Validation of the World Health Organization non-laboratory-based cardiovascular disease risk prediction models in ten diverse regions of China

Songchun Yang [1,2], Yinqi Ding [1], Canqing Yu [1,3], Yu Guo [4], Yuanjie Pang [1], Dianjianyi Sun [1,3], Pei Pei [3], Ling Yang [5,6], Yiping Chen [5,6], Huaidong Du [5,6], Dan Schmidt [6], Rebecca Stevens [6], Derrick Bennett [5,6], Robert Clarke [6], Junshi Chen [7], Zhengming Chen [6], Liming Li [1,3], Jun Lv [1,3]* on behalf of the China Kadoorie Biobank Collaborative Group [†]

[1] Department of Epidemiology & Biostatistics, School of Public Health, Peking University, Beijing 100191, China;

[2] Department of Dermatology, Xiangya Hospital, Central South University, Changsha 410008, China;

[3] Peking University Center for Public Health and Epidemic Preparedness & Response, Beijing 100191, China;

[4] Fuwai Hospital Chinese Academy of Medical Sciences, Beijing 100730, China;

[5] Medical Research Council Population Health Research Unit at the University of Oxford, Oxford OX3 7LF, United Kingdom;

[6] Clinical Trial Service Unit & Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, University of Oxford, Oxford OX3 7LF, United Kingdom;

[7] China National Center for Food Safety Risk Assessment, Beijing 100738, China;

[†] The members of the steering committee and collaborative group are listed in *Supplementary Materials*.

*** Correspondence to**

Jun Lv, MD, PhD

Department of Epidemiology and Biostatistics,

Peking University Health Science Center

38 Xueyuan Road, Beijing 100191, China

Phone: 86-10-82801528 (ext. 322)

Email: lvjun@bjmu.edu.cn

Word count: Abstract 304, Main text 2999

(3 figures, 2 tables, and 1 supplementary material)

12 Nov 2022

Abstract

Objective: To perform an external validation of the World Health Organization (WHO) non-laboratory-based cardiovascular disease (CVD) risk prediction models (hereafter, the WHO model) in ten diverse regions of China.

Methods: The China Kadoorie Biobank (CKB), an ongoing prospective study, recruited >512,000 adults aged 30 to 79 from 10 diverse regions (5 urban, 5 rural) in China. We performed an external validation of the WHO model for East Asia in CKB. We recalculated recalibration parameters for the original WHO model in each study region and evaluated the predictive performances before and after recalibration.

Findings: A total of 412,225 participants aged 40–79 years were included in the present study. During a median follow-up of 11 years, 58,035 and 41,262 incident CVD events were recorded in women and men, respectively. The Harrell's C index of the WHO model for East Asia was 0.682 (95% CI: 0.655–0.710) in women and 0.700 (95% CI: 0.681–0.719) in men but varied between regions. After recalibration in each study region, the discrimination was greatly improved in the overall population. The C index increased from 0.674 (95% CI: 0.672–0.677) to 0.749 (95% CI: 0.746–0.751) in women and from 0.698 (95% CI: 0.695–0.701) to 0.753 (95% CI: 0.750–0.755) in men. The original WHO model underestimated the 10-year CVD risk in most regions, which was greatly improved after separate recalibration in each study region. The ratios of predicted to observed events before and after recalibration in the overall population were 0.189 and 1.027 in women and 0.543 and 1.089 in men.

Conclusion: The WHO model for East Asia yielded only moderate discrimination for CVD in the Chinese population and had limited prediction for CVD risk in different regions in China. Recalibration for diverse regions greatly improved discrimination in the overall study population and calibration in each study region.

Abbreviations and acronyms

CAD: coronary artery disease

CI: confidence interval

CKB: China Kadoorie Biobank

CVD: cardiovascular diseases

HS: hemorrhagic stroke

IS: ischemic stroke

WHO: World Health Organization

Introduction

Cardiovascular diseases (CVD), including coronary artery disease (CAD) and stroke, are the leading causes of death and disability worldwide.[1] Risk prediction models are important tools for identifying high-risk individuals for early primary prevention of CVD.[2-6] Before a newly developed model is popularized and applied, external validation studies are required to evaluate the predictive performances of the model in the target population. Discrimination and calibration of a model are two important metrics. The discrimination indicates the ability of a model to distinguish between people who did and did not develop the event of interest. The calibration indicates the agreement between observed event risks and event risks predicted by the model.[7] When the calibration performance of a model is not good, recalibration is usually required for better application to the target population.[8]

The World Health Organization (WHO) has developed new models to estimate CVD risk for people aged 40-80 years in 21 Global Burden of Disease regions, including laboratory-based and non-laboratory-based models.[9] The non-laboratory-based model (hereafter "the WHO model") is more applicable in resource-limited regions where blood-based biomarkers, such as lipids, are not widely available for all individuals. Although the WHO model for East Asia has been recommended for CVD risk prediction in China, it does not take into account significant differences in the spatial patterns of incidence, prevalence, and mortality of CVD overall as well as its main subtypes, and the prevalence of major contributing risk factors in China.[10-12] An external validation study for the WHO model has been conducted in the China-PAR (Prediction for Atherosclerotic Cardiovascular Disease Risk in China) cohort. This study ignored different regions of China and reported that the WHO model overestimated the observed CVD risk.[13] However, the predictive performances of the WHO model by regions were not evaluated.

The present study was based on the China Kadoorie Biobank (CKB) and had two

aims: (1) to perform an external validation of the WHO model for East Asia in 10 diverse regions of China; (2) to consider the significant regional differences in the incidence of CVD subtypes in China and compare the predictive performance of the WHO model before and after separate recalibration in 10 regions.

