SSM - Population Health 25 (2024) 101605

Contents lists available at ScienceDirect

SSM - Population Health

journal homepage: www.elsevier.com/locate/ssmph

The development and validation of a nomogram-based risk prediction model for mortality among older adults

Jun Duan[a,*,1], MingXia Wang[b,1], Napoleon Bellua Sam[c], Qin Tian[d], TingTing Zheng[e], Yun Chen[e,**], XiaoMei Deng[f,***], Yan Liu[a,****]

[a] Department of Medical Record Statistics, Peking University Shenzhen Hospital, Shenzhen, China
[b] Department of Stomatology, Luohu Hospital of Traditional Chinese Medicine, Shenzhen, China
[c] Department of Epidemiology and Biostatistics, University for Development Studies, Tamale, Ghana
[d] Scientific Research Center, The Seventh Affiliated Hospital, Sun Yat-sen University, Shenzhen, Guangdong, 518107, China
[e] Department of Ultrasound, Peking University Shenzhen Hospital, Shenzhen Key Laboratory for Drug Addiction and Medication Safety, Institute of Ultrasound Medicine, Shenzhen-PKU-HKUST Medical Center, Shenzhen, China
[f] Department of Comprehensive Ward, Peking University Shenzhen Hospital, Shenzhen, China

ARTICLE INFO

Keywords:
Oldest old
Predictive model
Population-based longitudinal study
Longevity

ABSTRACT

Objective: This research aims to construct and authenticate a comprehensive predictive model for all-cause mortality, based on a multifaceted array of risk factors.
Methods: The derivation cohort for this study was the Chinese Longitudinal Healthy Longevity Survey (CLHLS), while the Healthy Ageing and Biomarkers Cohort Study (HABCS) and the China Health and Retirement Longitudinal Study (CHARLS) were used as validation cohorts. Risk factors were filtered using lasso regression, and predictive factors were determined using net reclassification improvement. Cox proportional hazards models were employed to establish the mortality risk prediction equations, and the model's fit was evaluated using a discrimination concordance index (C-index). To evaluate the internal consistency of discrimination and calibration, a 10x10 cross-validation technique was employed. Calibration plots were generated to compare predicted probabilities with observed probabilities. The prediction ability of the equations was demonstrated using nomogram.
Results: The CLHLS (mean age 88.08, n = 37074) recorded 28158 deaths (179683 person-years) throughout the course of an 8–20 year follow-up period. Additionally, there were 1384 deaths in the HABCS (mean age 86.74, n = 2552), and 1221 deaths in the CHARLS (mean age 72.48, n = 4794). The final all-cause mortality model incorporated demographic characteristics like age, sex, and current marital status, as well as functional status indicators including cognitive function and activities of daily living. Additionally, lifestyle factors like past smoking condition and leisure activities including housework, television viewing or radio listening, and gardening work were included. The C-index for the derivation cohort was 0.728 (95% CI: 0.724–0.732), while the external validation results for the CHARS and HABCS cohorts were 0.761 (95% CI: 0.749–0.773) and 0.713 (95% CI: 0.697–0.729), respectively.
Conclusion: This study introduces a reliable, validated, and acceptable mortality risk predictor for older adults in China. These predictive factors have potential applications in public health policy and clinical practice.

* Corresponding authors.
** Corresponding author. Department of Medical Record Statistics, Peking University Shenzhen Hospital, No.1120, Lianhua Road, Futian District, Shenzhen, 518036, China.
*** Corresponding authors.
**** Corresponding author.
E-mail addresses: duanjunywq@163.com (J. Duan), yunchen@sphmc.org (Y. Chen), dxmarilyn@163.com (X. Deng), liuyansz@163.com (Y. Liu).
[1] Jun Duan and Mingxia Wang contributed to the work equally.

https://doi.org/10.1016/j.ssmph.2024.101605
Received 7 April 2023; Received in revised form 15 October 2023; Accepted 5 January 2024
Available online 6 January 2024

1. Introduction

The global phenomenon of population aging is a significant concern for numerous countries, with projections indicating that approximately one-fifth of the general population will be 65 years and older by 2050 (Chang et al., 2019). For example, China experienced a notable increase in its population of individuals aged 65 years and above, with figures rising from 8.5% to 12.6% within the past decade, reaching a total of 176 million in 2019(Guo, 2020). In the pursuit of promoting healthy ageing, prognostic information pertaining to mortality has assumed a paramount significance.

Many studies have been reported on a single risk factor for mortality in older adults. For example, prognostic factors have been included general demographic characteristics, lifestyle (such as smoking, alcohol consumption, etc.) (Wei et al., 2021; Zhao et al., 2023), comorbidity (Menotti et al., 2021), functional status(Duan et al., 2020; Fan et al., 2020), health measures (such as blood pressure, lipids, etc.)(Lv et al., 2018, 2019), and genetic risk(Haberle et al., 2021). However, as previously reported, the construction of uncomplicated models incorporating a singular amalgamation of risk factors to predict mortality remains scarce.

Existing studies have used different predictors to construct death prediction models(Lee et al., 2006; Liu, Li, et al., 2021; Suemoto et al., 2017). In addition to age and sex, Lee's study focused on incorporating comorbidities, smoking, and impaired physical function to build predictive models of death (Lee et al., 2006). Liu's prediction of all-cause death was based on basic physiology measures, including blood pressure, lipids and BMI(Liu, Li, et al., 2021). Suemoto's study focused on age, sex, comorbidities, and impaired functioning. However, the follow-up loss in the Survey on Health, Ageing and Retirement in Europe (SHARE) was 36%, which would have introduced selection bias when assessing the correlation between predictors and mortality(Suemoto et al., 2017). In general, these models constructed by later studies have some different predictors, which is worthy of further consideration.

This may be because these mortality prediction studies in older adulters predominantly focusing on different age groups, aged approximately 60–85 years. Additionally, the proportion of the oldest old people (those over 79 years old) in the total elderly population is very small, which in reflection of the overall elderly death prediction is worth further consideration in today's aging society. It's worth noting that previous predicted risk variables may exhibit an inverse association for the oldest old such as hypertension and body mass index (BMI) (Lee et al., 2018; Lv et al., 2018, 2019; Wei et al., 2021). This also means that previous prediction models may not be universal in predicting death in the elderly population.

