.001 |
| 18.5–23.9 (N = 2063) | 718 (34.8) | 861 (41.7)* | 1327 (64.3)* | 1695 (82.2)* | < 0.001 |
| 24–27.9 (N = 1675) | 530 (31.6) | 686 (41.0)* | 1002 (59.8)* | 1360 (81.2)* | < 0.001 |
| ≥ 28 (N = 744) | 223 (30.0) | 225 (30.2) | 462 (62.1)* | 581 (78.1)* | < 0.001 |

**Table 4.** Comparison of people whose absolute value of (measured $VO_2$peak-predicted $VO_2$peak)/measured $VO_2$peak is > 20% in different prediction equations. Data are presented as n (%). *BMI* body mass index, *FRIEND* fitness and the importance of exercise: a national data base, *PUTH* Peking University Third Hospital, *RER* respiratory exchange rate, $VO_2$*peak* peak oxygen uptake. *Significantly different *vs.* PUTH equation ($p < 0.05$).

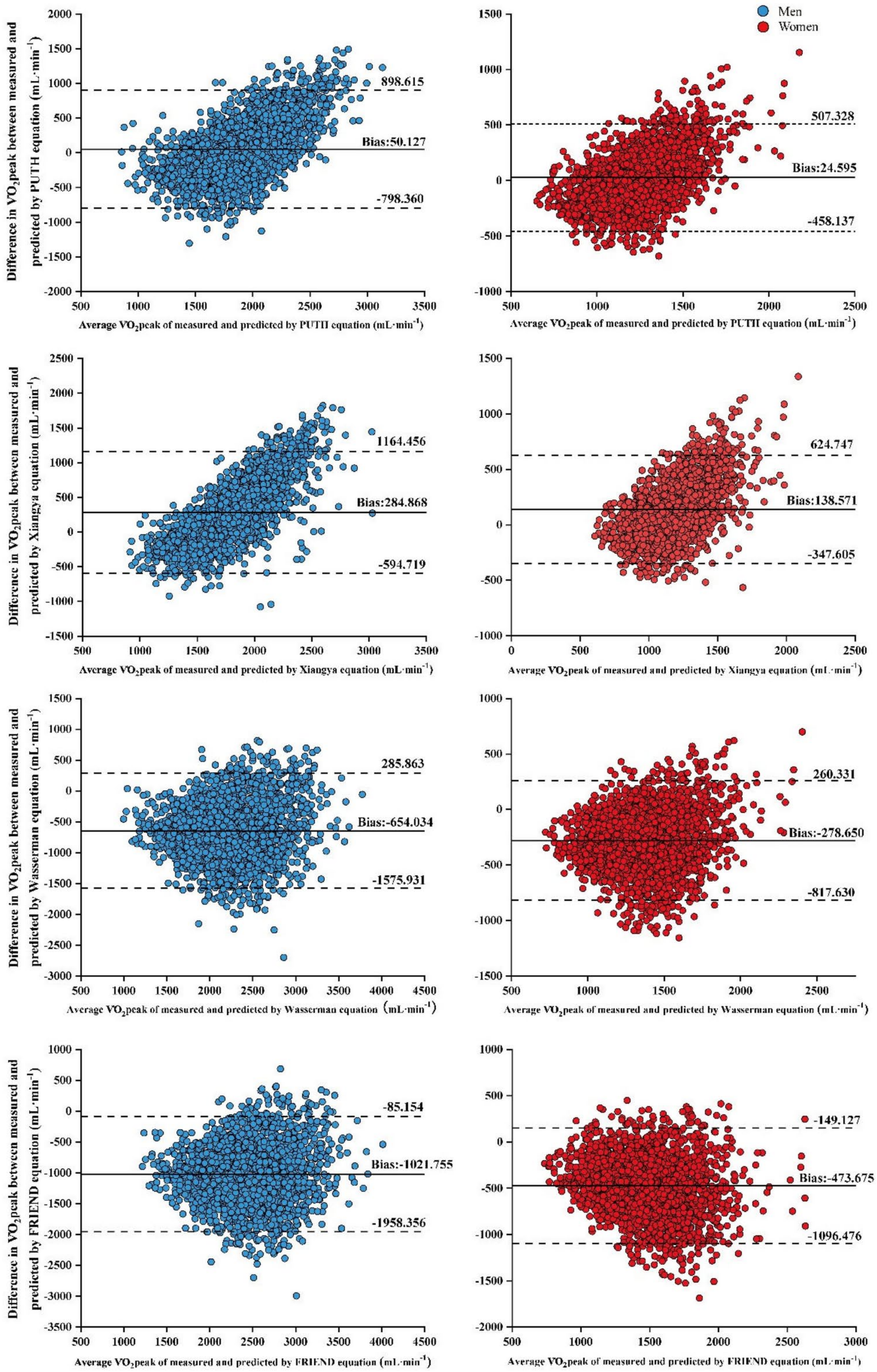


**Fig. 3**. The consistency evaluation between measured and predicted results for different equations. Bland–Altman plots were utilized to depict differences obtained from the PUTH, Xiangya, Wasserman, and FRIEND equations.

Our study further emphasizes the critical role of LBM and BSA in predicting $VO_2$peak, significantly improving the accuracy of $VO_2$peak prediction. Additionally, ethnic differences also play a crucial role in physiological functions, influenced by genetics, environment, and social factors[31], affecting the overall function of oxygen intake through respiration, transportation by the cardiovascular system to organs and skeletal muscles, and eventual elimination of carbon dioxide[32,33]. A study conducted in China by *Yan* et al. demonstrated significant

