Psychological factors substantially contribute to biological aging: evidence from the aging rate in Chinese older adults

Fedor Galkin[1], Kirill Kochetov[1], Diana Koldasbayeva[1], Manuel Faria[2], Helene H. Fung[3], Amber X. Chen[3], Alex Zhavoronkov[1,4,5]

[1]Deep Longevity Limited, Hong Kong, People's Republic of China
[2]Department of Psychology, Stanford University, Stanford, CA 94305, USA
[3]Department of Psychology, The Chinese University of Hong Kong, Hong Kong, People's Republic of China
[4]Insilico Medicine, Hong Kong, People's Republic of China
[5]Buck Institute for Research on Aging, Novato, CA 94945, USA

Correspondence to: Fedor Galkin; **email:** fedor@deeplongevity.com
Keywords: psychological aging, lifespan psychology, aging clocks, longevity
Received: April 4, 2022   **Accepted:** August 23, 2022   **Published:** September 27, 2022

ABSTRACT

We have developed a deep learning aging clock using blood test data from the China Health and Retirement Longitudinal Study, which has a mean absolute error of 5.68 years. We used the aging clock to demonstrate the connection between the physical and psychological aspects of aging. The clock detects accelerated aging in people with heart, liver, and lung conditions. We demonstrate that psychological factors, such as feeling unhappy or being lonely, add up to 1.65 years to one's biological age, and the aggregate effect exceeds the effects of biological sex, living area, marital status, and smoking status. We conclude that the psychological component should not be ignored in aging studies due to its significant impact on biological age.

INTRODUCTION

Aging clocks are statistical models that enable measurements of biological age, as opposed to chronological age. While the latter is determined by one's date of birth, the former depends on the intensity of aging processes and can be affected by genetics, life choices, and the environment. Most commonly, such aging clocks are regressors, trained to predict a person's chronological age based on a vector of input parameters, such as clinical blood test results, gene expression levels, or DNA methylation intensities [1–5]. Recently, aging clocks linked to frailty and mortality by design have been trained [6, 7].

Based on the pace of aging, people may be divided into slow, normal, and fast agers. Several studies have indicated that increased biological age is associated with higher all-cause mortality, vulnerability to

infection, and a number of non-communicable diseases [1, 2, 6, 8]. Conversely, a slow aging rate is associated with better health compared to age peers. Since aging clocks can detect various health conditions as accelerated aging, this technology holds the promise of becoming an all-purpose diagnostic and prognostic tool to benefit multiple healthcare-adjacent industries, such as insurance and nutrition. Aging clock studies are essential for the development of longevity therapeutics and geroprotectors (anti-aging medication). Aided by AI algorithms, aging clocks have the potential to become a staple of the modern drug design pipeline [5, 9–11]. This synthetic technology, which combines data-driven approaches, aging clocks, and deep learning, is being actively patented for use cases in global healthcare systems and the pharmaceutical industry [3, 12–17]. Longitudinal biogerontological studies have highlighted the ways to manipulate the pace of aging via exercise, diet, or

dietary supplements [18–21]. Some compounds extensively studied in this context show great promise on a global scale [22].

The impact of psychology on biological age, however, is a painfully understudied subject. In the 20th century, several personality development theories have been formed, such as Freud's psychosexual development and Erikson's psychosocial development [23–25]. While Freudian psychology is more focused on early-life development, the stages defined by Erikson describe personality evolution from childhood until old age. In the last three decades, many theories of aging have been proposed (see [26] for a review). In particular, in 1990s, Laura Carstensen proposed a theory, called socioemotional selectivity theory, that describes how individuals shift their motivations across adulthood [27]. The theory specifies how psychological factors such as emotions can affect aging. However, while human psychology is known to go through certain changes throughout the life span, few studies have attempted to describe these changes in the terms of biogerontology (i.e., using aging clocks) or determine their effect on physical health [28, 29]. A limited number of studies pointed out that psychological factors, such as stress [30], loneliness [31], mental health [32, 33] and negative perceptions of aging [34], could have significant influences on aging clocks derived from physical parameters. Other socio-psychological factors that might accelerate aging clocks include negative life events [35], modern lifestyles [19, 31, 36], and low socio-economic status [37].