To address the identified gap, we try to develop and validate a risk prediction model for all-mortality for older adults (those aged 65–105 years) with respect to the Chinese Longitudinal Healthy Longevity Survey (CLHLS). Our goal is to develop a universal predictive model that can be used for epidemiological and clinical purposes in the aging population.

2. Methods

2.1. Derivation cohort

In this study, the CLHLS was used as the derived cohort in this investigation. It aims to better understand human healthy longevity, identify factors that contribute to healthy longevity, and ascertain which factors, out of a large set of biological, behavioral, social, and traditional environment risk factors, play an important role in healthy longevity (Zeng, 2012). In 1998, half of the cities and counties in 23 Chinese provinces were randomly chosen, and then in 2000, 2002, 2005, 2008, 2011, 2014, and 2018, new older persons were recruited and followed up with. The CLHLS utilized a targeted random-sample methodology to ensure the attainment of representativeness. This involved conducting interviews with comparable proportions of male and female individuals in the nonagenarian, octogenarian, and young-old age groups (65–79 years), who resided in close proximity to centenarians (i.e., within the same village or street, if feasible, or within the same sampled county or city). The objective of this approach was to examine the factors influencing healthy longevity among distinct age and gender cohorts residing within a shared social and environmental context. Guidelines and regulations were followed during the conduct of this investigation, and CLHLS has been discussed in more depth (Zeng et al., 2017).

The initial sample for this study comprised 44612 older adults. However, we excluded 600 individuals aged 64 and below, as well as 6823 older adults who were lost participants. Additionally, there were 114 older adults whose death dates were incorrectly recorded (Fig. 1). The Supplementary material 1, provides a more thorough description of the CLHLS's design.

2.2. Validation cohort

The HABCS and CHARLS datasets were utillized as the validation cohort and their datasets used almost the same annotations as the CLHLS. In China, the HABCS was launched in eight longevity areas, which covered the southern, middle and northern areas(Lv, Mao, Gao, et al., 2019). The HABCS adopted a targeted random-sample design to ensure representativeness. For this study, Between 2008 and 2017, a total of 4167 older adults were surveyed in the HABCS datasets with a mean age of 86.74. This cohort consist of 677 centenarians, 730 nonagenarians, 814 octogenarians and1946 younger older adults (aged 65–79 years) (Supplementary material 2). On the other hand, the CHARLS study was carried out between June 2011 and March 2012 and involved a total of 17,708 individual participants. The samples for this study were chosen using a multistage probability sampling method, which covers 150 countries/districts and 450 villages/urban communities representing residents over 45 years old in China(Zhao et al., 2014). The baseline survey was conducted in 2011–2012, followed by waves 2 and 3 in 2015 and 2018, respectively. In order to ensure sample comparison, we excluded the participants aged under 65 years old. A total of 4794 older adults were included (mean age 72.48 years), with almost 86% younger adults (Supplementary material 3).

2.3. Definition of predictive factors

The CLHLS baseline examinations were conducted under strict quality control by trained research staff. A face-to-face standardized questionnaire was used to collect demographic characteristics, lifestyle factors, health conditions functional status, and leisure activities. Demographic characteristics such as sex, age, residence (urban or rural), current marital status (married and living with spouse or not), Co-residence (living with family members or not) and educational background (illiterate or not), occupation (farmer or not), physical labour (yes or no) were collected; functional status such as depression (yes/no), cognitive function (normal, mildly impairment, and severely impairment), disability in activity of daily living (ADL) (yes/no)(Duan et al., 2020; Zeng et al., 2017), physical function (good/poor) were also obtained; lifestyle factors including current/past regular exercise (yes/no), current/past smoke condition (yes/no), current/past drink condition (yes/no), dietary diversity (DD) (categorized as scores 1, 2, 3, 4, 5, ≥6) (Lv et al., 2020); leisure activities including growing vegetables and doing other field work, doing housework, reading newspapers/books, watching television or listening to radio, playing cards and/or mah-jong and gardening work(Li et al., 2020) were categorized as never, sometimes, and often; health conditions including high blood pressure(yes or no), self-reported hypertension (yes or no), self-reported respiratory disease (yes or no) and self-reported heart disease (yes or no). A systolic or diastolic blood pressure of more than 140 mmHg or 90 mmHg, respectively, is considered to be high blood pressure. More details about the definition of related predictive factors are described in the

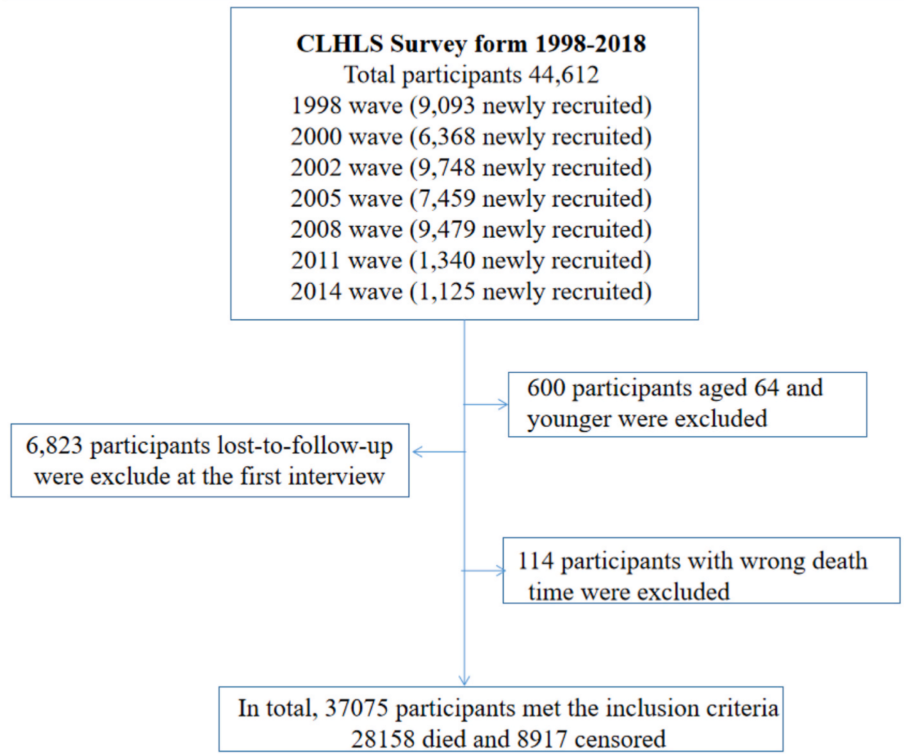

Fig. 1. Cohort Selection Criteria, Chinese Longitudinal Healthy Longevity Study (CLHLS), surveyed form 1998–2018.