Methods

Study population

CKB is an ongoing prospective study with 512,725 participants aged 30 to 79 enrolled from 10 diverse regions (5 urban, 5 rural) in China from 2004 to 2008 (**Table S1**). Details of the study have been described elsewhere.[14, 15] Briefly, the baseline questionnaire collected information on sociodemographic characteristics, lifestyle behaviors, dietary habits, personal health (including self-reported histories of CAD, stroke, and transient ischemic attacks), and family medical history. A 10-ml random blood sample was collected for each participant with the time of the last meal recorded. CKB obtained ethical approval from the Ethical Review Committee of the Chinese Center for Disease Control and Prevention (Beijing, China) and the Oxford Tropical Research Ethics Committee, University of Oxford (UK). All participants provided a written informed consent form. After excluding participants who were younger than 40 years old (n=77,623), those who had been diagnosed with CAD (n=15,286) or stroke or transient ischemic attack (n=7,590), or had missing body mass index (BMI; n=1) at the baseline survey. A total of 412,225 participants were included in the present analysis (**Fig. 1A**).

Assessment of predictors

Variables used in this study included sex, age, smoking status, systolic blood pressure (SBP), and BMI, which are predictors in the WHO model.[9] Details on the collection and definition of each predictor have been described in our previous work.[16]

Ascertainment of outcomes

All participants were followed up for incident disease outcomes since their baseline enrolment. Incident events were identified by using linkage with local disease and death registries and the national health insurance system, and supplemented by active follow-up.[14] More than 97% of participants were linked to the health insurance

system. Only 4,009 participants (0.97%) were lost to follow-up before censoring on December 31, 2017. We used information from underlying and multiple causes of death and primary and secondary discharge diagnoses. Trained staff of the CKB research team who were blinded to baseline information coded all events using the International Classification of Diseases, Tenth Revision (ICD-10). The medical records of the incident events (i.e., the first CVD event) were retrieved and reviewed by specialist physician adjudicators.[17] By October 2018, of the retrieved medical records of 33,515 CAD cases (I20-I25), 34,758 IS cases (I63), and 5023 HS cases (I60-I61), the confirmed rates of the diagnosis were 87.9%, 91.5%, and 80.4%, respectively.

Since the definitions of disease outcomes of the WHO model were slightly different in the derivation and recalibration processes,[1, 9] both definitions were used in the present study (**Table S2**). In the first component of the study (**Fig. 1B**), the outcome definition was consistent with the Global Burden of Disease (GBD) Study in 2017, one of the data sources to calculate the recalibration parameters for the WHO model.[1] In the second component (**Fig. 1B**), the definition was consistent with that in the derivation process of the WHO model.[9] Only the first CVD event during follow-up was included unless otherwise specified. If a participant was recorded with a CAD event and a stroke event successively or simultaneously, the date of the first of these two events was used in the analysis of CVD as a whole. When CAD or stroke was analyzed as different outcomes, the dates of the first CAD event and first stroke event were considered separately.

Statistical analysis

All analyses were conducted separately for women and men. In the first component, We recalibrated the original WHO model according to the latest recalibration parameters (in 2017) applicable to East Asia (**Fig. 1B**, **Supplementary Methods, Appendix 1, Table S3**).[9] In the second component, we recalibrated the original WHO model (i.e., the CAD model and the stroke model) in 10 study regions (**Fig. 1B**,

Supplementary Methods, Table S4). We evaluated the discrimination and calibration performance of the WHO model before and after recalibration across the overall study population and in each study region separately (**Supplementary Methods**).

The following sensitivity analyses were conducted. First, we used recalibration parameters derived in 2005, 2010, and 2015 to recalibrate the original WHO model (**Table S3**). Second, due to the higher incidence of stroke in China compared with Western countries, applying the disease outcome definition used by the derivation process of the WHO model could lead to an artificially low proportion of CAD in total CVD. Therefore, the outcome definition of the CKB model was adopted instead in the second component of this study (**Table S2**).

The study adhered to the TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) statement for reporting (**Supplementary Materials**).[18] Analyses were conducted with Stata 17.0. Figures were produced using R 3.6.0.

Results

Characteristics of the study population

A total of 412,225 participants from 10 different regions of China were included in the current study, 58.6% were female and 56.0% were rural residents. The mean age was 54.3 (SD=9.2) years (**Table 1**). There was substantial heterogeneity between levels of CVD risk factors between the 10 study regions. For example, the age-adjusted proportion of daily smokers was 0.2% among women and 39.2% among men in Meilan, while the corresponding proportion was 10.2% and 64.7% in Pengzhou (**Table S5**). During a median follow-up of 11 years, 58,035 and 41,262 incident CVD events defined according to WHO recalibration criteria were recorded among women and men, respectively. The number of hard CAD (nonfatal I21-I23 and fatal I21-I25) events was far fewer than that of soft CAD (any I20-I25) in the current study population (**Table 1**). There was significant heterogeneity in the 10-year risks of CVD and its subtypes between the 10 study regions. For example, the 10-year risk of broad stroke (any I60-I69) was over 30% in both sexes in Nangang, but lower than 10% in Wuzhong (**Fig. S1**).

External validation of the WHO model for East Asia

In external validation, Harrell's C index of the WHO model for East Asia was 0.682 (95% CI: 0.655–0.710) among women, with substantial variation between regions. The C index was lowest in Nangang (C=0.642, 95% CI: 0.637–0.647) and highest in Wuzhong (C=0.763, 95% CI: 0.753–0.772). C indexes among men were similar to those among women (**Fig. 2**). As for calibration, the WHO model for East Asia underestimated the 10-year risk of CVD for the overall population and all study regions except Wuzhong and Tongxiang (**Fig. S2**). After recalibrating the WHO model with recalibration parameters derived in different years, the discrimination performance did not materially change (**Table S6**), and the underestimation of the 10-year risk of CVD persisted (**Fig. S3**).

