

Preplanned Studies

Novel Genetic Loci Associated with PhenoAge Acceleration — Changzhou City, Jiangsu Province, China, 2012–2019

Yifan Wang^{1,8}; Yue Jiang^{2,3,8}; Hui Wang^{4,8}; Shuang Liang¹; Muhammed Lamin Sambou¹; Mei Wang¹; Meiqi Yang¹; Yakubu Salimata¹; Feifei Xu¹; Chenjie Li¹; Xiao Wang¹; Junpeng Yang⁵; Meng Zhu^{1,4}; Cheng Wang^{1,2,3,6}; Guangfu Jin^{1,2,3,4}; Hongxia Ma^{1,2,3,4}; Honggang Yi⁷; Hongbing Shen^{1,3,4,8,#}; Juncheng Dai^{1,2,4,9,#}

Summary

What is already known about this topic?

China is rapidly encountering population aging, yet studies on aging are limited by the traditional aging measure: chronological age, particularly in the field of genomics. Several promising aging measures have been proposed, but they lack comparative evaluation.

What is added by this report?

PhenoAge was identified as a measure of aging that demonstrated greater applicability in contemporary populations. Based on this, several novel genetic variants were found to enhance the predictive accuracy of aging.

What are the implications for public health practice?

These findings might provide new insights into aging and facilitate the development of a practical screening program based on PhenoAge, which aims to promote healthy aging in China.

China is rapidly experiencing population aging, but studies on aging are limited by the traditional aging measure: chronological age (CA), particularly in the field of genomics. Several novel aging measures have been proposed, but they lack comparative evaluation. Therefore, this study aimed to develop and select a more accurate measure for aging and identify novel aging-associated genetic variants. We developed three aging measures with potential for promotion based on biomarkers screened by Lasso regression in 7,584 participants from the Changzhou cohort. We assessed their performance and chose PhenoAge as the best aging measure by area under the receiver operating characteristic (ROC) curve (AUC). To identify aging-associated loci, we conducted a genome-wide association study (GWAS) using PhenoAge acceleration (PhenoAgeAccel) as the aging phenotype among 1,215 genotyped participants through a linear regression model. A total of 24 aging-associated

variants were identified, with 3 previously recognized as aging loci and 21 suggesting novel contributions. This study then further explored the functions of these loci on a multi-omics scale. These findings may provide new insights into aging and facilitate the development of a practical screening program based on PhenoAge, which aims to promote healthy aging in China.

Aging measures are important strategies for constructing and quantifying aging rate (1). Chronological age (CA) is the principal aging measure, offering advantages in simplicity and widespread acceptance. However, it introduces heterogeneity across studies (2). To overcome CA's limitations, several promising aging measures have been proposed (3). Among these, biological age (BA) measures based on composite biomarkers are cost-effective and demonstrate improved performance in replicating reported aging loci and suggesting genetic correlations with age-related phenotypes (4). Given the numerous BA types, evaluating the most suitable BA for the Chinese population is crucial.

The Changzhou cohort is an independent prospective cohort from Wujin District, Changzhou City, Jiangsu Province, China (5). A total of 7,584 individuals were included in this study. The study design workflow is illustrated in [Supplementary Figure S1](#) (available at <https://weekly.chinacdc.cn/>). Lasso regression was used to screen for candidate biomarkers. Three BA measures — KDM-BA, PhenoAge, and HD (6) — are more appropriate for application in China. These measures were developed based on candidate biomarkers, and the measure with the highest AUC value was selected. Aging acceleration was estimated as the residual of these aging measures after adjusting for CA using a linear regression model. Details regarding the subjects and BA measures are provided in the [Supplementary Material](#) (available at <https://weekly.chinacdc.cn/>).

We performed quality control for 1,244 genotyped participants and included 1,215 participants with

approximately 8.15 million variants. PLINK2.0 was used to perform linear regression analysis on the aging phenotype (7), adjusting for the first 10 principal components (PCs), age, sex, and genotyping batch. In the identification of aging-associated loci, we set the significance level at $P < 10^{-5}$ and considered leading SNPs with a threshold of $r^2 \geq 0.6$. ANNOVAR, Reactome, GWAS Catalog, IEUgwas, and GTEx were used for functional mapping and annotation of identified loci. To further explore the functions of these loci on the proteome and phenome, we analyzed protein-aging association and performed a phenome-wide association analysis (PheWAS) in 430,000 UK Biobank (UKB) participants. In the protein-aging association analysis, we adjusted for sex, age, lifestyle, and other appropriate factors. The details of the above analyses are provided in the [Supplementary Material](#).