Supplementary material 4.

2.4. All-cause mortality

In the follow-up survey, all-cause mortality was the primary outcome. From family members or relatives, survival status was assessed, including whether subjects had completed the study, died, or were lost to follow-up. Lost to follow-up was ascertained when older adults were not found when contacted. A censored older adult was one who survived but was lost to follow-up. In addition, the definition of both outcome and predictor variables across the derivation and validation cohorts were similar.

2.5. Statistical analysis

Continuous variables were represented by mean (SD), and either an unpaired, two-tailed *t*-test or the Mann-Whitney test was employed to compare them. Categorical variables were compared using either a X^2 test or Fisher exact test. Follow-up person years were calculated by determining the time difference between the initial interview date and the occurrence of all-cause mortality, the last follow-up interview, or the point of loss to follow-up, whichever came first.

The three-year and five-year mortality risks were used to build a model as a result of unpublished data showing a 67% higher death risk over a period of 3 years and an average follow-up duration of 5.14 years among older people. To find the independent determinants of mortality, lasso regression was used. The predictive factors for the final models were determined using the enhanced net reclassification improvement (NRI). The NRI, as proposed by Pencina et al. assessed the improvement in categorizing participants into absolute risk categories when new risk factors are added to a risk prediction model(Pencina et al., 2011). Cox proportional hazards models were employed to construct equations for predicting mortality risk. The model fit was assessed using the discrimination concordance index (C index).

In order to evaluate the internal consistency of discrimination and calibration, a 10x10 cross-validation technique was used to assess the internal consistency of the discrimination and calibration. This method randomly divided the dataset into 10 subsets, nine of which were used as training data and one as test data, respectively, for analysis. The complete dataset was divided into 30% and 70% portions for internal validation, which was then done again using the 10x10 cross-validation method.(Tibshirani, 1997). Calibration plots were employed to assess the concordance between predicted and observed probabilities. Bootstrapped samples were used for the calibration evaluation, making it easier to create a calibration plot showing the relationship between observed outcome frequencies and predicted probabilities(Ajnakina et al., 2021). Predictions from properly calibrated models should ideally form a 45-degree angle with the diagonal. The results of the multivariate COX regression analysis were combined with the R rms tool to create a nomogram. Each regression coefficient must be transformed into a proportional scale with a range of 0–100 points as part of the method for creating a nomogram. The nomogram was employed to illustrate the prediction capacity of the equations(Wang et al., 2013). The independent variables were added together to get total points, which were then converted into probabilities. The C index and calibration were used to assess the prediction effectiveness of the nomogram, and 1000 bootstrap samples were used to reduce overfitting bias. To evaluate the effectiveness of the equations, external validations using the HACBS and CHARLS datasets were performed. In the validation cohorts, the discrimination C index and calibration plots were also used to evaluate the equations in a similar manner.

R software (version 3.3.4) was used to analyze the data with the packages "rms", "survival" and "glmnet". Statistical significance was considered using *p*-values <0.05.

3. Results

A total of 37,075 older persons were included in the 179,638 person-years of the study, with a mean age of 88.08 years (SD = 11.24) and 80.22% of them being over 79 years old. The baseline characteristics and mortality risk of the older adults in the derivation cohort were depicted in Fig. 2. A total of 76% of the older adults were followed until the outcome. The older adults in the study were predominantly female, exhibited normal cognition and ADL, live without their spouse, had no history of smoking, never engaged in housework, watched TV or listened radio and participated in garden work. When compared to female, male had the hazard ratio (HR) 1.34(95%CI, 1.29–1.38). In comparison to individuals with normal cognition, older adults with severe cognitive impairment exhibited a HR 1.26(95%CI, 1.22–1.30). Similarly, older adults with impaired ADL showed a HR of 1.27(95%CI, 1.23–1.31) when compared to those with normal ADL. Fig. 2 and Supplementary material 5 both contain additional baseline characteristics and details on mortality risk.

Predictive equations were developed from the derived cohort for assessing all-cause mortality risk among the Chinese older adults. Lasso method was employed to filter variables, resulting in the careful selection of the top 9 based on the score sorting NRI (Supplementary material 6 and 7). Besides the age and sex, we incorporated seven significant risk factor variables; namely cognitive function, ADL function, current marital status, past smoking status, frequency of engaging in housework, watching television or listening to radio, and performing garden work into the equation. Fig. 2 displays the equations' parameters and risk factors for hazards. The internal calibration plot, which evaluated survival probability, showed a good degree of agreement between predictions and actual observation (Fig. 3). The outcomes shown in Table 1 further demonstrated that our equations have excellent mortality discriminatory capacity. In the derivation cohort, the C-index was determined to be 0.728 (95% CI: 0.724–0.732). External validation of the CHARS and HABCS yielded C-index results of 0.761 (95% CI: 0.749–0.773) and 0.713 (95% CI: 0.697–0.729), respectively. The external calibration plot also showed a good agreement (Supplementary material 8 and 9).

Fig. 4 presents the nomogram constructed using the COX model. Age emerged as the most significant predictor of all-cause mortality, followed by ADL function and cognitive function. The score for each predictive factor was calculated using a point scale. By figuring out the cumulative score and mapping it onto the comprehensive points scale, it becomes possible to estimate the likelihood of mortality within a three to five-year timeframe. For instance, individuals aged of 80–84 was allocated a score of 61 point, while males were assigned 10 points. Additional allocations included 16 points for ADL impairment, 16 points for severe cognition impairment, 0 points for married and live with spouse, 7 points for smoking in the past; 15 points for never performing in housework; 12 points for never watching television or listening to the radio, and 13 points for a lack of participation in gardening activity. Notably, the cumulative score of 150 in this instance corresponds to a 60% probability of mortality within three years.