A study involving baboons linked social stress caused by competition to reversible aging acceleration [38]. Similar findings have been observed for humans: lifetime accumulated stress can be responsible for up to 3.6 years of extra biological age [39]. These findings suggest that the effectors of stress in organismal aging are glucocorticoid hormones, known to be involved in memory and cognition. Stress-induced epigenetic changes by glucocorticoid signaling colocalize with aging-associated DNA methylation sites, which may promote arteriosclerosis and other health conditions. Other effectors of stress-related aging include oxidative damage and telomere shortening, which may add up to ten years of extra biological age [40]. These stress-induced deviations from the normal aging pattern may persist for years, as shown in studies of childhood trauma and post-traumatic stress disorder [30, 41, 42]. Importantly, there is reason to believe that acting upon these psychosocial domains may confer a longevity benefit, as some studies have found that intense relaxation and emotional coping skills modulate biological age metrics [43, 44].

The link between physical aging and psychology is not unidirectional, as some studies have indicated. People with a high pace of aging are more likely to develop psychopathologies in later life [33], whereas epigenetic age at birth has been implicated in childhood temperament development [45]. In most studies, however, uncovering the causal links is an impossible task due to their cross-sectional design. The task is even further complicated by the inherent ambiguity of the term "psychological aging." Psychological age may refer to the mental state most typical for people of a certain age group. At the same time, psychological age is something that corresponds to one's subjective perception of the passing time, framed within their experiences and expectations. Interestingly, psychological age may be associated with opposite psychological trends, depending on the definition. For example, people with prominent neurotic behavior are more prevalent among chronologically younger people, yet neurotic people perceive themselves to be older than they really are [46].

In this article, we continue exploring psychology in the context of aging. We have developed a novel aging clock based on blood panels from the China Health and Retirement Longitudinal Study (CHARLS) dataset in order to define biological age in a cohort of Chinese adults. We represent biological age as a function of various psychological and social variables to measure their effects on the pace of aging. We also showed that AI-based psychological aging clocks are viable across cultures, building on prior work based on a U.S. dataset (i.e., an individualist culture) and now a Chinese dataset (i.e., a collectivist culture) [46]. Finally, we demonstrate the biological relevance of the aging clock by analyzing its error distribution in people with certain diseases.

CHARLS is a nationwide study of Chinese population older than 45 years that features information on respondents' social and economic statuses, health history, biometrics, and blood panels [47]. Since its inception, CHARLS has been used to study the effects of nutrition, social bonds and economic status on healthy aging [48–51]. These insights have been used to measure the impact of social and economic policies on the Chinese national longevity potential [52–58].

Among East Asian countries, China has the lowest portion of successful agers: 15.7% [59]. A successful ager is a person over 65 years old with no major disabilities and with normal cognitive function and social engagement. For comparison, the portion of successful agers among older adults is 29.2% in Japan and 25.5% in Korea. Yet, due to the large population of China, the number of persons over 65 years old in China is higher than such number in the entire Europe

[60]. Understanding aging in China can thus provide important insights on aging in the world.

In this article, we demonstrate that mental and physical health are interlinked and both can be described by a single aging clock trained on blood test data (Figure 1). We also demonstrate that the detrimental impact of low psychological well-being is of the same magnitude as serious diseases and smoking.

RESULTS

Blood-based age predictor

We used a combination of 16 blood biomarkers, seven biometric parameters, and the biological sex of CHARLS participants (see "Model Training", Figure 2A) to create a deep neural network (DNN) age predictor. The model outperformed mean age assignment in both cross validation (CV) and test sets, as well as the discovery set comprised of participants with a history of serious health conditions (Figure 3, Table 1). To be more specific, we report a mean absolute error (MAE) of 5.68 years in a collection of 4451 test samples from generally healthy people.

To illustrate the benefit of the deep learning approach, we also trained an elastic net (EN) using the same set division and variables and obtained a model displaying an MAE of 6.19 years for the same 4,451 people.