A total of 7,584 participants were followed up from 2012 to 2019, with a mean follow-up time of $6.37 \pm$

0.53 years. There were 3 BA measures constructed based on sex and 18 biomarkers screened by Lasso regression. Based on the AUC, PhenoAge (AUC=0.79) was selected as the optimal aging measure for this study (Figure 1). The formula and distribution of PhenoAge and PhenoAgeAccel are shown in [Supplementary Figure S2](#) (available at <https://weekly.chinacdc.cn/>).

We identified 24 aging-associated loci tagged by 24 lead SNPs for PhenoAgeAccel (Table 1), 3 of which were previously reported. The genes mapped by these loci were primarily involved in programmed cell death, cell cycle, and immune system pathways. The effects of several identified variants on aging were significantly heterogeneous among age and sex subgroups. Notably, lead SNPs rs9376269, rs76772550, and rs67548191 were identified as eQTLs of *IFNGR1*, *SLC24A4*, and *GGPS1*, respectively, in both the IEUgwas and GTEx databases. *IFNGR1* protein was also positively correlated with PhenoAgeAccel $P < 0.001$ in the UKB.

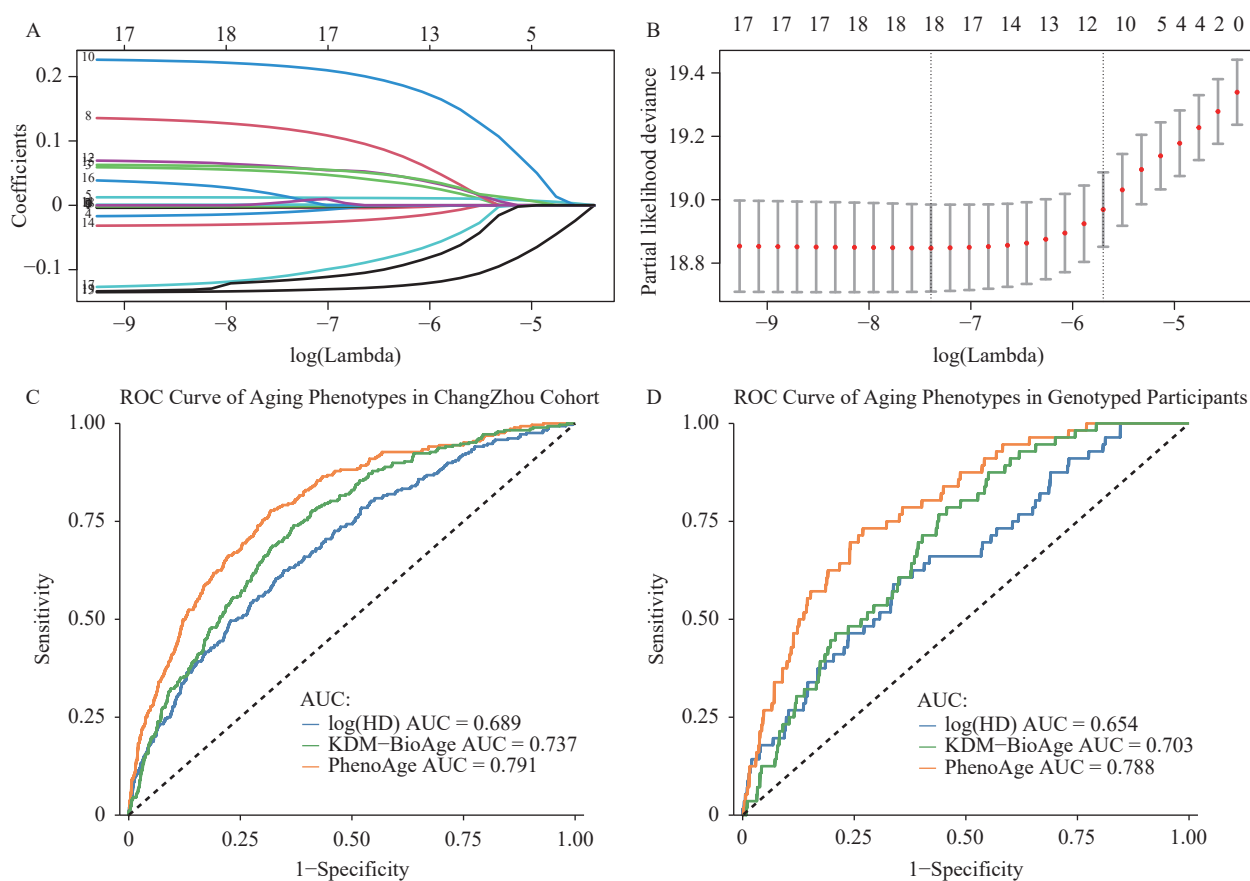


FIGURE 1. Screening of variables based on Lasso regression and evaluation of aging measures based on AUC. (A) The variation characteristics of the coefficient of variables in Lasso regression; (B) The selection process of the optimum value of the parameter λ in the Lasso regression model by cross-validation method; (C) AUC of three aging measures in the Changzhou Cohort; (D) AUC of three aging measures in the Changzhou genotyped group.

Abbreviation: AUC=area under the receiver operating characteristic curve; HD=homeostatic dysregulation; ROC=receiver operating characteristic curve.

TABLE 1. Results for 24 identified aging-associated risk loci.

SNP	CHR	POS	RA	EA	MAF	BETA	SE	P	Related genes	Region
rs78629466	1	156087706	G	A	0.098	1.539	0.344	7.49×10 ⁻⁶	LMNA	Intronic
rs34186915	1	206806424	C	A	0.034	2.547	0.557	4.87×10 ⁻⁶	EIF2D, DYRK3	Intergenic