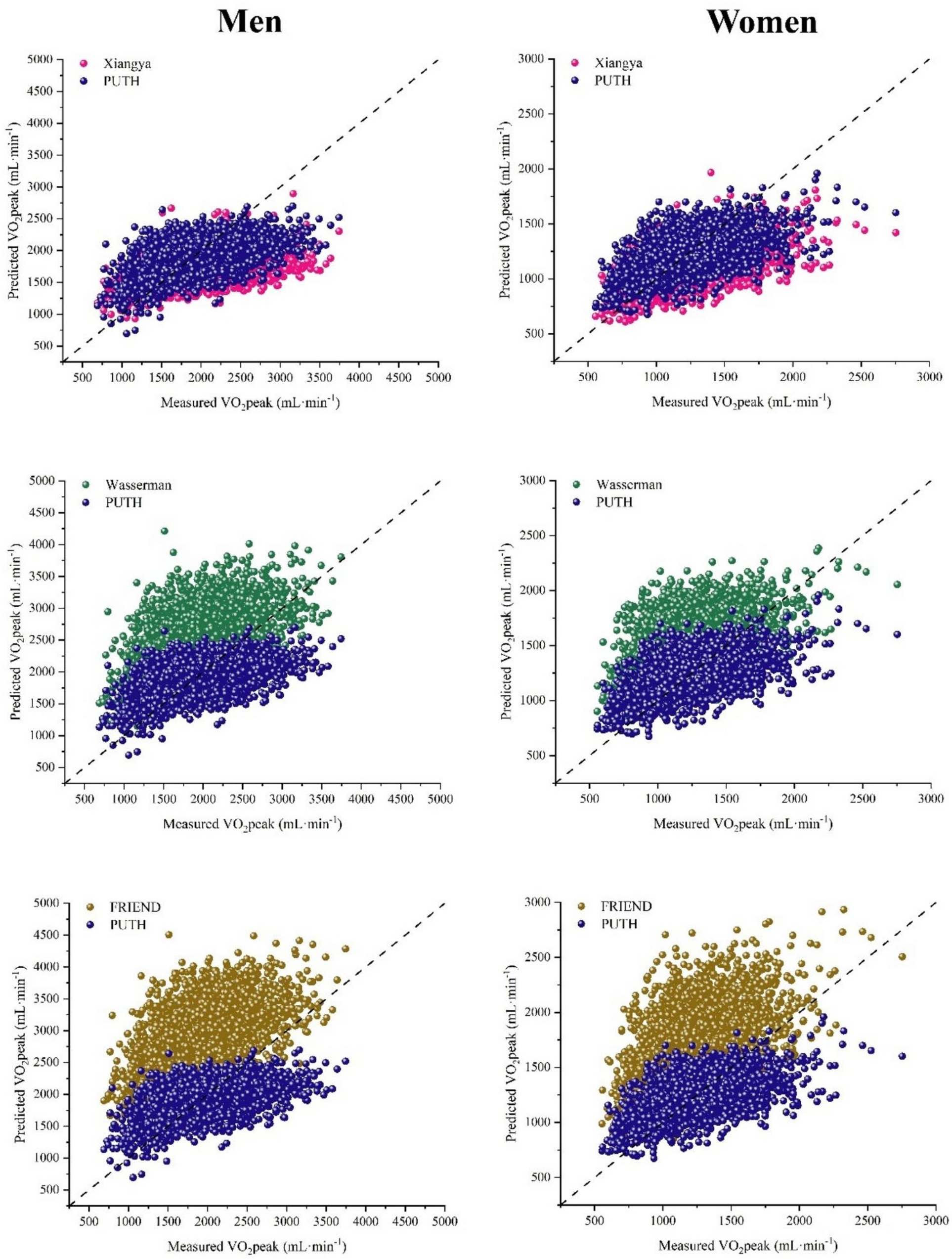


**Fig. 4**. Scatter plots comparing measured $VO_2$peak values with predicted values from four equations (Xiangya, Wasserman, FRIEND, and PUTH). The dashed 45-degree line represents the line of identity, where perfect agreement between measured and predicted values would lie. Points deviating from the identity line indicate discrepancies between measured and predicted values.

differences in forced expiratory volume in 1 s ($FEV_1$) and forced vital capacity (FVC) among different ethnic groups, even after adjusting for age, height, sex, and residence differences[34]. These findings further underscored the complex interplay between body composition, genetic background, and other environmental factors, suggesting that a comprehensive consideration of multiple factors was essential when assessing $VO_2$peak, particularly across different national and ethnic groups for personalized analysis.

## Significance of establishing $\dot{V}O_2$peak prediction equation in China

$VO_2$peak is a key indicator of CRF and has established reference values based on cohort studies in the United States and Europe. However, as shown by the results of this study, widely used foreign equations are not applicable to the Chinese population, possibly due to significant differences in genetics, lifestyle, body type, and environmental factors. Most of the existing $VO_2$peak prediction equations for Chinese adults have been derived from small sample sizes, which may not fully reflect the physiological characteristics of the Chinese population[25,35,36]. Knowing an individual's CRF as a percentage of the normal range for their age and sex helps classify their fitness level, providing a basis for activity counseling, risk stratification, and therapy evaluation. Establishing a reference equation tailored to the Chinese population is crucial for clinical practice and public health, as it will provide a reliable tool for evaluating CRF across diverse age groups and health conditions in China.

### Advantages of the PUTH equation

By including 9155 participants from China and conducting temporal validation on the validation cohort, we developed the PUTH equation to predict $VO_2$peak reference values for the Chinese population. Consistent with most studies in Western countries, sex, age, and height were identified as important variables in calculating $VO_2$peak reference values. Additionally, our PUTH equation included BMI, BSA, and LBM as independent variables, introducing new variables not previously considered in Chinese population-based cohort studies, which is essential for personalized health assessments, disease prevention, and the development of targeted interventions.