Based on the results shown in Supplementary Figure 1, the five most important features of our predictor are systolic and diastolic blood pressure, cystatin C, BMI and spirometry. We used permutation feature importance technique to measure feature weights for
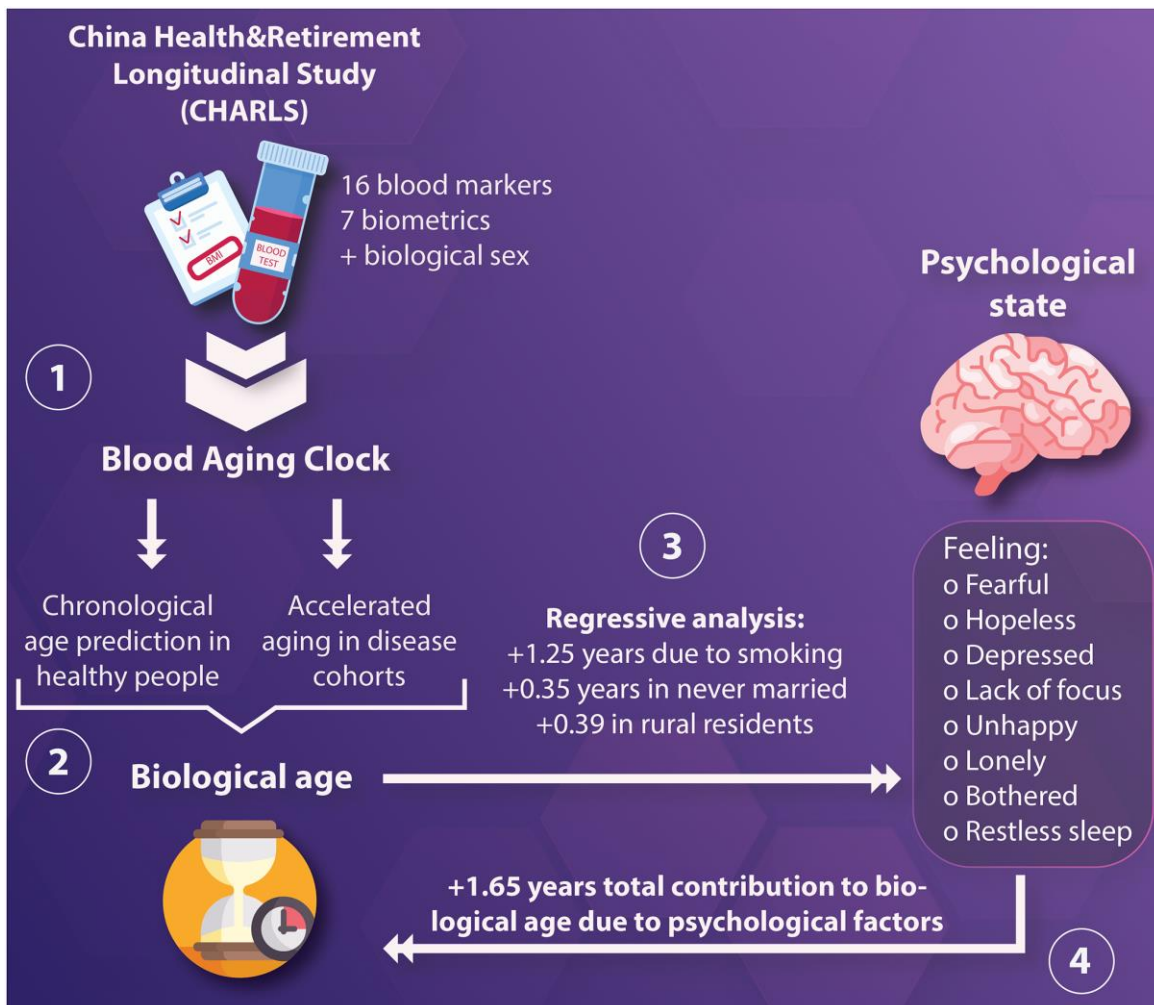

Figure 1. (**1**) Blood and biometric data, in addition to biological sex, were used to construct a neural network aimed to predict chronological age; (**2**) The age predicted by the model, hereby denoted as "biological age" was tested in healthy and ill participants to identify conditions interpreted as accelerated aging; (**3**) Then, regressive analysis was performed using an elastic net to quantify the total contribution of demographic, lifestyle, and psychological factors, hereby denoted as "psychological state," to biological age; (**4**) The weight of each variable was understood as age acceleration, with the aggregate effect of one's psychological state being able to accelerate biological aging by 1.65 years.

age prediction. The permutation feature importance (PFI) for a specific feature is defined to be the decrease in MAE when this specific feature column is randomly shuffled [61].