rs67548191	1	235324984	C	T	0.017	4.721	1.065	9.30×10 ⁻⁶	RBM34	Upstream
rs6712152	2	155748871	G	C	0.044	2.552	0.492	2.11×10 ⁻⁷	KCNJ3, LINC01876	Intergenic
rs75159275	2	156053319	T	A	0.036	2.929	0.544	7.15×10 ⁻⁸	KCNJ3, LINC01876	Intergenic
rs17266628	3	122127926	G	A	0.279	-1.096	0.225	1.09×10 ⁻⁶	FAM162A	Intronic
rs534553017	4	186795694	C	T	0.015	-5.479	1.145	1.72×10 ⁻⁶	SORBS2	Intronic
rs143045396	4	190324195	G	T	0.054	2.077	0.45	4.02×10 ⁻⁶	LINC02508, LINC01262	Intergenic
rs199764613	5	120002898	G	A	0.026	-4.132	0.891	3.56×10 ⁻⁶	PRR16	Intronic
rs761656272	5	159746937	TAATA	T	0.018	-3.620	0.819	9.99×10 ⁻⁶	CCNJL	Intergenic
rs9376269	6	137539505	G	C	0.476	-1.082	0.205	1.25×10 ⁻⁷	IFNGR1	Intronic
rs10237037	7	7848087	C	T	0.036	2.655	0.546	1.18×10 ⁻⁶	UMAD1	Intronic
rs78007164	7	75998408	C	T	0.028	3.006	0.619	1.18×10 ⁻⁶	YWHAG, SSC4D	Intergenic
rs137923974	7	148380093	G	A	0.029	-2.961	0.607	1.10×10 ⁻⁶	C7orf33, CUL1	Intergenic
rs1446270	9	12046808	C	T	0.476	-0.937	0.199	2.57×10 ⁻⁶	PTPRD-AS2, TYRP1	Intergenic
rs546542	9	77332549	C	T	0.278	-1.028	0.228	6.49×10 ⁻⁶	RORB, TRPM6	Intergenic
rs79594032	14	20724036	C	T	0.022	4.566	0.959	1.95×10 ⁻⁶	OR11H4, TTC5	Intergenic
rs2060609	14	43338730	G	A	0.162	1.251	0.277	6.21×10 ⁻⁶	LRFN5, FSCB	Intergenic
rs76772550	14	92970259	C	A	0.184	-1.277	0.266	1.56×10 ⁻⁶	SLC24A4, RIN3	Intergenic
rs4778079	15	93593528	C	T	0.207	1.157	0.251	3.90×10 ⁻⁶	RGMA	Intronic
rs138259742	16	73965374	T	A	0.016	4.998	1.123	8.57×10 ⁻⁶	LINC01568, LOC101928035	Intergenic
rs1432070	18	70162291	C	T	0.231	-1.073	0.233	4.01×10 ⁻⁶	LINC01899, CBLN2	Intergenic
rs11881034	19	6281607	C	T	0.292	0.985	0.22	7.87×10 ⁻⁶	MLLT1, ACER1	Intergenic
rs285200	20	42366834	C	T	0.013	6.257	1.241	4.63×10 ⁻⁷	GTSF1L, LINC01728	Intergenic

Abbreviation: BETA=beta coefficient; CHR=chromosome; EA=effect allele; MAF=minor allele frequency; POS=position; RA=reference allele; SE=standard error; SNP=single nucleotide polymorphism.