The derivation of the PUTH equation is widely representative, with participants spanning a broad range of ages (18–88 years), weights (40.0–186.9 kg), and heights (140.0–207.0 cm), including a substantial number of women (Women = 2321, 51.2%). Furthermore, the effectiveness of PUTH equation was also assessed in a temporal validation cohort within the Chinese population. In the validation sample, women accounted for 51.5%, and the age (18–89 years), height (143.0–196.0 cm), and weight (38.4–171.5 kg) ranges were similarly broad, effectively covering various subgroups, including women, youth, elderly, and obese individuals (see Table 1). This is uncommon in previous equations and enables a more precise assessment of CRF in these particular populations. Furthermore, in the validation cohort, the ratio of measured to predicted $VO_2$peak values was 102.5%, indicating minimal differences between measured and predicted $VO_2$peak values across different sex, age groups, and BMI categories, particularly in young (< 30 years old) and elderly (> 60 years old) individuals, as well as underweight (BMI < 18.5 $kg \cdot m^{-2}$) and obese individuals (BMI ≥ 28 $kg \cdot m^{-2}$), with no significant statistical differences (see Table 2).

### Comparison of the PUTH prediction equation with other prediction equations

Our study yielded several key findings. First, compared to the Xiangya, Wasserman, and FRIEND equations used for calculating $VO_2$peak reference values, the PUTH equation produced values that remained more stable across different sex, age groups, and BMI categories. Specifically, the mean difference between measured and predicted $VO_2$peak values using the PUTH equation was minimal (37.0 $mL \cdot min^{-1}$), whereas the mean differences for the Xiangya, Wasserman, and FRIEND equations were considerably larger (209.6, −456.5, and −739.7 $mL \cdot min^{-1}$, respectively; see Table 3). Consequently, the $VO_2$peak levels observed in our participants were lower than those reported in the Wasserman and FRIEND studies.

Secondly, when comparing the absolute values of (measured $VO_2$peak—predicted $VO_2$peak)/measured $VO_2$peak > 20% among different predictive equations, the PUTH equation demonstrated the lowest percentage (33.0%), compared to 39.5, 62.9, and 81.4% for the Xiangya, Wasserman, and FRIEND equations, respectively. This trend was consistent across sex, age, and BMI subgroups (see Table 4).

Third, Bland–Altman analyses revealed that, for both males and females, the PUTH equation exhibited smaller deviations than the other equations (see Fig. 3). These findings were corroborated by the scatter plots shown in Fig. 4.

Regarding the disparities in predicting $VO_2$peak reference values in the Chinese population using different equations, we observed that the Xiangya equation underestimated these values, while the Wasserman and FRIEND equations overestimated them, with the greatest disparities found in the Wasserman and FRIEND equations (see Table 3). The primary reason for the disparities in the Xiangya equation may be its smaller sample size (964), leading to limited representativeness.

Several factors may contribute to our $VO_2$peak reference values being significantly lower than those of other countries. Firstly, this difference may be associated with the overall body type, weight, and height of the study population. Whether compared to the participants of the Wasserman and FRIEND equations in this study or to those in relevant Western studies, the Chinese population generally exhibits lower BMI, body weight, and height[14–16,37]. Secondly, differences in predicted values may also be attributed to ethnicity. Asian populations demonstrate distinct differences in skeletal muscle physiology, lung function, and chest wall anatomical structure compared to Western populations, potentially resulting in further differences in exercise capacity[38,39]. Additionally, variations in heart structure and function may also contribute to disparities in $VO_2$peak reference values. Lastly, differences in methodology of testing[15,28], participant selection bias[14,40], habitual activity levels[40], and obesity levels[15,21] may all contribute to variations in predicted $VO_2$peak. When the cohort used to develop the prediction equation includes a substantial number of individuals whose characteristics closely resemble those of the target population, the resulting prediction equation is preferable. Based on these results, we propose that the PUTH equation may outperform traditional equations in predicting $VO_2$peak reference values within the Chinese population, thereby providing a more accurate standard for assessing CRF.

## Strengths and limitations

The primary strengths of this study include its large sample size and the inclusion of specific populations, namely women, older adults, and individuals with obesity. Additionally, this study emphasizes the impact of body composition and BMI, which has not been adequately considered in previous prediction equations for the Chinese population.