Recalibrating the WHO model in 10 CKB study regions

After recalibrating the WHO model in each study region, the discrimination performance of the WHO model was greatly improved in the overall study population. Among women, the C index increased by 0.075 from 0.674 (95% CI: 0.672–0.677) to 0.749 (95% CI: 0.746–0.751). Among men, the increment of the C index was 0.055 (**Fig. 3A**). However, recalibration had a minimal effect on the discrimination performance in each study region (**Table 2**). The calibration performance of the recalibrated WHO model was close to unity in the overall study population (**Fig. 3B**) and 10 study regions (**Fig. S4**). The discrimination performance of the original WHO model in older people (≥65 years), individuals with hypertension, diabetes, low education level, and low household income were all improved after recalibration (**Table S7**). The recalibrated WHO model was well-calibrated in older people and those with hypertension, but slightly underestimated the risk of CVD in people with diabetes (**Fig. S5**). When the outcome definition of the CKB model was adopted instead, the discrimination and calibration performance was improved in the overall study population (**Fig. S6**). However, the recalibrated model still slightly underestimated the 10-year risk of CVD in participants with diabetes (**Fig. S7**).

Discussion

In the present study, we conducted an external validation of the WHO model for East Asia in 10 diverse regions of China. The overall discrimination of the WHO model was moderate and the 10-year CVD risk of the CKB participants was underestimated in most regions. After recalibration of the WHO model in each study region, the discrimination and calibration performances of the model were greatly improved in the overall study population.

The pooled Harrell's C of the WHO model for East Asia was only about 0.7 in both sexes, lower than previous studies conducted in the Chinese populations. In an external validation study based on the Asia Pacific Cohorts Studies Collaboration (APCSC) and the China Multi-Provincial Cohort Study (CMCS), the pooled C index of the non-laboratory-based WHO model for East Asia was 0.741 (95% CI: 0.725–0.757).[9] When applying the WHO model in the China-PAR cohort, the C index was 0.754 (95% CI: 0.731–0.777) in women and 0.762 (95% CI: 0.744–0.781) in men.[13] Differences in the definition of outcome and study population could have influenced the discrimination. In the present study, we adopted an identical definition as that used in the recalibration process of the WHO model. This definition includes nonfatal angina (I20) for CAD and other cerebrovascular diseases (I65-I69) for stroke,[1] and is broader than that adopted by APCSC, CMCS, and China-PAR.[9, 13] These differences could partly explain the overestimation of CVD risk in the external validation study of China-PAR.[13]

The WHO model for East Asia underestimated the CVD risk to a variable extent in most study regions. Separate recalibration of the WHO model in each region achieved almost ideal calibration. The findings suggest a universal model is unsuitable for direct application to different regions in China due to the significant regional differences in the incidence of CVD subtypes. Models need to be localized according to the local prevalence of CVD risk factors and disease incidence rates

before being applied to a specific region. The study based on China-PAR did not evaluate the calibration performance of the WHO model by region, thus making the conclusions less reliable.[13] The current participants of each study region came from a relatively small geographical area in China. It is not feasible to update the model across the country according to the current regional size. A possible approach is to update the model in a larger area (such as a province) first, and then to update the model in smaller geographical areas. External validation studies conducted in other regions of China are warranted to examine our findings.

Recalibration significantly improved the discrimination of the WHO model in the overall population, highlighting the significance of recalibration in different regions of China. Recalibration by region is equivalent to adding the region as a predictor. Due to the significant differences in the spatial patterns of incidence of CVD and the prevalence of major CVD risk factors in China,[10-12] we observed a significant improvement in discrimination in the overall population. However, recalibration had little impact on discrimination of each study region. Since recalibration changed the absolute risk but not the order of absolute risk for each participant,[19] both the discrimination of the CAD and the stroke submodel remained unchanged in each study region. Therefore, the discrimination of the total CVD model was not significantly affected.

The differences in discrimination of the WHO model among 10 study regions persisted after recalibration and could not be explained by the aforementioned spatial patterns. China is the largest developing country, where CVD risk factors might be more complex than those in developed countries. In addition to well-established risk factors, other potential risk factors, such as environmental hazards in the home, work, and broader outdoor environment, might also influence the incidence of CVD. These risk factors were not included in the current model and might affect the discrimination of the model to varying degrees in different regions. The discrimination of the WHO model was not good (C index < 0.7) in some study regions, such as Nangang,

Liuyang, and Huixian. For these regions, on the one hand, there might be some specific risk factors required to be uncovered. On the other hand, other known CVD risk factors might help improve risk prediction. Specifically, the current model could be used to screen a subgroup of people with a relatively high risk of CVD; subsequently, other CVD risk factors could be evaluated in the selected populations. Among non-laboratory-based indicators, previous studies have found waist circumference was a better predictor than BMI, and antihypertensive treatment improved risk prediction.[16, 20] Diastolic blood pressure, level of education and waist-hip ratio in men, and physical activity in women could improve the risk prediction of hemorrhagic stroke in CKB.[16] Other physical examination indicators such as ankle-brachial index and arterial stiffness, psychosocial and work stress, and environmental exposure are also expected to improve risk prediction,[3, 4] however, these indicators are not easily available and measurable, limiting the possible application.