Therefore, rs9376269 may affect aging through its influence on the transcription and translation of *IFNGR1*.

In the PheWAS, we found 212 significant associations between 24 variants and 75 outcomes (Figure 2). Half of these variants were associated with age-related phenotypes such as diabetes, hypertension, and cardiovascular diseases. Additionally, other variant-phenotype associations involved metabolic diseases, infectious diseases, blood diseases, and certain tumors. These results suggested that these aging-associated variants were associated with aging-related phenotypes, and that metabolic, infectious, and cardiovascular diseases contributed most to aging in the Changzhou population.

DISCUSSION

This study investigated appropriate aging measures

and aging-associated variants in the Changzhou cohort. Among BA measures, PhenoAge was identified as the most applicable for this population through comparative evaluation. Twenty-four aging-associated loci were identified based on PhenoAgeAccel.

Previous aging studies, primarily based on CA, identified numerous aging-associated loci. Of these, only Apolipoprotein E (*APOE*) gene has been widely recognized as an aging-associated gene (2). A limitation of these studies is the difficulty in eliminating heterogeneity introduced by varying longevity thresholds (2). To address this, alternative aging measures, particularly BA measures, have been developed. However, the optimal measure for quantifying aging remains uncertain. Our research compared aging measures from similar previous studies and identified PhenoAge as optimal for the current population.

Compared with previous similar studies, our study employed lasso regression to further screen biomarkers,

famine periods. According to the Developmental Origins of Health and Disease theory, such conditions might be associated with a higher incidence of cardiovascular and metabolic diseases in their offspring. Thus, the effect of aging genes mainly related to inflammation and metabolism is more pronounced in this study.

This study has several limitations. The participants were recruited from Changzhou in Jiangsu Province, a relatively developed city in China with a predominantly Han population, potentially limiting the generalizability of our findings to the broader Chinese population. Additionally, the small sample size and stringent quality control standards may have limited our ability to replicate previously reported aging-associated variants, particularly those related to *APOE*.

In summary, our study established PhenoAge as an aging measure that may demonstrate greater applicability in the Chinese population. Furthermore, we identified several novel aging-associated variants. Our findings showed significant potential for developing a practical, PhenoAge-based aging screening method to promote healthy aging in China.

Conflicts of interest: No conflicts of interest.

Acknowledgements: The well-dedicated participants and staff who participated in the Changzhou cohort study.

Funding: Supported by the National Natural Science Foundation of China (Grant ID: 81941020, 82192903, 82192904), the Science Fund for Distinguished Young Scholars of Jiangsu Province (BK20211533), the Natural Science Foundation of Jiangsu Province (BK20221183), and the Nanjing Medical Science and Technique Development Foundation (JQX18009). These supporting sources had no involvement and restrictions regarding the publication.

doi: 10.46234/ccdcw2024.258

Corresponding authors: Juncheng Dai, djc@njmu.edu.cn; Hongbing Shen, hbshen@njmu.edu.cn.

¹ Department of Epidemiology, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing City, Jiangsu Province, China; ² State Key Laboratory of Reproductive Medicine, Nanjing Medical University, Nanjing City, Jiangsu Province, China; ³ China International Cooperation Center for Environment and

Human Health, Nanjing Medical University, Nanjing City, Jiangsu Province, China; ⁴ State Key Lab of Cancer Biomarkers, Prevention and Treatment; Collaborative Innovation Center for Cancer Personalized Medicine, Nanjing Medical University, Nanjing City, Jiangsu Province, China; ⁵ The First School of Clinical Medicine, Nanjing Medical University, Nanjing, China; ⁶ Department of Bioinformatics, School of Biomedical Engineering and Informatics, Nanjing Medical University, Nanjing City, Jiangsu Province, China; ⁷ Department of Biostatistics, School of Public Health, Nanjing Medical University, Nanjing City, Jiangsu Province, China; ⁸ The Chinese Center for Disease Control and Prevention, Beijing, China; ⁹ Jiangsu Life and Health Industry Academician Collaborative Innovation Center, Nanjing City, Jiangsu Province, China.
[‡] Joint first authors.