4. Discussion

This study effectively developed and validated a model for estimating the risk of all-cause mortality in Chinese populations across a three-year and five-year period in three prospective cohort studies. The predictive factors encompassed demographic characteristics like age,

Variables	Total (n = 37075)	Survival(n = 8917)	Death(n = 28158)	P Value	Total (HR, 95% CI)
Age, Mean (SD)				<0.01	
	88.08±11.24	79.61±10.99	91.73±9.73		1.055(1.053,1.056)
sex, n (%)				<0.01	
Female	21827 (59)	4876 (55)	16951 (60)		Reference
Male	15248 (41)	4041 (45)	11207 (40)		1.34(1.29,1.38)
Cognition, n (%)				<0.01	
Nomal	21430 (58)	7251 (81)	14179 (50)		Reference
Mildly impairment	5523 (15)	898 (10)	4625 (16)		1.13(1.09,1.17)
Severly impairment	10122 (27)	768 (9)	9354 (33)		1.26(1.22,1.30)
ADL, n (%)				<0.01	
Nomal	26261 (71)	7900 (89)	18361 (65)		Reference
Impairment	10814 (29)	1017 (11)	9797 (35)		1.27(1.23,1.31)
Spouse, n (%)				<0.01	
With	9548 (26)	4178 (47)	5370 (19)		Reference
Without	27527 (74)	4739 (53)	22788 (81)		1.24(1.19,1.28)
Smoke, n (%)				<0.01	
No	25537 (69)	6023 (68)	19514 (69)		Reference
Yes	11538 (31)	2894 (32)	8644 (31)		1.13(1.10,1.16)
Housework, n (%)				<0.01	
Often	12727 (34)	5136 (58)	7591 (27)		Reference
Sometimes	5681 (15)	1391 (16)	4290 (15)		1.14(1.10,1.19)
Never	18667 (50)	2390 (27)	16277 (58)		1.29(1.25,1.34)
Watch TV or listen radio, n (%)				<0.01	
Often	13451 (36)	5141 (58)	8310 (30)		Reference
Sometimes	8820 (24)	2011 (23)	6809 (24)		1.13(1.09,1.17)
Never	14804 (40)	1765 (20)	13039 (46)		1.21(1.17,1.25)
Garden work, n (%)				<0.01	
Often	2434 (7)	1193 (13)	1241 (4)		Reference
Sometimes	1867 (5)	739 (8)	1128 (4)		1.09(1.00,1.18)
Never	32774 (88)	6985 (78)	25789 (92)		1.21(1.14,1.29)

Fig. 2. Baseline characteristics and mortality risk of older adults in the seven successive and non-overlapping cohorts categorized by survival status. Hazard ratio (HR); Confidence interval (CI); Activities of daily living (ADL); With spouse: married and living with spouse; Smoke: smoke in the past.

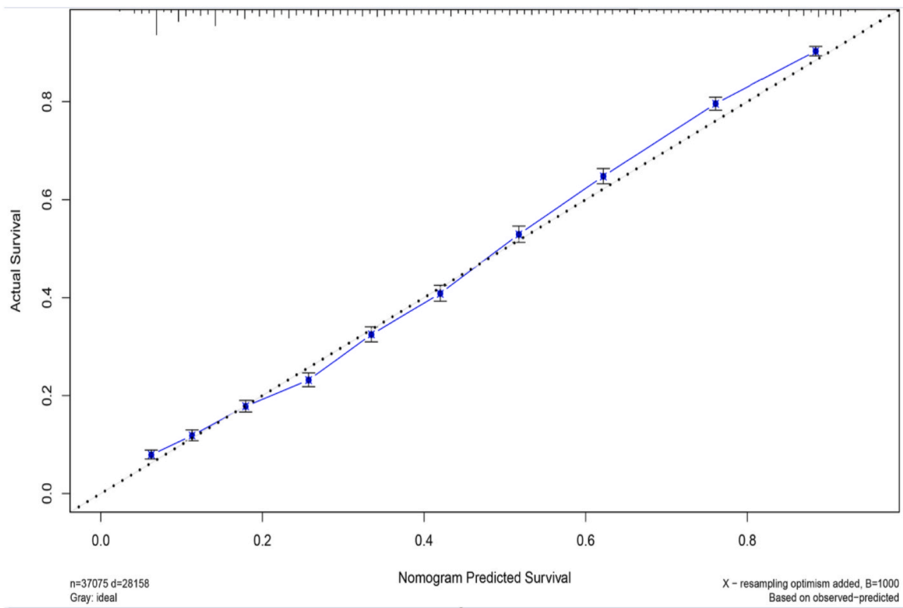

Fig. 3. The internal calibration curve of 10x10 cross-validation for predicting older adults survival in the CLHLS.

Table 1
Internal validation and external validation of mortality risk prediction equations using 10 × 10 cross validation technique.

	Total N	Actual events[a]	C-index	(95% CI)
Derivation Cohort (CLHLS)	37075	28158	0.728	(0.724,0.732)
Internal validation	11120	8387	0.727	(0.723,0.731)
	25955	19771	0.731	(0.725,0.737)
Validation Cohort				
HABCS validation	2522	1384	0.761	(0.749,0.773)
	756	407	0.773	(0.751,0.791)
	1766	977	0.760	(0.746,0.774)
CHARS validation	4794	1210	0.713	(0.697,0.729)
	1436	362	0.705	(0.678,0.732)
	3358	848	0.718	(0.700,0.736)

Abbreviations: CLHLS, Chinese Longitudinal Healthy Longevity Study, surveyed form 1998–2018; HABCS, Healthy Ageing and Biomarkers Cohort Study, surveyed form 2008–2018; CHARLS, China Health and Retirement Longitudinal Study, surveyed form 2011–2018; CI, confidence interval.
Internal validation means that the total data is divided into 30% and 70%, and 10x10 cross-validation is entered again.
[a] Actual number of events through follow-up period.

sex, and current marital status, functional status including cognitive function and ADL, lifestyle factors like past smoking history, and leisure activities like housework, television viewing or radio listening, and gardening work. The nomogram effectively illustrates the predictive capacity of our findings. Mortality risk prediction demonstrated good internal consistency in the CLHLS and external validation in the HABCS and CHARLS datasets.