The present study adopted different outcome definitions in the recalibration process of the original WHO model. The WHO model exhibited robustness in calibration performance and approached the ideal level regardless of the outcome definition. The findings suggest the flexibility of the recalibration method. Outcome definitions adopted in the application process could have differed from that used in the model derivation process. The main factor affecting the calibration performance after recalibration is more likely to be the reliability of the data source used to generate the recalibration parameters. However, different outcome definitions affect the interpretation of the model. The ratio of incidence rates of the stroke to CAD is higher in China than in Western countries.[9] In the present population, the WHO model mainly predicts the risk of stroke when adopting the outcome definition in the derivation dataset of the WHO model, which adopted a narrower definition for CAD and a broader definition for stroke than the CKB definition. In addition, the hard endpoint for CAD is well-defined and measured consistently across studies, however, this narrower definition might underestimate the overall CAD burden. Currently, there

is no consensus on the definition of the disease outcomes for use in CVD risk prediction models. Future studies are required to determine the most appropriate outcome definition by comprehensively considering the health, economic, and social benefits.

The present study is the largest external validation study of the WHO model in the Chinese population. CKB recruited participants from 5 urban and 5 rural regions of China, providing good coverage of regions with different burdens of CVD subtypes. All 10 study regions adopted identical procedures and standardized protocols, allowing comparison and pooling of results from the different regions. Less than 1% of CKB participants were lost after an average of 11 years of follow-up.

Several limitations merit consideration. First, we were unable to validate the laboratory-based WHO model since the blood lipid information was only available in a subset of participants. Previous studies have suggested the laboratory-based and non-laboratory-based WHO models have similar predictive performances.[9, 13, 21] However, the non-laboratory-based WHO model is much more likely to be used in resource-limited regions. Second, the recalibrated WHO model for each study region should be considered a new model. External validation studies are warranted before application. Third, like most large-scale cohorts, the participants recruited at baseline were almost volunteers. However, the selection bias caused by the loss of follow-up is very small in CKB. Fourth, participants with low socioeconomic status might have delayed hospital visits. The current analyses included only inpatient events, mainly corresponding to more severe conditions and narrowing, to some extent, the difference in hospital visits among groups with different socioeconomic statuses. In addition, recalibration significantly improved the discrimination among participants with low socioeconomic status. Therefore, this should not affect our conclusion in the second component.

Based on the large population-based cohort of Chinese adults, we validated the non-laboratory-based WHO model for East Asia and found it was not directly

applicable to different regions of China. The WHO model needs to be localized before being used in a specific region in China. In the future, it is necessary to establish basic parameter databases required for model recalibration and model evaluation in the real world.

Supplementary Materials

Members of the China Kadoorie Biobank collaborative group

Supplementary Methods; Figure S1-S7; Table S1-S7; eReferences[1, 9, 16, 22-25]

Appendix 1; TRIPOD Checklist

Contributors

JL conceived and designed the study. LL, ZC, and JC designed and supervised the whole study, obtained funding, and, together with CY, YG, PP, LY, YC, HD, D.Schmidt, and RS acquired the data. SY and YD analyzed the data. SY drafted the manuscript. YH, CY, YP, and D.Sun helped to interpret the results. JL, DB, and RC contributed to the critical revision of the manuscript for important intellectual content. All authors reviewed and approved the final manuscript. JL is the guarantor.

Acknowledgment

The most important acknowledgment is to the participants in the study and the members of the survey teams in each of the 10 regional centers, as well as to the project development and management teams based at Beijing, Oxford, and the 10 regional centers.

Study funding

This work was supported by National Natural Science Foundation of China (82192904, 82192901, 82192900, 91846303). The CKB baseline survey and the first re-survey were supported by a grant from the Kadoorie Charitable Foundation in Hong Kong. The long-term follow-up is supported by grants from the UK Wellcome Trust (212946/Z/18/Z, 202922/Z/16/Z, 104085/Z/14/Z, 088158/Z/09/Z), grants (2016YFC0900500) from the National Key R&D Program of China, National Natural Science Foundation of China (81390540), and Chinese Ministry of Science and Technology (2011BAI09B01). Zhengming Chen was supported by the UK Medical Research Council. The funders had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

Disclosure

We report no disclosures relevant to the manuscript.