Submitted: July 07, 2024

Accepted: September 02, 2024

Issued: December 06, 2024

REFERENCES

- Ferrucci L, Gonzalez-Freire M, Fabbri E, Simonsick E, Tanaka T, Moore Z, et al. Measuring biological aging in humans: a quest. *Aging Cell* 2020;19(2):e13080. <https://doi.org/10.1111/acel.13080>.
- Deelen J, Evans DS, Arking DE, Tesi N, Nygaard M, Liu XM, et al. A meta-analysis of genome-wide association studies identifies multiple longevity genes. *Nat Commun* 2019;10(1):3669. <https://doi.org/10.1038/s41467-019-11558-2>.
- Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY)* 2018;10(4):573 - 91. <https://doi.org/10.18632/aging.101414>.
- Kuo CL, Pilling LC, Liu ZY, Atkins JL, Levine ME. Genetic associations for two biological age measures point to distinct aging phenotypes. *Aging Cell* 2021;20(6):e13376. <https://doi.org/10.1111/acel.13376>.
- Chen W, Lu F, Liu SJ, Du JB, Wang JM, Qian Y, et al. Cancer risk and key components of metabolic syndrome: a population-based prospective cohort study in Chinese. *Chin Med J (Engl)* 2012;125(3):481 - 5. <https://doi.org/10.3760/cma.j.issn.0366-6999.2012.03.015>.
- Kwon D, Belsky DW. A toolkit for quantification of biological age from blood chemistry and organ function test data: BioAge. *GeroScience* 2021;43(6):2795 - 808. <https://doi.org/10.1007/s11357-021-00480-5>.
- Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 2015;4:7. <https://doi.org/10.1186/s13742-015-0047-8>.
- van de Wetering D, de Paus RA, van Dissel JT, van de Vosse E. Functional analysis of naturally occurring amino acid substitutions in human IFN- γ R1. *Mol Immunol* 2010;47(5):1023 - 30. <https://doi.org/10.1016/j.molimm.2009.11.016>.
- Rasa SMM, Annunziata F, Krepelova A, Nunna S, Omrani O, Gebert N, et al. Inflammaging is driven by upregulation of innate immune receptors and systemic interferon signaling and is ameliorated by dietary restriction. *Cell Rep* 2022;39(13):111017. <https://doi.org/10.1016/j.celrep.2022.111017>.
- Zeng Y, Nie C, Min JX, Liu XM, Li MM, Chen HS, et al. Novel loci and pathways significantly associated with longevity. *Sci Rep* 2016;6: 21243. <https://doi.org/10.1038/srep21243>.

SUPPLEMENTARY MATERIAL

SUBJECTS AND GENETIC DATA

Changzhou Cohort

The Changzhou cohort is an independent prospective cohort from the Wujin District in Changzhou City, Jiangsu Province, China (1). The institutional review boards of Nanjing Medical University approved the study protocol. We included 7,797 participants with comprehensive blood biomarkers information from the 2012–2013 follow-up. To ensure the reliability of our study, the following criteria were used to exclude participants (2–3). Finally, 7,584 individuals were included for further evaluation. Genetic data were available for 1,215 out of the 7,584 included participants, which were derived from our two previously published datasets (4–5). In this study, we performed quality controls and 1,215 samples with ~8.15 million variants were included in the following evaluation.

UK Biobank

The UK Biobank (UKB) cohort has recruited 502,461 individuals aged between 40 and 69 years in the UK between 2006 and 2010 (6). The UK Biobank Pharma Proteomics Project (UKB-PPP) is a sub-project of UKB. It conducted proteomic profiling on blood plasma samples from 54,306 participants using the Olink platform (7). A total of 41,083 UKB-PPP participants were included for aging-related protein analysis. This study was carried out under the auspices of application number 79151 inside the UK Biobank resource.

METHODS

Aging Measures Development and Evaluation

To construct aging measures, candidate biomarkers were considered based on previous reports, data availability and their correlation with CA. Twenty candidate biomarkers were included in the first round of consideration. We used the capping method to address outliers beyond the 2.5% and 97.5% quantiles in blood biochemical biomarkers, as well as the 0.5% and 99.5% quantiles in body measurement indexes. Lasso regression was used to further screen for the candidate biomarkers. The optimal model was selected using a Cox proportional hazard elastic net model for mortality based on 10-fold cross-validation. Before calculating aging measures, non-normally distributed biomarkers were log-transformed.