A limited number of studies have examined the predictive model for all-cause mortality in the Chinese older adults. Liu et al. Suemoto et al. and Lee et al. have developed a predictive model that includes various factors such as age, sex, self-reported diseases, current smoking status, alcohol use, BMI, physical activity(Lee et al., 2006; Liu, Li, et al., 2021; Suemoto et al., 2017). Previous investigations predominantly focusing on individuals aged approximately 60–85 years. In contrast, our prognostic model incorporates an average age of 88.08 years for the elderly, thereby necessitating corresponding adjustments in the prediction model. As the age increased, previous studies have reported inverse associations between hypertension, triglycerides, and BMI among the oldest old individuals(Lv et al., 2019, 2022; Streit et al., 2018). For instance, when it comes to hypertension, the aging process in the elderly is accompanied by noticeable changes in the blood vessels, which are defined by the vessel wall becoming thinner, harder, and more brittle, as well as losing some of its flexibility. Hypertension is essential for supporting the appropriate operation of numerous organs since it works to keep the body's blood supply adequate(Streit et al., 2018). Regarding BMI, the possible explanation for this inverse association is that overweight and obesity could serve as indicators of improved nutritional status, wherein the benefits of enhanced nutrition outweigh the negative effects of higher BMI(Lv et al., 2022). In a similar vein, our study failed to take into account other crucial but less urgent indications, such as hypertension and BMI and so on, which have been investigated in multiple studies concentrating on the oldest elderly population(Lv, Gao, et al., 2018; Sydo et al., 2018; Thinggaard et al., 2010; Wuorela et al., 2020).

Some risk prediction models were developed to predict a cause mortality based on clinical indexes or in specific populations, thus limiting their applicability to the broader older adult population(Bello et al., 2015; Horne et al., 2009; Walter et al., 2001). For example, high total cholesterol levels may be considered significant mortality risk factors for mortality based on clinical indicators(Zhao et al., 2022). However, these indicators are often costly, time-consuming, and not universally accessible, known only to the individuals and their healthcare providers. In the majority of instances, the solitary measurement of biochemical indicators might exhibit immediate stochasticity and fail to capture an individual's distinctive attributes. Consequently, the general public may not be adequately aware of these clinical factors. It is prevalent for the elderly to experience various ailments, and our study also indicates that cardiovascular and respiratory diseases serve as risk factors for mortality in this demographic, displayed in Supplementary material 5. However, the identification and monitoring of disease progression may not always be readily apparent, necessitating regular clinical assessments for confirmation. Identifying impaired cognitive and ADL in older individuals is a straightforward task, wherein the heightened mortality risk can be attributed to dementia resulting from impaired cognitive function and falls stemming from impaired physical function, thereby amplifying the likelihood of death(Duan et al., 2020; Sachs et al., 2011; Stineman et al., 2012). In instances where the elderly experience a decline in functional ability, their family members often serve as the first to recognize this as an early indication of mortality,

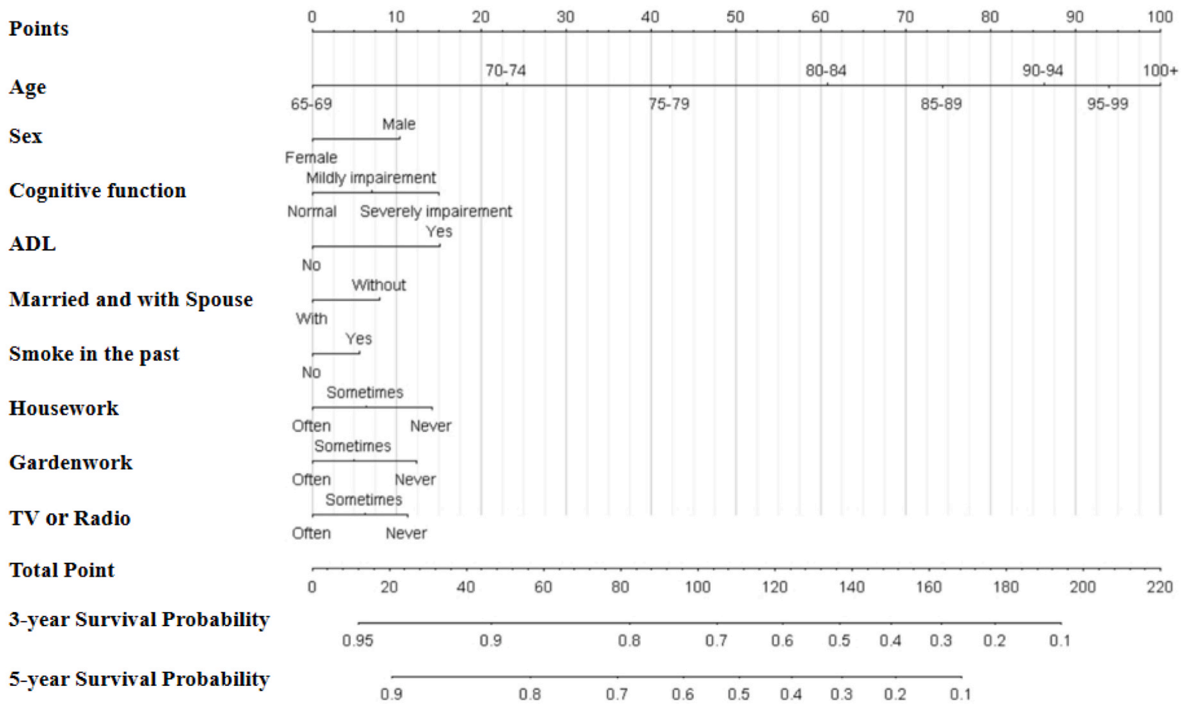

Fig. 4. Survival nomogram. (To use the nomogram, a older adult value is located on each variable axis, and a line is drawn upward to determine the number of points received for each variable value. The sum of these numbers is located on the Total Points axis, and a line is drawn downward to the survival axes to determine the likelihood of 3- or 5-year survival).

prompting them to adopt protective measures in their daily routines. Furthermore, prior models for forecasting mortality among the elderly have not undergone subsequent validation against predictive models utilizing external data.