References

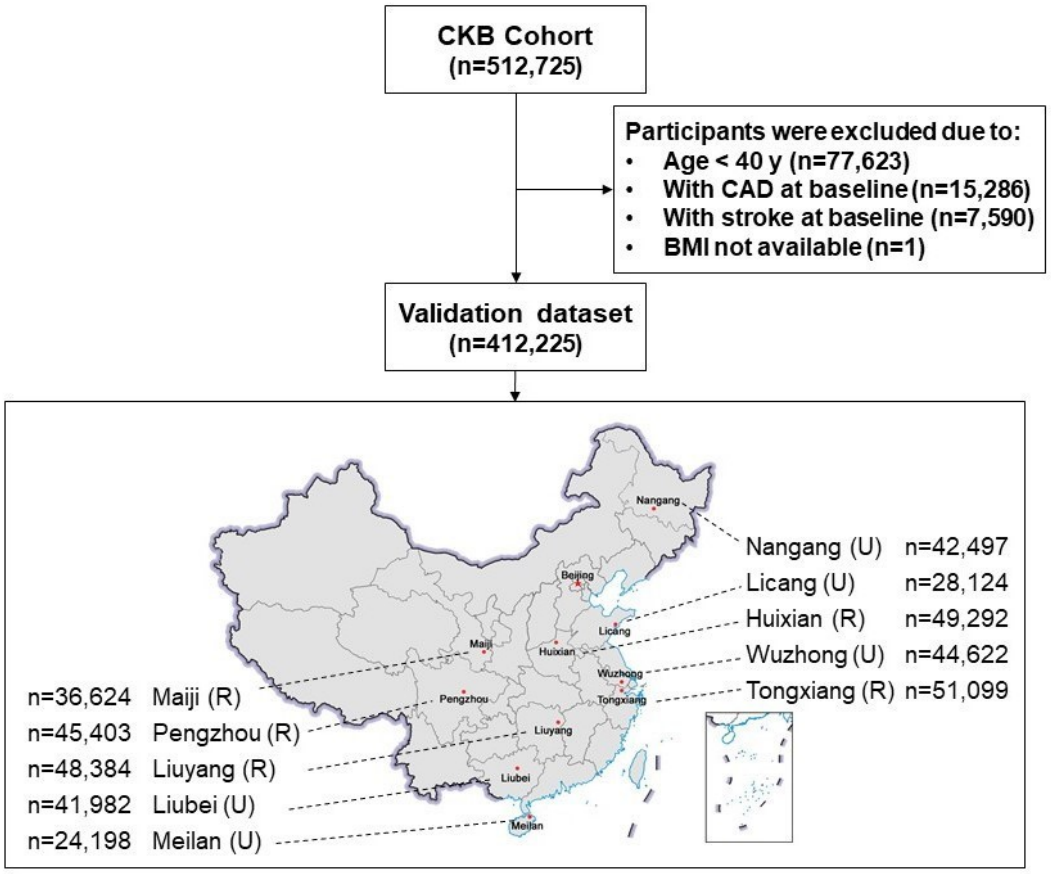
1. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2018;392(10159):1789-858. Epub 2018/11/30.
2. World Health Organization. HEARTS technical package for cardiovascular disease management in primary health care: risk based CVD management. Geneva: World Health Organization; 2020.
3. Arnett DK, Blumenthal RS, Albert MA, Buroker AB, Goldberger ZD, Hahn EJ, et al. 2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol. 2019;74(10):1376-414. Epub 2019/03/22.
4. Visseren FLJ, Mach F, Smulders YM, Carballo D, Koskinas KC, Back M, et al. 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice. Eur Heart J. 2021;42(34):3227-337. Epub 2021/08/31.
5. Chinese Society of Cardiology of Chinese Medical Association, Cardiovascular Disease Prevention and Rehabilitation Committee of Chinese Association of Rehabilitation Medicine, Cardiovascular Disease Committee of Chinese Association of Gerontology and Geriatrics, Thrombosis Prevention and Treatment Committee of Chinese Medical Doctor Association. Chinese guideline on the primary prevention of cardiovascular diseases. Chin J Cardiol. 2020;48(12):1000-38. Epub 2020/12/29.
6. The Joint Task Force for Guideline on the Assessment and Management of Cardiovascular Risk in China. Guideline on the assessment and management of cardiovascular risk in China. Chin J Prev Med. 2019;53(1):13-35.
7. Damen JA, Pajouheshnia R, Heus P, Moons KGM, Reitsma JB, Scholten R, et al. Performance of the Framingham risk models and pooled cohort equations for predicting 10-year risk of cardiovascular disease: a systematic review and meta-analysis. BMC Med. 2019;17(1):109. Epub 2019/06/14.
8. Pennells L, Kaptoge S, Wood A, Sweeting M, Zhao X, White I, et al. Equalization of four cardiovascular risk algorithms after systematic recalibration: individual-participant meta-analysis of 86 prospective studies. Eur Heart J. 2019;40(7):621-31. Epub 2018/11/27.
9. The WHO CVD Risk Chart Working Group. World Health Organization cardiovascular disease risk charts: revised models to estimate risk in 21 global regions. Lancet Glob Health. 2019;7(10):e1332-e45. Epub 2019/09/07.
10. Liu S, Li Y, Zeng X, Wang H, Yin P, Wang L, et al. Burden of Cardiovascular Diseases in China, 1990-2016: Findings From the 2016 Global Burden of Disease Study. JAMA Cardiol. 2019;4(4):342-52. Epub 2019/03/14.

11. Zhao D, Liu J, Wang M, Zhang X, Zhou M. Epidemiology of cardiovascular disease in China: current features and implications. Nat Rev Cardiol. 2019;16(4):203-12. Epub 2018/11/24.
12. Li X, Wu C, Lu J, Chen B, Li Y, Yang Y, et al. Cardiovascular risk factors in China: a nationwide population-based cohort study. Lancet Public Health. 2020;5(12):e672-e81. Epub 2020/12/04.
13. Li J, Liu F, Yang X, Cao J, Chen S, Chen J, et al. Validating World Health Organization cardiovascular disease risk charts and optimizing risk assessment in China. Lancet Reg Health West Pac. 2021;8:100096.
14. Chen Z, Chen J, Collins R, Guo Y, Peto R, Wu F, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. Int J Epidemiol. 2011;40(6):1652-66. Epub 2011/12/14.
15. Chen Z, Lee L, Chen J, Collins R, Wu F, Guo Y, et al. Cohort profile: the Kadoorie Study of Chronic Disease in China (KSCDC). Int J Epidemiol. 2005;34(6):1243-9. Epub 2005/09/01.
16. Yang S, Han Y, Yu C, Guo Y, Pang Y, Sun D, et al. Development of a model to predict 10-year risk of ischemic and hemorrhagic stroke and ischemic heart disease using the China Kadoorie Biobank. Neurology. 2022;98(23):e2307-e17.
17. Chen Y, Clarke R. Verification and Adjudication of Health Outcomes in Prospective Cohort Studies. In: Chen Z, editor. Population Biobank Studies: A Practical Guide. Singapore: Springer Singapore; 2020. p. 123-43.
18. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. Ann Intern Med. 2015;162(1):55-63. Epub 2015/01/07.
19. Su TL, Jaki T, Hickey GL, Buchan I, Sperrin M. A review of statistical updating methods for clinical prediction models. Stat Methods Med Res. 2018;27(1):185-97. Epub 2016/07/28.
20. Yang X, Li J, Hu D, Chen J, Li Y, Huang J, et al. Predicting the 10-Year Risks of Atherosclerotic Cardiovascular Disease in Chinese Population: The China-PAR Project (Prediction for ASCVD Risk in China). Circulation. 2016;134(19):1430-40. Epub 2016/09/30.
21. Rezaei F, Seif M, Gandomkar A, Fattahi MR, Malekzadeh F, Sepanlou SG, et al. Comparison of laboratory-based and non-laboratory-based WHO cardiovascular disease risk charts: a population-based study. J Transl Med. 2022;20(1):133. Epub 2022/03/18.
22. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc. 1958;53:457-81.
23. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med. 1996;15(4):361-87. Epub 1996/02/28.
24. Cook NR, Ridker PM. Calibration of the Pooled Cohort Equations for Atherosclerotic Cardiovascular Disease: An Update. Ann Intern Med.