Biological relevance of the predictor

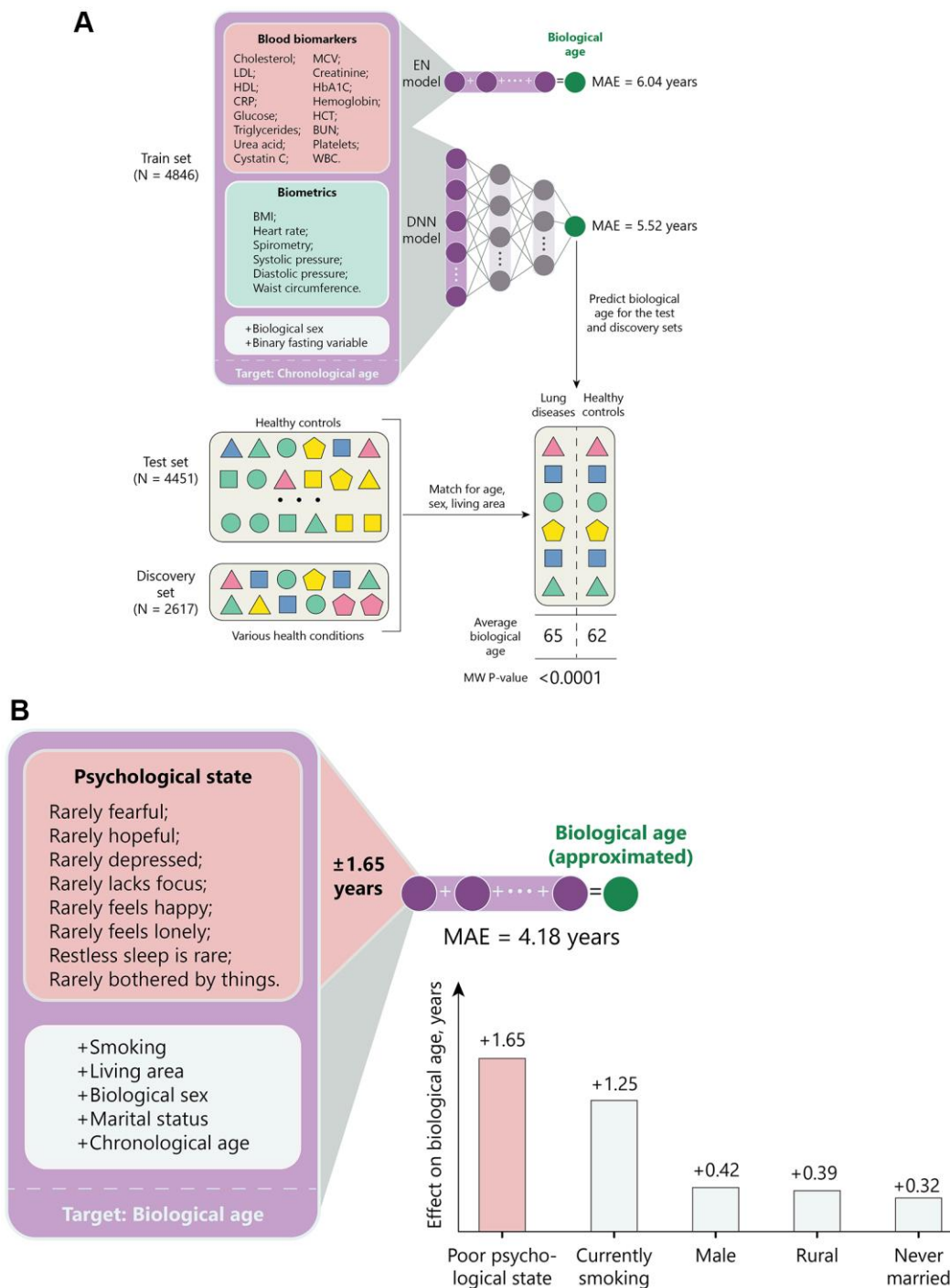To verify that the predictions of the new deep learning aging clock represented the pace of aging, we