Following the candidate biomarkers selection, we calculate three promising aging measures: KDM-BA, PhenoAge, and HD. All these measures can be constructed by the above biomarkers. Detailed methods for constructing PhenoAge, KDM-BA and Log(HD) can be found in a previous study (8).

The quantitative ability of these aging measures was evaluated with ROC curves. According to the AUC, we selected the best aging measure with the highest AUC value. Subsequently, we estimated the aging acceleration phenotype by the residual of these aging measures after subtracting the effect of CA using a linear regression model. A positive value of aging acceleration indicates a person appears to be physiologically older than expected and vice versa.

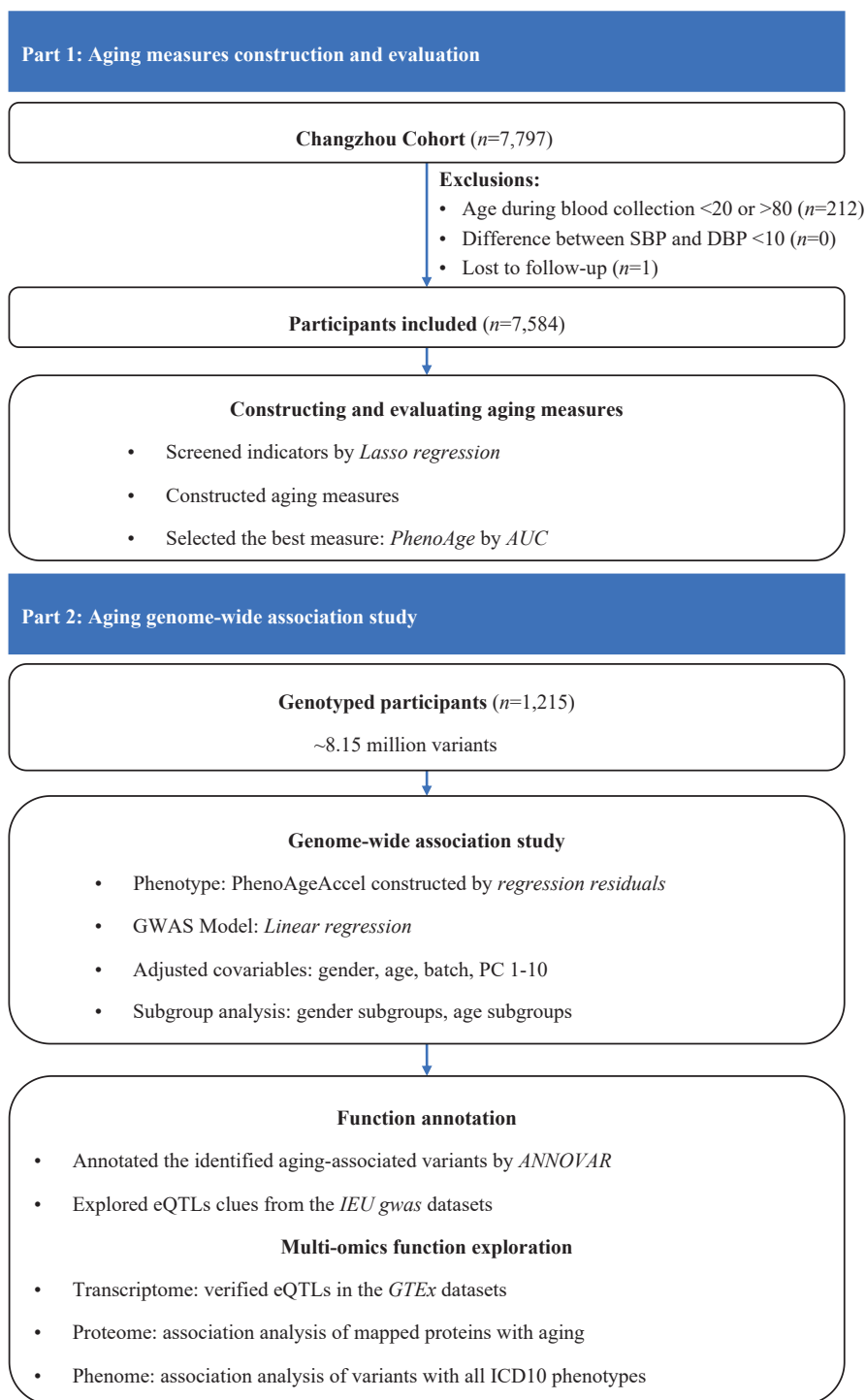
Genome-Wide Association Analysis and Meta-Analysis