This study has several limitations. First, not all participants reached $VO_2$peak, as some did not attain the plateau phase of oxygen consumption, potentially leading to an underestimation of $VO_2$peak. Second, there was an uneven distribution of participants across both age and BMI groups, with a particular lack of individuals aged 70 and older, as well as insufficient representation in the underweight and obese groups. This imbalance may limit the generalizability of the findings, especially to older adults and individuals with extreme BMI values. Finally, being a single-center study, despite nationwide recruitment, selection bias cannot be fully excluded, which may affect the external validity of the findings.

For future research, a multi-center prospective design is recommended to achieve a more balanced age and BMI distribution. We will also explore indicators like ventilation and oxygen uptake efficiency slope to better understand cardiopulmonary function and overall health.

## Conclusion

This study introduces the PUTH equation, a new $VO_2$peak prediction equation for Chinese adults. The equation shows reduced bias and superior accuracy compared to existing equations, with particular emphasis on the role of body composition (such as LBM and BSA) in predicting cardiovascular health, demonstrating broad clinical applicability potential. However, further validation through large-scale, multi-center studies is required to confirm its broader applicability.

## Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: 26 July 2024; Accepted: 12 March 2025
Published online: 29 March 2025

## References

1. Sui, X., Sarzynski, M. A., Lee, D. C. & Kokkinos, P. F. Impact of changes in cardiorespiratory fitness on hypertension, dyslipidemia and survival: An overview of the epidemiological evidence. *Prog. Cardiovasc. Dis.* **60**, 56–66. https://doi.org/10.1016/j.pcad.2017.02.006 (2017).
2. Lavie, C. J., Ozemek, C., Carbone, S., Katzmarzyk, P. T. & Blair, S. N. Sedentary behavior, exercise, and cardiovascular health. *Circ. Res.* **124**, 799–815. https://doi.org/10.1161/circresaha.118.312669 (2019).
3. Ehrman, J. K. et al. Cardiorespiratory fitness change and mortality risk among black and white patients: Henry ford exercise testing (FIT) project. *Am. J. Med.* **130**, 1177–1183. https://doi.org/10.1016/j.amjmed.2017.02.036 (2017).
4. De Schutter, A. et al. Cardiac rehabilitation fitness changes and subsequent survival. *Eur. Heart J. Qual. Care Clin. Outcomes* **4**, 173–179. https://doi.org/10.1093/ehjqcco/qcy018 (2018).
5. Franklin, B. A., Lavie, C. J., Squires, R. W. & Milani, R. V. Exercise-based cardiac rehabilitation and improvements in cardiorespiratory fitness: implications regarding patient benefit. *Mayo Clin. Proc.* **88**, 431–437. https://doi.org/10.1016/j.mayocp.2013.03.009 (2013).
6. Ross, R. et al. Importance of assessing cardiorespiratory fitness in clinical practice: A case for fitness as a clinical vital sign: A scientific statement from the american heart association. *Circulation* **134**, e653–e699. https://doi.org/10.1161/cir.0000000000000461 (2016).
7. Lavie, C. J. et al. Exercise and the cardiovascular system: clinical science and cardiovascular outcomes. *Circul. Res.* **117**, 207–219. https://doi.org/10.1161/circresaha.117.305205 (2015).
8. Guazzi, M. et al. 2016 focused update: clinical recommendations for cardiopulmonary exercise testing data assessment in specific patient populations. *Eur. Heart J.* **39**, 1144–1161. https://doi.org/10.1093/eurheartj/ehw180 (2018).