Figure 2. Graphic description of the study. (**A**) 4,846 entries from CHARLS were used to train a deep neural network regressor of chronological age, whose performance was compared to an elastic net model. The age predicted by a regressor is further referenced as "biological age". Biological age of healthy (test set) participants and participants with a history of serious diseases (discovery set) was compared to identify conditions that are interpreted as accelerated aging by the model. To control for confounders, prediction averages were compared in matched cohorts. (**B**) The effect of psychological and other factors has been compared using elastic net. All variables, except chronological age, were converted to binary to predict biological age returned by the neural network described in panel (**A**). The weight of each variable was interpreted as age acceleration.

Table 1. Deep neural network prediction of the chronological age of CHARLS participants both in cross-validation and test sets.

	Metric	Cohort		
		CV	**Test**	**Discovery**
Mean age assignment	**MAE, years**	8.38	7.11	7.32
Elastic net	**MAE, years**	6.04	6.19	8.06
Neural network	**MAE, years**	5.52	5.68	5.94
	MAPE, years	9.88	9.84	10
	R^2	0.51	0.43	0.44
	***N*, people**	4846	4451	2617

The baseline model is mean age assignment. Abbreviations: CV: cross-validation; MAE: mean absolute error; MAPE: mean absolute percentage error; R^2: coefficient of determination.

investigated whether people with aging-associated diseases were predicted to be older than healthy people.

The discovery set defined before training contained people with a history of cancer, heart attack, stroke, liver disease, or lung disease, while all participants in the test set had none of these conditions. To minimize the confounding effect of other variables, we analyzed isolated conditions (i.e., no comorbidities) juxtaposed against a subsample of healthy people with matching age, sex, and living area distributions.

People with a history of stroke, liver disease, and lung conditions were predicted to be significantly older (*P* value < 0.0005, *U* test) than their healthy counterparts (Table 2). However, the mean effect of these conditions on the predicted age did not exceed $+1.50$ years, which is below the model's MAE of 5.56 years.

No significant differences in age predictions were detected for people with a history of heart disease or cancer when compared to healthy controls. It should be noted, however, that cancer is the least frequent condition in CHARLS and the lack of significance may also have been caused by insufficient sample sizes.

Effects of psychological states on the pace of aging

To compare the effect of multiple factors on biological age, we used EN to regress aging clock's predictions as a function of demographic and psychological variables (Figure 2B). All factors included in the analysis (except for being widowed) showed significant effects on the pace of aging (Table 3). The three largest effects on the aging rate belonged to the following binary factors: smoking ($+1.25$ years), being currently married (-0.59 years), and suffering from restless sleep ($+0.44$ years). However, if the eight psychological variables are considered as a scale representing psychological well-being, being on its lower end has an effect of accelerating aging by 1.65 years.
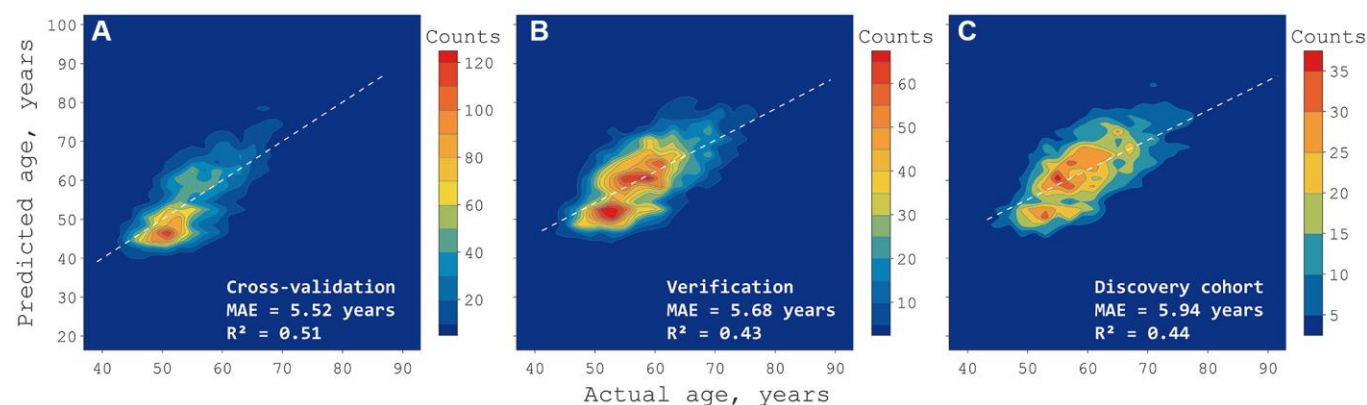