PLINK2.0 was used to perform linear regression analysis on biological aging acceleration (9), including the first 10 principal components with age and sex as covariates. Batch was additionally adjusted in the analysis of the NJMU Array. Furthermore, a meta-analysis was performed to combine the two GWAS of NJMU Array and NJMU WGS using an inverse-variance based method in random-metal (10). In the meta-analysis, additional random effects (DerSimonian-Laird estimator) were used for estimating variants with significant heterogeneity ($P\text{-het}<0.05$ or $I^2>50$) and fixed effects were used for estimating other variants. In the identification of aging-associated loci, we set significant level at $P<1.00\times 10^{-5}$ and considered leading SNPs with a threshold of $r^2\geq 0.6$.

DATA AND CODE AVAILABILITY

The variation data reported in this paper have been deposited in the Genome Variation Map (GVM) in National

Genomics Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences and China National Center for Bioinformation, under accession number GVP000035 (11–12). The codes for the development of the three Aging measures were provided at <https://github.com/dayoonkwon/BioAge>. In addition, detailed models, codes for the three measures in the Changzhou cohort are available from the corresponding author by request.



SUPPLEMENTARY FIGURE S1. Workflow chart of study design.

Abbreviation: SBP=systolic blood pressure; DBP=diastolic blood pressure; AUC=area under curve; GWAS=genome-wide association study; PC=principal component; eQTLs=expression quantitative trait locus; ICD=international classification of diseases.

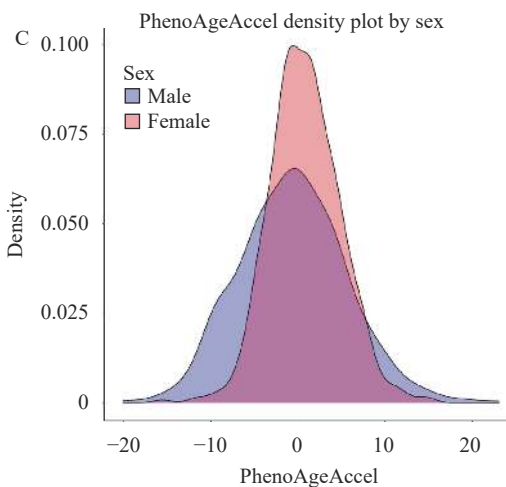
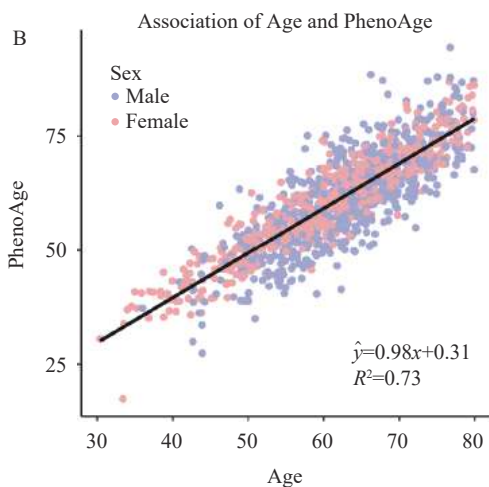
A Formula of PhenoAge:

$$\begin{aligned}
 xb_male = & -12.813 - 0.018 * \text{weight} - 0.007 * \text{waist} + 0.042 * \text{hip} \\
 & + 0.001 * \text{SBP} + 0.014 * \text{DBP} - 0.010 * \text{Hb} + 0.168 * \text{WBC} \\
 & - 0.003 * \text{PLT} + 1.389 * \log(\text{FBG}) - 0.555 * \log(\text{TBIL}) \\
 & + 0.221 * \log(\text{DBIL}) - 0.120 * \text{ALB} - 0.619 * \log(\text{ALT}) \\
 & + 1.047 * \log(\text{AST}) + 0.184 * \text{TC} - 0.364 * \log(\text{TG}) \\
 & - 0.769 * \log(\text{HDL-C}) + 0.000,04 * \text{BMI}
 \end{aligned}$$

$$\text{PhenoAge}_{male} = 119.540 + \frac{\log(-0.029 * \log(\exp(\frac{-33.395 * \exp(xb_male)}{0.029}))))}{0.088}$$