Our prognostic factors holds potential utility within the realms of public health policy. For instance, the prevalence of marital status without a spouse is largely attributable to spousal loss. Individuals who are married but without a spouse often experience significant negative effects on their mental well-being and undergo substantial changes in both physical and psychological dimensions, consequently elevating the likelihood of mortality(Brenn & Ytterstad, 2016; Wright et al., 2015). The emergence of an aging society has led to a growing prominence of issues faced by the elderly, thereby increasing the attention towards social problems. In light of the social challenges encountered by widowed elderly individuals, it is crucial for adults to assume the responsibility of supporting the elderly and providing them with increased companionship to alleviate feelings of loneliness. In the conclusive model on all-cause mortality, this study incorporated leisure activities. These predictions contribute to the enhancement of physical and mental well-being among older individuals and are worthy of public health promotion. Notably, Leisure activities in our study may be a representative part of the leisure activities of the elderly. Housework could be a long-term lifestyle factor that promotes well-being and health in later life(Murphy et al., 2013). Spending more time on housework has also been found to be associated with a higher likelihood of meeting standard physical activity guidelines and better self-reported health among older adults. Engaging in activities such as watching TV or listening to the radio has been found to have positive effects on thinking abilities, language skills, and social interaction among patients and elderly individuals. Additionally, these forms of entertainment and recreation can effectively contribute to the utilization of leisure time, promoting relaxation and happiness(Li et al., 2020). Furthermore, gardening has been identified as a coping strategy that effectively reduces stress levels and fosters a sense of well-being, thereby potentially influencing overall survival rates(Litt et al., 2023).

Therefore, our prediction model may contribute to the field of public health by examining the implications for predicting population mortality. At the governmental level, strategies such as enhancing public awareness, prioritizing prevention efforts, addressing functional limitations among the elderly, promoting regular physical activity, and fostering increased engagement in recreational activities can be implemented to strengthen public health outcomes. Moreover, our predictive model holds potential in identifying patients at high and low risk, thereby facilitating targeted interventions tailored to each group within clinical settings. For example, An octogenarian male who smokes exhibits pronounced cognitive impairment and compromised physical functioning. Additionally, he lacks the presence of a spouse and abstains from engaging in any leisurely pursuits. In contrast, an octogenarian non-smoker possesses intact cognitive function, normal physical capabilities, enjoys the companionship of a spouse, and regularly participates in leisure activities. Consequently, the three-year mortality risk for the former individual may be approximately 50% greater than that of the latter. For clinical decision making, physicians can employ this information to evaluate the likelihood of survival for elderly patients. While treating the elderly, doctors can also encourage family members to pay more attention to the functional status of the elderly and encourage the elderly to participate in leisure activities.

In this study, the lasso regression was used to establish the all-cause mortality prognostic risk factors. Compared with previous studies, it might be more practical and might play a decision-making role in the formulation of public health policies. This study is one of the few researches to develop and validate mortality prediction models that include a large sample of the oldest old. It also explored many important prognostic factors and eventually incorporated predictive factors from demographics, self-reported diseases and lifestyle factors. Moreover, in order to optimize our prediction model, we decided the final prognostic factor to be included based on the improvement of the contribution of NRI and public health significance. Although some of the predictive factors for mortality were not included in the model, our prediction model was more reasonable in practical and statistical filed.

This study is not devoid of limitations. Firstly, although individuals were considered representative of the general older adults, the verification of different countries and different races were still missing because of the unavailability of data and the scarcity of representative samples of the older adults. Secondly, exploration of cause mortality, such as the accidental death, was not performed because of the unavailable cause-specifific mortality data. Thirdly, we can prove that traditional predictive factors were good performance in predicting mortality but cannot prove that traditional predictive factors were better than clinical indicators. However, findings from this study might lay a certain foundation for further research in the future.

5. Conclusions

This study developed and validated a predictive model for estimating all-cause mortality risk with respect to the CLHLS and evaluated the performance of the developed equations in the HABCS and the CHARLS. It was a rare study on the prediction of death in the oldest old with a large sample and use non-clinical indicators. The predictive factors encompassed demographic characteristics like age, sex, and current marital status, functional status including cognitive function and ADL, lifestyle factors like past smoking history, and leisure activities like housework, television viewing or radio listening, and gardening work. The nomogram effectively illustrates the predictive capacity of our findings. The incorporation of these prognostic factors could potentially provide support for public health policy and medical decision-making. However, the limitations of this study are mainly that it has not been extensively independent evaluated across different populations, countries, and regions.

Availability of data and materials

Data will be made available on request. In the CLHLS, HABCS, and CHARLS cohorts, there are open cohorts.

Data can be obtained from the English site https://sites.duke.edu/centerforaging/programs/chinese-longitudinal-healthy-longevity-survey-clhls/, as well as the Chinese sites https://opendata.pku.edu.cn/https://opendata.pku.edu.cn/andhttps://charls.pku.edu.cn/index/zh-cn.html.

Funding

The “San-ming” Project of Medicine in Shenzhen funded this work (NO. SZSM201812088). Shenzhen Science and technology planning project (JCYJ20190809111017163), Shenzhen Sustainable Development Technology Special Project (KCXFZ20200201101048774) and Scientific Research Foundation of Shenzhen Health Economics Association (202304).

Ethics approval

The biomedical ethics committee of Peking University approved the CLHLS Protection of Human Subjects protocol (IRB00001052-13074). All participants and/or their relatives provided informed and written consent.

Author Statement

Jun Duan, MingXia Wang Conceptualization, Methodology, Software, WritingOriginal Draft. Yun Chen, XiaoMei Deng, Yan Liu, Jun Duan: Conceptualization, Methodology, Supervision, Writing-Review & Editing, Funding Acquisition. Napoleon Bellua Sam, Qin Tian, TingTing Zheng: Conceptualization, Formal analysis, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

Our sincere thanks go out to all participants and their families. We acknowledge all staffs for their great contributions to the success of the three cohorts and the teams that provided the shared database. A special thank you goes out to Prof. Yi Zeng, the principal investigator, and Dr. Huashuai Chen, the data manager.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ssmph.2024.101605.