Figure 3. Density map of deep neural network predictions in three sets. The training set (**A**) was used to tune the neural network in a cross-validated manner, healthy samples from the test set (**B**) were used for verification, and the discovery set (**C**) contains only people with health conditions to explore cases of accelerated aging. The color bars mark continuous areas containing the same number of people. MAE stands for mean absolute error, R^2 stands for coefficient of determination.

Table 2. Differences in deep neural network age predictions between afflicted people in the discovery cohort and healthy people in the test cohort.

Disease	Afflicted error, years	Healthy error, years	Delta, years	*N*, people	Median age, years	*P* value
Heart disease	0.427	0.459	−0.03	873	63	0.873
Cancer	−0.333	0.74	−1.07	73	59	0.225
Stroke	1.557	0.068	1.49	115	65	*0.0156
Lung disease	1.869	−0.336	2.2	782	63	*<0.0001
Liver disease	0.138	−0.829	0.97	280	59	*0.0166

The comparison has been carried out in samples matched for age, sex, and living area (rural or urban). "*" marks the diseases with significantly accelerated pace of aging (*P* value < 0.05, Mann-Whitney *U* test).

Table 3. Biological age obtained from the deep neural network predictor was regressed as a function of chronological age and additional features to estimate their effect on the pace of aging.

Type	Variable	Coefficient	Standard deviation (σ)
Demographic	**Chronological age, years**	0.49	0.00
	Is male	0.42	0.02
Social	**Is married**	−0.59	0.04
	Is widowed	0.27	0.09
	Is rural	0.39	0.01
Psychological	**Rarely bothered by things**	0.09	0.03
	Rarely lacks focus	−0.05	0.02
	Rarely depressed	−0.09	0.02
	Rarely hopeful	0.28	0.01
	Rarely fearful	−0.29	0.02
	Restless sleep is rare	−0.44	0.01
	Rarely feels happy	0.35	0.02
	Rarely feels lonely	0.05	0.03
Smoking	**Currently smoking**	1.25	0.02
	Intercept	28.37	0.02
Metrics			
	MAPE (CV), %	7.57	
	MAPE (Test), %	6.37	
	MAE (CV), years	4.18	
	MAE (Test), years	4.09	
	N (CV), people	4846	
	N (Test), people	4451	

All variables, except for chronological age, are binary. All coefficients, except for "Is widowed", are significant within 3σ, as estimated with ten random cross-validation splits. Abbreviations: CV: cross-validation; MAE: mean absolute error; MAPE: mean absolute percentage error.

DISCUSSION

In this article, we have demonstrated a novel blood aging clock developed using only Chinese data from CHARLS. Previous installments of BloodAge have been validated in Russian, Canadian, European, and Korean populations [1, 62]. The same technology has also been utilized in numerous settings, including mortality-linked datasets [1, 8, 63]. In these studies, BloodAge has demonstrated strong population-specific effects, with some nations or races showing significantly different error distributions. Thus, to ensure that biological age can be accurately estimated in the Chinese population, we trained a new DNN with CHARLS data.

While CHARLS contains an abundance of medical and demographic profiles, blood panels are available in two of its waves. We chose the 2015 wave due to a larger number of measured blood parameters, yet it is still below the number of biomarkers available in other public datasets, such as NHANES. While the original BloodAge was trained on blood panels with 46 biomarkers from over 60,000 people, CHARLS offers only 16 blood parameters [62]. Our initial experiments with this limited variable space failed to produce an accurate predictor of chronological age; hence, we augmented the training data with biometric data, such as BMI, waist circumference, and blood pressure. In total, we were able to train a DNN with 24 variables and 4,846 healthy samples to reach an accuracy of MAE = 5.68 years, which is almost identical to 5.55 years in the original BloodAge publication from 2016 [62].