$$\begin{aligned}
 xb_female = & -16.088 - 0.006 * \text{weight} + 0.057 * \text{waist} - 0.015 * \text{hip} \\
 & - 0.001 * \text{SBP} + 0.007 * \text{DBP} - 0.002 * \text{Hb} - 0.010 * \text{WBC} \\
 & + 0.003 * \text{PLT} + 0.677 * \log(\text{FBG}) + 0.170 * \log(\text{TBIL}) \\
 & + 0.109 * \log(\text{DBIL}) - 0.047 * \text{ALB} - 0.264 * \log(\text{ALT}) \\
 & + 0.699 * \log(\text{AST}) - 0.197 * \text{TC} + 0.057 * \log(\text{TG}) \\
 & + 0.881 * \log(\text{HDL-C}) - 0.110 * \text{BMI}
 \end{aligned}$$

$$\text{PhenoAge}_{female} = 118.189 + \frac{\log(-0.028 * \log(\exp(\frac{-30.860 * \exp(xb_female)}{0.029}))))}{0.103}$$



SUPPLEMENTARY FIGURE S2. The distribution of PhenoAge and PhenoAgeAccel. (A) Formulas of PhenoAge by gender. (B) Scatter plot of age versus PhenoAge. (C) Density plot of PhenoAgeAccel by sex.

Abbreviation: ALB=albumin; ALT=alanine aminotransferase; AST=aspartate aminotransferase; BMI= body mass index; DBIL=direct bilirubin; DBP=diastolic blood pressure; FBG=fasting blood glucose; Hb=hemoglobin; HDLC=high-density lipoprotein cholesterol; PLT=platelet count; SBP=systolic blood pressure; SD=standard deviation; TC=total cholesterol; TBIL=total bilirubin; TG=triglycerides; WBC=white blood cell count.

REFERENCES

- Chen W, Lu F, Liu SJ, Du JB, Wang JM, Qian Y, et al. Cancer risk and key components of metabolic syndrome: a population-based prospective cohort study in Chinese. *Chin Med J (Engl)* 2012;125(3):481 – 5. <https://doi.org/10.3760/cma.j.issn.0366-6999.2012.03.015>.
- Liu ZY. Development and validation of 2 composite aging measures using routine clinical biomarkers in the Chinese population: analyses from 2 prospective cohort studies. *J Gerontol A Biol Sci Med Sci* 2021;76(9):1627 – 32. <https://doi.org/10.1093/gerona/glaa238>.
- Qi SF, Cao YJ, Wang HJ, Zhang B, Yan J, Tian QB. Associations of carbohydrate intake with new-onset hypertension subtypes: results from the China health and nutrition survey (2000–2011). *Front Nutr* 2021;8:728774. <https://doi.org/10.3389/fnut.2021.728774>.
- Dai JC, Lv J, Zhu M, Wang YZ, Qin N, Ma HX, et al. Identification of risk loci and a polygenic risk score for lung cancer: a large-scale prospective cohort study in Chinese populations. *Lancet Respir Med* 2019;7(10):881 – 91. [https://doi.org/10.1016/S2213-2600\(19\)30144-4](https://doi.org/10.1016/S2213-2600(19)30144-4).
- Wang C, Dai JC, Qin N, Fan JY, Ma HX, Chen CC, et al. Analyses of rare predisposing variants of lung cancer in 6,004 whole genomes in Chinese. *Cancer Cell* 2022;40(10):1223 – 39.e6. <https://doi.org/10.1016/j.ccell.2022.08.013>.
- Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015;12(3):e1001779. <https://doi.org/10.1371/journal.pmed.1001779>.
- Sun BB, Chiou J, Traylor M, Benner C, Hsu YH, Richardson TG, et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature* 2023;622(7982):329 – 38. <https://doi.org/10.1038/s41586-023-06592-6>.
- Kwon D, Belsky DW. A toolkit for quantification of biological age from blood chemistry and organ function test data: BioAge. *GeroScience* 2021;43(6):2795 – 808. <https://doi.org/10.1007/s11357-021-00480-5>.
- Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* 2015;4:7. <https://doi.org/10.1186/s13742-015-0047-8>.
- Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010;26(17):2190 – 1. <https://doi.org/10.1093/bioinformatics/btq340>.
- CNCB-NGDC Members and Partners. Database resources of the national genomics data center, China national center for bioinformation in 2022. *Nucleic Acids Res* 2022;50(D1):D27 – 38. <https://doi.org/10.1093/nar/gkab951>.
- Li CP, Tian DM, Tang BX, Liu XN, Teng XF, Zhao WM, et al. Genome variation map: a worldwide collection of genome variations across multiple species. *Nucleic Acids Res* 2021;49(D1):D1186 – 91. <https://doi.org/10.1093/nar/gkaa1005>.