Abbreviations

ADL:	Disability in Activity of Daily Living
BMI	Body mass index
CHARLS	China Health and Retirement Longitudinal Study
CI	Confidence interval
CLHLS	Chinese Longitudinal Healthy Longevity Survey
DD	Dietary Diversity
HABCS	Healthy Ageing and Biomarkers Cohort Study
HR	Hazard ratio
MMSE	Mini-Mental State Examination
NRI	Net Reclassification Improvement

References

Ajnakina, O., Agbedjro, D., Mccammon, R., Faul, J., Murray, R. M., Stahl, D., et al. (2021). Development and validation of prediction model to estimate 10-year risk of all-cause mortality using modern statistical learning methods: A large population-based cohort study and external validation. *BMC Medical Research Methodology, 21*(1), 8. https://doi.org/10.1186/s12874-020-01204-7

Bello, G. A., Dumancas, G. G., & Gennings, C. (2015). Development and validation of a clinical risk-assessment tool predictive of all-cause mortality. *Bioinformatics and Biology Insights, 9*(Suppl 3), 1–10. https://doi.org/10.4137/BBI.S30172

Brenn, T., & Ytterstad, E. (2016). Increased risk of death immediately after losing a spouse: Cause-specific mortality following widowhood in Norway. *Preventive Medicine, 89*, 251–256. https://doi.org/10.1016/j.ypmed.2016.06.019

Chang, A. Y., Skirbekk, V. F., Tyrovolas, S., Kassebaum, N. J., & Dieleman, J. L. (2019). Measuring population ageing: An analysis of the global burden of disease study 2017. *The Lancet Public Health, 4*(3), e159–e167. https://doi.org/10.1016/S2468-2667(19)30019-2

Duan, J., Lv, Y. B., Gao, X., Zhou, J. H., Kraus, V. B., Zeng, Y., et al. (2020). Association of cognitive impairment and elderly mortality: Differences between two cohorts ascertained 6-years apart in China. *BMC Geriatrics, 20*(1), 29. https://doi.org/10.1186/s12877-020-1424-4

Fan, J., Yu, C., Guo, Y., Bian, Z., Sun, Z., Yang, L., et al. (2020). Frailty index and all-cause and cause-specific mortality in Chinese adults: A prospective cohort study. *The Lancet Public Health, 5*(12), e650–e660. https://doi.org/10.1016/S2468-2667(20)30113-4

Guo, R. (2020). *China ethnic statistical yearbook 2020*.

Haberle, A. D., Biggs, M. L., Cushman, M., Psaty, B. M., Newman, A. B., Shlipak, M. G., et al. (2021). Level and change in n-terminal pro-b-type natriuretic peptide and kidney function and survival to age 90. *J Gerontol A Biol Sci Med Sci, 76*(3), 478–484. https://doi.org/10.1093/gerona/glaa124

Horne, B. D., May, H. T., Muhlestein, J. B., Ronnow, B. S., Lappe, D. L., Renlund, D. G., et al. (2009). Exceptional mortality prediction by risk scores from common laboratory tests. *The American Journal of Medicine, 122*(6), 550–558. https://doi.org/10.1016/j.amjmed.2008.10.043

Lee, D. H., Keum, N., Hu, F. B., Orav, E. J., Rimm, E. B., Willett, W. C., et al. (2018). Predicted lean body mass, fat mass, and all cause and cause specific mortality in men: Prospective us cohort study. *BMJ, 362*, Article k2575. https://doi.org/10.1136/bmj.k2575

Lee, S. J., Lindquist, K., Segal, M. R., & Covinsky, K. E. (2006). Development and validation of a prognostic index for 4-year mortality in older adults. *JAMA, 295*(7), 801–808. https://doi.org/10.1001/jama.295.7.801

Li, Z. H., Zhang, X. R., Lv, Y. B., Shen, D., Li, F. R., Zhong, W. F., et al. (2020). Leisure activities and all-cause mortality among the Chinese oldest-old population: A prospective community-based cohort study. *Journal of the American Medical Directors Association, 21*(6), 713–719. https://doi.org/10.1016/j.jamda.2019.08.003

Litt, J. S., Alaimo, K., Harrall, K. K., Hamman, R. F., Hebert, J. R., Hurley, T. G., et al. (2023). Effects of a community gardening intervention on diet, physical activity, and anthropometry outcomes in the USA (caps): An observer-blind, randomised controlled trial. *The Lancet Planetary Health, 7*(1), e23–e32. https://doi.org/10.1016/S2542-5196(22)00303-5

Liu, C. S., Li, C. I., Wang, M. C., Yang, S. Y., Li, T. C., & Lin, C. C. (2021). Building clinical risk score systems for predicting the all-cause and expanded cardiovascular-specific mortality of patients with type 2 diabetes. *Diabetes, Obesity and Metabolism, 23*(2), 467–479. https://doi.org/10.1111/dom.14240

Liu, L., Lo, K., Huang, C., Feng, Y. Q., Zhou, Y. L., & Huang, Y. Q. (2021). Derivation and validation of a simple nomogram prediction model for all-cause mortality among middle-aged and elderly general population. *Annals of Palliative Medicine, 10*(2), 1167–1179. https://doi.org/10.21037/apm-20-580

Lv, Y. B., Gao, X., Yin, Z. X., Chen, H. S., Luo, J. S., Brasher, M. S., et al. (2018). Revisiting the association of blood pressure with mortality in oldest old people in China: Community based, longitudinal prospective study. *BMJ, 361*, k2158. https://doi.org/10.1136/bmj.k2158

Lv, Y., Kraus, V. B., Gao, X., Yin, Z., Zhou, J., Mao, C., et al. (2020). Higher dietary diversity scores and protein-rich food consumption were associated with lower risk of all-cause mortality in the oldest old. *Clinical Nutrition, 39*(7), 2246–2254. https://doi.org/10.1016/j.clnu.2019.10.012

Lv, Y. B., Liu, S., Yin, Z. X., Gao, X., Kraus, V. B., Mao, C., et al. (2018). Associations of body mass index and waist circumference with 3-year all-cause mortality among the oldest old: Evidence from a Chinese community-based prospective cohort study. *Journal of the American Medical Directors Association, 19*(8), 672–678. https://doi.org/10.1016/j.jamda.2018.03.015

Lv, Y., Mao, C., Gao, X., Ji, J. S., Kraus, V. B., Yin, Z., et al. (2022). The obesity paradox is mostly driven by decreased noncardiovascular disease mortality in the oldest old in China: A 20-year prospective cohort study. *Nat Aging, 2*(5), 389–396. https://doi.org/10.1038/s43587-022-00201-3

Lv, Y. B., Mao, C., Gao, X., Yin, Z. X., Kraus, V. B., Yuan, J. Q., et al. (2019). Triglycerides paradox among the oldest old: "the lower the better?". *Journal of the American Geriatrics Society, 67*(4), 741–748. https://doi.org/10.1111/jgs.15733

Lv, Y., Mao, C., Yin, Z., Li, F., Wu, X., & Shi, X. (2019). Healthy ageing and biomarkers cohort study (habcs): A cohort profile. *BMJ Open, 9*(10), Article e26513. https://doi.org/10.1136/bmjopen-2018-026513