We proceeded to validate the output of the DNN as a general health marker for a subset of CHARLS subjects with known health issues (Table 2). Our hypothesis was that people with aging-related diseases should have DNN prediction errors shifted to the right, thus indicating the increased pace of aging. To correct for confounders, we compared the errors within subsets of people afflicted by only one condition and healthy subsets matched for sex, age, and living area. Among the five conditions, significant age acceleration was detected for the participants with a history of stroke, liver, and lung conditions. At the same time, people who had suffered from heart attack or cancer had a DNN error distribution indistinguishable from healthy controls. It should be noted that the lack of significance for cancer may have been caused by the heterogeneity of oncological diagnoses, as well as seemingly anti-aging properties of malicious cell lines (such as proliferation) [64].

The health conditions accelerated the aging processes by 1–2 years. Given the high average age of the afflicted participants, this seemingly minor age acceleration can lead to a significant increase in mortality risk. The effect of age acceleration on survival, however, cannot be assessed in the absence of mortality-linked data. Altogether, we consider that the described aging clock possesses the property of disease-relevance that makes its prediction error representative of a participant's health.

Since accelerated aging was demonstrated in cohorts matched for age, sex, and living area, this effect offers additional predictive value compared to chronological age. However, an increased pace of aging may have multiple causes and thus lacks specificity. This drawback may be eradicated in future ensemble models that classify inputs in disease-specific categories and return biological age as a metric for severity. Among the blood biomarkers included in the DNN aging clock, cystatin C was the second most important, ranking below systolic pressure. Cystatin C is a cysteine protease used as a biomarker of renal function and its elevation is a major risk factor in all-cause mortality in the Chinese older adults [65]. Some other studies of mental and physical health have linked abnormal cystatin C levels to new-onset depressive symptoms [66].

Finally, we quantified the effect of psychological features on the pace of aging with the EN approach. The biological age estimate, obtained with the DNN, was regressed as a function of chronological age, sex, marital status, living area, and several mental state variables (Table 3). We deliberately chose not to include biometric features used by the DNN since they are implicitly present in the biological age target itself.

To make the interpretation easier, we converted all variables, except for chronological age, into a binary form. Since CHARLS contains mostly middle-aged or older participants, the EN features a high intercept (28.37 years), and its results may not be applicable to younger datasets. According to the EN interpretation, the biological age of any person is almost fully comprised of the intercept and chronological age. Nonetheless, most other features had significant non-zero coefficients, as estimated with ten-fold CV. In aggregate, the psychological state of a participant contributes 1.65 years of biological age if all the associated variables assume the values indicating lower well-being. For example, feeling unhappy will contribute 0.35 years and having no trouble with sleep will detract 0.44 years.

The survey collection method may have affected the significance of the psychological factors. The participants were asked to rank the frequency of certain

feelings or issues during the past week. We believe that in more thorough assessments of psychological well-being, the contribution of such features would be larger.

The impact of demographic features is of a similar magnitude as that of psychological factors. Among them, marital status had the greatest effect on the pace of aging. Currently married participants were predicted to be 0.6 years younger than the never married, while widowed participants were predicted to be 0.3 years older. Previous CHARLS-based studies showed that widowhood was associated with signs of accelerated aging, particularly memory decline [50, 51]. However, the data did not allow us to study marital satisfaction as an aging factor, although we hypothesize that it positively correlates with the other quantified psychological effects and thus is beneficial for the pace of aging [67, 68]. We also interpret this finding as the importance of family support in older age. Both instrumental and emotional support from the spouse and offspring significantly improve life satisfaction, and thus decrease the aging rate [69]. At the same time, feeling lonely had a nonsignificant effect on the pace of aging, indicating that family support is much more important for successful aging than weak ties with low emotional intensity [49]. However, some studies have reported that the Chinese older adults (except the oldest old) display slightly lower life satisfaction scores if they still live with their children [70]. Altogether, these findings show that family support may improve longevity potential only in combination with a reasonable amount of autonomy, although there might be cross-cultural differences.