9. Herdy, A. H. et al. Cardiopulmonary exercise test: Background, applicability and interpretation. *Arqu. Brasileiros de Cardiol.* **107**, 467–481. https://doi.org/10.5935/abc.20160171 (2016).
10. Imboden, M. T. et al. Cardiorespiratory fitness and mortality in healthy men and women. *J. Am. Coll. Cardiol.* **72**, 2283–2292. https://doi.org/10.1016/j.jacc.2018.08.2166 (2018).
11. Kaminsky, L. A., Myers, J. & Arena, R. Determining cardiorespiratory fitness with precision: Compendium of findings from the FRIEND registry. *Prog. Cardiovasc. Dis.* **62**, 76–82. https://doi.org/10.1016/j.pcad.2018.10.003 (2019).
12. Vecchiato, M. et al. Comparison of cardiorespiratory fitness prediction equations and generation of new predictive model for patients with obesity. *Med. Sci. Sports Exerc.* **56**, 1732–1739. https://doi.org/10.1249/mss.0000000000003463 (2024).
13. Jurov, I., Toplišek, J. & Cvijić, M. Prediction of maximal oxygen consumption in cycle ergometry in competitive cyclists. *Life* https://doi.org/10.3390/life13010160 (2023).
14. Koch, B. et al. Reference values for cardiopulmonary exercise testing in healthy volunteers: the SHIP study. *Eur. Respir. J.* **33**, 389–397. https://doi.org/10.1183/09031936.00074208 (2009).
15. de Souza, E. S. C. G. et al. A reference equation for maximal aerobic power for treadmill and cycle ergometer exercise testing: Analysis from the FRIEND registry. *Eur. J. Prevent. Cardiol.* **25**, 742–750. https://doi.org/10.1177/2047487318763958 (2018).
16. Wasserman, K., et al. W. Principles of exercise testing and interpretation including pathophysiology and clinical aplications. (Lippincott Williams & Wilkins, 2012).
17. Königstein, K. et al. The obesity factor: How cardiorespiratory fitness is estimated more accurately in people with obesity. *Obesity (Silver Spring, Md.)* **26**, 291–298. https://doi.org/10.1002/oby.22078 (2018).
18. von Elm, E. et al. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ (Clin. Res. Ed.)* **335**, 806–808. https://doi.org/10.1136/bmj.39335.541782.AD (2007).
19. World Medical Association Declaration of Helsinki. ethical principles for medical research involving human subjects. *JAMA* **310**, 2191–2194. https://doi.org/10.1001/jama.2013.281053 (2013).
20. Balady, G. J. et al. Clinician's Guide to cardiopulmonary exercise testing in adults: a scientific statement from the American heart association. *Circulation* **122**, 191–225. https://doi.org/10.1161/CIR.0b013e3181e52e69 (2010).
21. Abe, T., Loenneke, J. P. & Thiebaud, R. S. Fat-free adipose tissue mass: impact on peak oxygen uptake (VO(2peak)) in adolescents with and without obesity. *Sports Med. (Auckland N.Z.)* **49**, 9–15. https://doi.org/10.1007/s40279-018-1020-3 (2019).
22. Wiecha, S. et al. VO(2max) prediction based on submaximal cardiorespiratory relationships and body composition in male runners and cyclists: a population study. *eLife* https://doi.org/10.7554/eLife.86291 (2023).
23. Hu, Y. M. et al. Study of formula for calculating body surface areas of the Chinese adults. *Sheng li xue bao : [Acta physiol. Sin.]* **51**, 45–48 (1999).
24. Li, J., Shang, J., Guo, B., Gong, J. & Xu, H. Establishment of prediction equations of lean body mass suitable for Chinese adults. *BioMed Res. Int.* **2019**, 1757954. https://doi.org/10.1155/2019/1757954 (2019).
25. Dun, Y. et al. Characteristics and reference values for cardiopulmonary exercise testing in the adult Chinese population—The Xiangya hospital exercise testing project (the X-ET project). *Int. J. Cardiol.* **332**, 15–21. https://doi.org/10.1016/j.ijcard.2021.03.013 (2021).