Menotti, A., Puddu, P. E., & Tolonen, H. (2021). Age at death in elderly cohorts of four european countries of the seven countries study: The role of comorbidities. *Aging Clinical and Experimental Research, 33*(3), 521–528. https://doi.org/10.1007/s40520-020-01568-2

Murphy, M. H., Donnelly, P., Breslin, G., Shibli, S., & Nevill, A. M. (2013). Does doing housework keep you healthy? The contribution of domestic physical activity to meeting current recommendations for health. *BMC Public Health, 13*, 966. https://doi.org/10.1186/1471-2458-13-966

Pencina, M. J., D'Agostino, R. S., & Steyerberg, E. W. (2011). Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Statistics in Medicine, 30*(1), 11–21. https://doi.org/10.1002/sim.4085

Sachs, G. A., Carter, R., Holtz, L. R., Smith, F., Stump, T. E., Tu, W., et al. (2011). Cognitive impairment: An independent predictor of excess mortality: A cohort study. *Annals of Internal Medicine, 155*(5), 300–308. https://doi.org/10.7326/0003-4819-155-5-201109060-00007

Stineman, M. G., Xie, D., Pan, Q., Kurichi, J. E., Zhang, Z., Saliba, D., et al. (2012). All-cause 1-, 5-, and 10-year mortality in elderly people according to activities of daily living stage. *Journal of the American Geriatrics Society, 60*(3), 485–492. https://doi.org/10.1111/j.1532-5415.2011.03867.x

Streit, S., Poortvliet, R., & Gussekloo, J. (2018). Lower blood pressure during antihypertensive treatment is associated with higher all-cause mortality and accelerated cognitive decline in the oldest-old. Data from the leiden 85-plus study. *Age and Ageing, 47*(4), 545–550. https://doi.org/10.1093/ageing/afy072

Suemoto, C. K., Ueda, P., Beltran-Sanchez, H., Lebrao, M. L., Duarte, Y. A., Wong, R., et al. (2017). Development and validation of a 10-year mortality prediction model: meta-analysis of individual participant data from five cohorts of older adults in developed and developing countries. *J Gerontol A Biol Sci Med Sci, 72*(3), 410–416. https://doi.org/10.1093/gerona/glw166

Sydo, N., Merkely, B., Carta, K., Becker, D., Hussain, N., Murphy, J. G., et al. (2018). Effect of cardiorespiratory fitness on co-morbidities and mortality in never, past, and current smokers. *The American Journal of Cardiology, 122*(10), 1765–1772. https://doi.org/10.1016/j.amjcard.2018.08.012

Thinggaard, M., Jacobsen, R., Jeune, B., Martinussen, T., & Christensen, K. (2010). Is the relationship between bmi and mortality increasingly u-shaped with advancing age? A 10-year follow-up of persons aged 70-95 years. *J Gerontol A Biol Sci Med Sci, 65*(5), 526–531. https://doi.org/10.1093/gerona/glp214

Tibshirani, R. (1997). The lasso method for variable selection in the cox model. *Statistics in Medicine, 16*(4), 385–395. https://doi.org/10.1002/(sici)1097-0258(19970228)16:4<385::aid-sim380>3.0.co;2-3

Walter, L. C., Brand, R. J., Counsell, S. R., Palmer, R. M., Landefeld, C. S., Fortinsky, R. H., et al. (2001). Development and validation of a prognostic index for 1-year mortality in older adults after hospitalization. *JAMA, 285*(23), 2987–2994. https://doi.org/10.1001/jama.285.23.2987

Wang, Y., Li, J., Xia, Y., Gong, R., Wang, K., Yan, Z., et al. (2013). Prognostic nomogram for intrahepatic cholangiocarcinoma after partial hepatectomy. *Journal of Clinical Oncology, 31*(9), 1188–1195. https://doi.org/10.1200/JCO.2012.41.5984

Wei, Y., Lv, Y., Zhou, J., Gao, X., Duan, J., Zhao, C., et al. (2021). Smoking cessation in late life is associated with increased risk of all-cause mortality amongst oldest old people: A community-based prospective cohort study. *Age and Ageing, 50*(4), 1298–1305. https://doi.org/10.1093/ageing/afaa280

Wright, D. M., Rosato, M., & O'Reilly, D. (2015). Urban/rural variation in the influence of widowhood on mortality risk: A cohort study of almost 300,000 couples. *Health & Place, 34*, 67–73. https://doi.org/10.1016/j.healthplace.2015.04.003

Wuorela, M., Lavonius, S., Salminen, M., Vahlberg, T., Viitanen, M., & Viikari, L. (2020). Self-rated health and objective health status as predictors of all-cause mortality among older people: A prospective study with a 5-, 10-, and 27-year follow-up. *BMC Geriatrics, 20*(1), 120. https://doi.org/10.1186/s12877-020-01516-9

Zeng, Y. (2012). Towards deeper research and better policy for healthy aging –using the unique data of Chinese longitudinal healthy longevity survey. *China Economic Journal, 5*(2–3), 131–149. https://doi.org/10.1080/17538963.2013.764677

Zeng, Y., Feng, Q., Hesketh, T., Christensen, K., & Vaupel, J. W. (2017). Survival, disabilities in activities of daily living, and physical and cognitive functioning among the oldest-old in China: A cohort study. *Lancet, 389*(10079), 1619–1629. https://doi.org/10.1016/S0140-6736(17)30548-2

Zhao, B., Gan, L., Graubard, B. I., Mannisto, S., Albanes, D., & Huang, J. (2022). Associations of dietary cholesterol, serum cholesterol, and egg consumption with overall and cause-specific mortality: Systematic review and updated meta-analysis. *Circulation, 145*(20), 1506–1520. https://doi.org/10.1161/CIRCULATIONAHA.121.057642

Zhao, Y., Hu, Y., Smith, J. P., Strauss, J., & Yang, G. (2014). Cohort profile: The China health and retirement longitudinal study (charls). *International Journal of Epidemiology, 43*(1), 61–68. https://doi.org/10.1093/ije/dys203

Zhao, J., Stockwell, T., Naimi, T., Churchill, S., Clay, J., & Sherk, A. (2023). Association between daily alcohol intake and risk of all-cause mortality: A systematic review and meta-analyses. *JAMA Network Open, 6*(3), Article e236185. https://doi.org/10.1001/jamanetworkopen.2023.6185