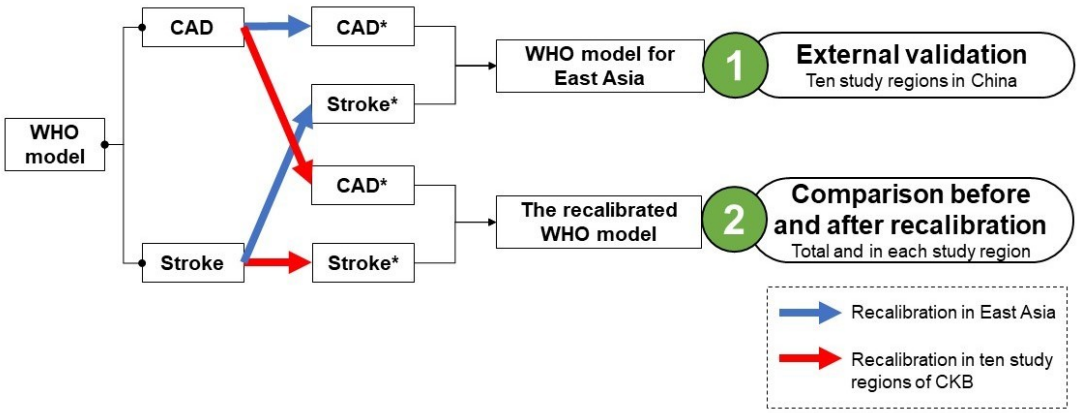
2016;165(11):786-94.
25. D'Agostino RB, Nam B-H. Evaluation of the Performance of Survival Analysis Models: Discrimination and Calibration Measures. Handbook of Statistics: Elsevier; 2003. p. 1-25.

Fig. 1. Study overview

(A) Flowchart for the present study



(B) Study design



BMI: body mass index; CAD: coronary artery disease; CKB: China Kadoorie Biobank; CVD: cardiovascular disease; HS: hemorrhagic stroke; IS: ischemic stroke; WHO: World Health Organization.

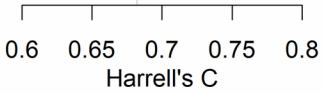
(A) Flowchart for the present study. "U" represents the urban region, and "R" represents the rural region. For detailed geographic information on each study region, please see **Table S1**.

(B) Study design. "*" represents the model after recalibration. The WHO model represents the WHO non-laboratory-based CVD risk prediction model. The detailed method for recalibration was described in the **Supplementary Methods**.

Fig. 2. Harrell's C for the WHO model for East Asia in the 10 study regions

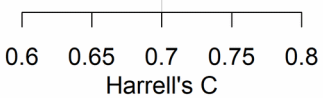
(A) Women

Region	Cases	Participants	Harrell's C (95% CI)
Nangang (U)	12,374	25,380	0.642 (0.637, 0.647)
Liuyang (R)	8,783	26,485	0.648 (0.642, 0.654)
Huixian (R)	9,446	27,887	0.659 (0.653, 0.664)
Meilan (U)	3,363	15,334	0.688 (0.679, 0.696)
Pengzhou (R)	5,283	27,626	0.699 (0.692, 0.706)
Maiji (R)	4,594	21,594	0.702 (0.694, 0.710)
Licang (U)	4,022	15,794	0.708 (0.701, 0.716)
Liubei (U)	5,081	26,147	0.710 (0.703, 0.717)
Tongxiang (R)	2,723	29,563	0.750 (0.741, 0.758)
Wuzhong (U)	2,366	25,746	0.763 (0.753, 0.772)
Combined (meta)	**58,035**	**241,556**	**0.682 (0.655, 0.710)**
Total	58,035	241,556	0.664 (0.662, 0.666)