26. Zhou, B. F. 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. BES* **15**, 83–96 (2002).
27. Fletcher, G. F. et al. Exercise standards for testing and training: a scientific statement from the American heart association. *Circulation* **128**, 873–934. https://doi.org/10.1161/CIR.0b013e31829b5b44 (2013).
28. Edvardsen, E., Hansen, B. H., Holme, I. M., Dyrstad, S. M. & Anderssen, S. A. Reference values for cardiorespiratory response and fitness on the treadmill in a 20- to 85-year-old population. *Chest* **144**, 241–248. https://doi.org/10.1378/chest.12-1458 (2013).
29. Paap, D. & Takken, T. Reference values for cardiopulmonary exercise testing in healthy adults: a systematic review. *Expert Rev. Cardiovasc. Ther.* **12**, 1439–1453. https://doi.org/10.1586/14779072.2014.985657 (2014).
30. Takken, T. et al. Reference values for cardiopulmonary exercise testing in healthy subjects—An updated systematic review. *Expert Rev. Cardiovasc. Ther.* **17**, 413–426. https://doi.org/10.1080/14779072.2019.1627874 (2019).
31. Williams, D. R. & Sternthal, M. Understanding racial-ethnic disparities in health: sociological contributions. *J. Health Soc. Behav.* **51**, S15-27. https://doi.org/10.1177/0022146510383838 (2010).
32. Wasserman, K. Coupling of external to cellular respiration during exercise: the wisdom of the body revisited. *Am. J. Physiol.* **266**, E519-539. https://doi.org/10.1152/ajpendo.1994.266.4.E519 (1994).
33. Sun, X. G. New theory of holistic integrative physiology and medicine. I: New insight of mechanism of control and regulation of breathing. *Zhongguo ying yong sheng li xue za zhi Zhongguo yingyong shenglixue zazhi Chin. J. Appl. Physiol.* **31**, 295–301 (2015).
34. Yan, R. et al. Ethnic differences in spirometry measurements in China: Results from a large community-based epidemiological study. *Respirology (Carlton, Vic.)* **23**, 704–713. https://doi.org/10.1111/resp.13258 (2018).
35. Ong, K. C. et al. Predictive values for cardiopulmonary exercise testing in sedentary Chinese adults. *Respirology (Carlton, Vic.)* **7**, 225–231. https://doi.org/10.1046/j.1440-1843.2002.00393.x (2002).
36. Wang, Y. et al. Normal references of peak oxygen uptake for cardiorespiratory fitness measured with cardiopulmonary exercise testing in chinese adults. *J. Clin. Med.* https://doi.org/10.3390/jcm11164904 (2022).
37. Kaminsky, L. A. et al. Updated reference standards for cardiorespiratory fitness measured with cardiopulmonary exercise testing: data from the fitness registry and the importance of exercise national database (FRIEND). *Mayo Clin. Proc.* **97**, 285–293. https://doi.org/10.1016/j.mayocp.2021.08.020 (2022).
38. Ceaser, T. & Hunter, G. Black and white race differences in aerobic capacity, muscle fiber type, and their influence on metabolic processes. *Sports Med. (Auckland, N.Z.)* **45**, 615–623. https://doi.org/10.1007/s40279-015-0318-7 (2015).
39. Korotzer, B., Ong, S. & Hansen, J. E. Ethnic differences in pulmonary function in healthy nonsmoking Asian-Americans and European-Americans. *Am. J. Respir. Crit. Care Med.* **161**, 1101–1108. https://doi.org/10.1164/ajrccm.161.4.9902063 (2000).
40. Almeida, A. E. et al. An equation for the prediction of oxygen consumption in a Brazilian population. *Arqu. Brasileiros de Cardiol.* **103**, 299–307. https://doi.org/10.5935/abc.20140137 (2014).


## Author contributions

JN. W wrote the main manuscript text; YD. T, W. G, Sl. X and W. Z revised the manuscript; YX. S, DD. X, KX. W and L. Y performed the CPET and collected the data; T. S and L. X prepared tables and figures; YF. Y and C. R calculated and examined the data. All authors reviewed the manuscript.


## Funding

This work was supported by the "Comprehensive assessment of exercise safety, exercise prescription design and application of monitoring equipment in elderly population (2020YB02)", "Clinical key project of the Peking University Third Hospital of Peking University (BYSYZD2021008)" and "Youth incubation fund of the Peking University Third Hospital of Peking University (BYSYFY2021013)".


## Declarations

## Competing interests

The authors declare no competing interests.

## Additional information


**Correspondence** and requests for materials should be addressed to W.Z.