(B) Men

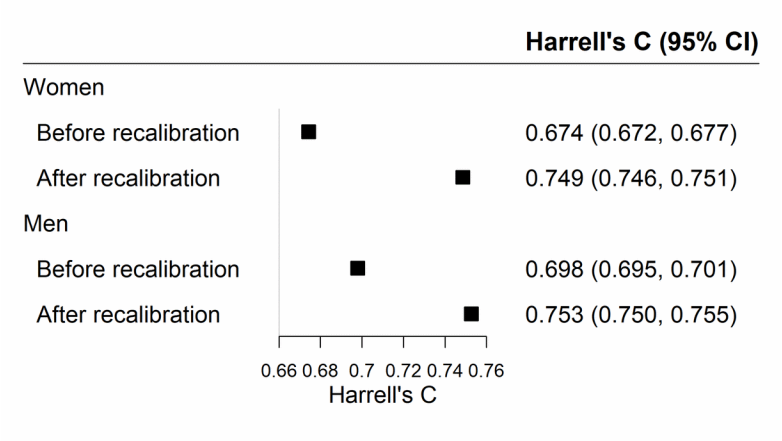
Region	Cases	Participants	Harrell's C (95% CI)
Nangang (U)	7,887	17,117	0.667 (0.661, 0.673)
Huixian (R)	6,815	21,405	0.680 (0.673, 0.686)
Liuyang (R)	7,110	21,899	0.681 (0.675, 0.688)
Maiji (R)	3,216	15,030	0.710 (0.701, 0.719)
Meilan (U)	2,131	8,864	0.716 (0.706, 0.726)
Liubei (U)	3,802	15,835	0.718 (0.710, 0.726)
Licang (U)	2,671	12,330	0.723 (0.713, 0.732)
Pengzhou (R)	3,125	17,777	0.725 (0.716, 0.734)
Wuzhong (U)	1,919	18,876	0.736 (0.725, 0.747)
Tongxiang (R)	2,586	21,536	0.745 (0.735, 0.754)
Combined (meta)	**41,262**	**170,669**	**0.700 (0.681, 0.719)**
Total	41,262	170,669	0.690 (0.687, 0.692)



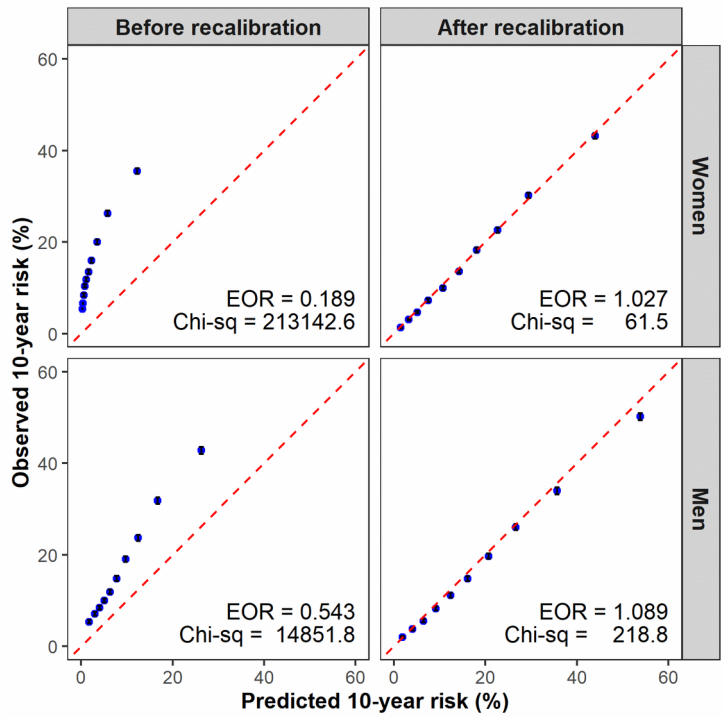
CAD: Coronary artery disease; CI: confidence interval; CVD: cardiovascular disease; WHO: World Health Organization. "U" represents the urban study region, and "R" represents the rural study region. CVD events were defined according to the definition used by the recalibration process of the WHO model (see definition 1 in **Table S2**). The recalibration parameters applied to China in 2017 were used for the recalibration of the WHO model (**Supplementary Methods, Table S3**). The C index estimates the probability of the model correctly predicting who will have a CVD event first in a randomly selected pair of participants. The value of this index is between 0.5 and 1. Generally, a C index above 0.7 indicates a good prediction model. The C index of the "total" row was calculated based on the whole sample ignoring the region. The Combined C index was the sum of the C index of each study region weighted by inverse variance. The confidence interval of the combined C index was calculated by using the *t* distribution with 9 degrees of freedom.

Fig. 3. Predictive performance of the WHO model before and after recalibration in the overall study population

(A) Discrimination performance



(B) Calibration performance



CI: confidence interval; CVD: cardiovascular disease; EOR: ratios of predicted to observed events; WHO: World Health Organization.

CVD events were defined according to the definition used by the derivation process of the WHO model (see definition 2 in **Table S2**). The WHO model was recalibrated in each study region (**Supplementary Methods, Table S4**). The 95% CIs in both figures were too narrow to display clearly. **(A) Discrimination performance**. The C index estimates the probability of the model correctly predicting who will have a CVD event first in a randomly selected pair of participants. The value of this index is between 0.5 and 1. Generally, a C index above 0.7 indicates a good prediction model. **(B) Calibration performance**. The closer the points are to the diagonal line (i.e., the 45° line), the better the calibration performance is. When the point is above the red line, it indicates that the model underestimates the actual risk; conversely, it indicates that the model overestimates the actual risk. A ratio of 1.0 indicates perfect calibration; ratios greater than 1.0 indicate overestimation, and those less than 1.0 indicate underestimation. The Chi-sq represents the Nam-D'Agostino test chi-square with nine degrees of freedom. A smaller chi-square value represents a better calibration performance.

Table 1. Selected characteristics of the CKB study participants by sex.

	Women	Men
Baseline characteristics		
Total participants	241,556	170,669
Rural area	133,155 (55.1)	97,647 (57.2)
Age, years	52.8 (46.2-59.9)	54.0 (47.4-62.0)
Primary school and below	145,314 (60.2)	78,040 (45.7)
Household income < 10,000 RMB/year	71,431 (29.6)	45,286 (26.5)
Current daily smoker	5,602 (2.3)	97,132 (56.9)
Use of blood pressure-lowering treatment	33,332 (13.8)	19,800 (11.6)
Systolic blood pressure, mmHg	128.5 (116.0-144.0)	130.5 (119.5-144.5)
Diastolic blood pressure, mmHg	76.5 (70.0-84.0)	79.0 (71.5-86.5)
Body mass index, kg/m^2	23.7 (21.5-26.1)	23.1 (21.0-25.5)
Waist circumference, cm	79.0 (72.8-85.6)	81.3 (74.3-88.5)
Self-reported diabetes	8,128 (3.4)	4,876 (2.9)
All diabetes [a]	15,700 (6.5)	9,732 (5.7)
Follow-up period [b]		
Total person-years observed	2,642,595	1,805,894
Follow-up time, years	11.1 (10.2-12.1)	11.0 (10.0-12.0)
Participants with follow-up≥10 years	198,767 (82.3)	130,806 (76.6)
Cardiovascular outcomes [c]		
WHO recalibration definition		
CAD	24,403	16,696
Stroke	43,351	30,778
Total	58,035	41,262
WHO derivation definition		
CAD	3,835	4,689
Stroke	40,171	29,241
Total	42,919	32,658
CKB derivation definition		
CAD	24,403	16,696
Stroke	27,222	22,818
Total	45,200	34,798

CAD: Coronary artery disease; CKB: China Kadoorie Biobank; RMB, Renminbi; WHO: World Health Organization.

Data are n (%) or median (25–75th percentile range) unless otherwise specified.

[a] All diabetes was defined as self-reported diabetes or screen-detected diabetes, which was defined as (i) random blood glucose ≥7.0 mmol/L and fasting time ≥8 hours, (ii) random blood glucose ≥11.1 mmol/L and fasting time <8 hours, or (iii) fasting blood glucose ≥7.0 mmol/L on subsequent testing.

[b] Person-year was calculated as the time from the baseline date to the first of the following: death, loss to follow-up, or the global censoring date (December 31, 2017).

[c] Only the first CVD event was included. If a participant was recorded with a CAD event and a stroke event successively or

simultaneously, he/she was counted as one CVD event, one CAD event, and one stroke event. For detailed definitions of each outcome, please see **Table S2**.

Table 2. Harrell's C for the WHO model before and after recalibration in the 10 study regions

Regions	Women			Men		
	Cases	Harrell's C before recalibration [a]	Harrell's C after recalibration	Cases	Harrell's C before recalibration [a]	Harrell's C after recalibration
Nangang (U)	9,290	0.654 (0.648, 0.659)	0.654 (0.648, 0.659)	6,352	0.680 (0.673, 0.687)	0.682 (0.676, 0.689)
Licang (U)	1,439	0.768 (0.756, 0.779)	0.767 (0.755, 0.779)	1,457	0.747 (0.734, 0.759)	0.752 (0.740, 0.764)
Wuzhong (U)	1,840	0.762 (0.751, 0.773)	0.761 (0.750, 0.771)	1,547	0.740 (0.728, 0.752)	0.747 (0.735, 0.759)
Liubei (U)	3,518	0.725 (0.717, 0.734)	0.725 (0.716, 0.733)	2,867	0.729 (0.720, 0.738)	0.734 (0.725, 0.742)
Meilan (U)	2,913	0.692 (0.683, 0.701)	0.691 (0.682, 0.700)	1,850	0.718 (0.707, 0.729)	0.722 (0.711, 0.733)
Huixian (R)	7,222	0.667 (0.661, 0.673)	0.668 (0.661, 0.674)	5,594	0.683 (0.676, 0.690)	0.689 (0.682, 0.696)
Maiji (R)	3,779	0.719 (0.711, 0.727)	0.720 (0.712, 0.728)	2,724	0.728 (0.718, 0.738)	0.732 (0.723, 0.742)
Tongxiang (R)	2,035	0.756 (0.746, 0.767)	0.756 (0.745, 0.766)	2,024	0.747 (0.736, 0.757)	0.752 (0.741, 0.762)
Pengzhou (R)	3,620	0.699 (0.691, 0.708)	0.700 (0.692, 0.709)	2,373	0.730 (0.720, 0.740)	0.736 (0.726, 0.746)
Liuyang (R)	7,263	0.643 (0.636, 0.649)	0.643 (0.636, 0.649)	5,870	0.678 (0.671, 0.685)	0.684 (0.677, 0.691)
Combined (meta)	42,919	0.690 (0.660, 0.721)	0.690 (0.660, 0.720)	32,658	0.706 (0.686, 0.727)	0.711 (0.691, 0.732)

CI: confidence interval; CKB: China Kadoorie Biobank; CVD: cardiovascular disease; WHO: World Health Organization.

"U" represents the urban study region, and "R" represents the rural study region. CVD events were defined according to the definition used by the derivation process of the WHO model (see definition 2 in **Table S2**). The WHO model was recalibrated in each study region (**Supplementary Methods, Table S4**). The C index estimates the probability of the model correctly predicting who will have a CVD event first in a randomly selected pair of participants. The value of this index is between 0.5 and 1. Generally, a C index above 0.7 indicates a good prediction model. The combined C index was the sum of the C index of each study region weighted by inverse variance. The numbers in brackets represent 95% CI. The 95% CI of the combined C index was calculated by using the *t* distribution with 9 degrees of freedom.

[a] Since the outcome definition here is different from that in **Fig. 2**, there are some differences between the C indices before recalibration and those in **Fig. 2**.