Additionally, people living in rural areas are predicted to be 0.4 years older than those living in urban areas. The adverse effect of rural living has been recently shown in a study of multimorbidity in China and has been attributed to "lower socio-economic status, poorer social services, and lower access to quality medical services" [71]. Malnutrition and physical labor are proposed as the key contributors to the increased prevalence of certain conditions, such as osteoporosis and arthritis. These factors are also contributing to poorer mental health in the older Chinese rural population, who are four to six times more likely to be depressed that the urban population [72, 73].

The new DNN successfully represents both mental and physical health decline as cases of accelerated aging. This property makes this new aging clocks a great tool for developing anti-aging therapeutics. The impact of the interventions targeting the body may be assessed with the original DNN, while the efficiency of the interventions targeting one's mental state may be estimated with the EN summarized in Table 3. Both modes of this new aging clock may be combined to construct personalized regimens reducing the pace of aging. The utility of this aging clock may be further increased by combining it with a SOM-based recommendation engine we presented earlier this year, FuturSelf [74]. This model of human psychology is available at https://futurself.ai/ and provides personalized recommendations aimed to maximize users' future mental well-being.

We are planning to validate this approach in other public datasets before releasing the model. However, the exact wordings for the psychological assessment in CHARLS make the model less compatible with Western datasets, such as UK Biobank or NHANES.

Altogether, we have demonstrated that the pace of aging is significantly associated with psychological features. The observed contribution of one's mental state to one's biological age is significant and is comparable to the effect of smoking. Thus, promoting mental health may be considered a potential anti-aging intervention with possible benefits at par with more tangible, physical therapeutic approaches.

CONCLUSION

Using data from the Chinese CHARLS database, we have demonstrated that organismal aging is not only determined by physical factors but also, to a certain degree, affected by mental state and social status. We interpreted biological age as a proxy for the general state of health and show that positive feelings (happiness, hope, safety) have a significant impact on the former.

The study findings further support the necessity of companionship and a psychologically pleasant environment for healthy longevity.

METHODS

Data overview

This study features data from CHARLS collected in waves between 2008 and 2018, with two waves (2011, 2015) featuring blood panels. We used data from the 2015 wave due to a larger number of available blood biomarkers. In total, 11,914 entries were imported from CHARLS and divided into training ($N = 4,846$), test ($N = 4,451$), and discovery ($N = 2,617$) sets. Samples with missing sex or outlier values in the predictor variables were excluded. The training set comprised generally healthy people. The test set aggregated healthy people and people with the following health issues: hypertension ($N = 1,277$), arthritis ($N = 915$), dyslipidemia ($N = 278$), kidney

disease (N = 207), and diabetes (N = 115). The discovery set contained only samples with a history of cancer (N = 73), chronic heart condition or heart attack (N = 873), stroke (N = 115), lung disease (N = 782) or liver conditions (N = 280).

The distributions of all CHARLS variables mentioned in this article are available in Supplementary Tables 1, 2 and Supplementary Figure 2.

Model training

To train the predictor, we used 16 blood panel variables: total cholesterol, LDL cholesterol, HDL cholesterol, C-reactive protein, glucose, triglycerides, urea acid, cystatin C, MCV, creatinine, HbA1C, hemoglobin, hematocrit, BUN, platelet count, and WBC count. These variables were further complemented with biological sex, a binary variable indicating whether the person was fasting before blood drawing, and six biometric variables: systolic and diastolic pressure, BMI, waist circumference, heart rate, spirometry. In cases of CHARLS participants having multiple measurements for a biometric parameter, the average was used.

We used feed-forward DNN with more than three hidden layers for the age prediction task. We treated age prediction as a regression task, so linear activation function was used on the DNN output. All features were normalized and mapped to the 0–1 space using MinMaxScaler from the scikit-learn Python package in order to prevent gradient explosion and make all features equally important for the DNN [75, 76]. The MAE loss function between real and predicted age was used as an objective and minimized during training. Grid search was applied to find the best model parameters in terms of an objective metric (MAE). The best performing model had the following parameters: 5 hidden layers with 256 neurons each, an exponential linear unit (ELU) activation function after each layer, dropout with 35% probability, and L2 with coefficient 1e-5 after each layer for the purpose of regularization [77, 78]. All models were trained until convergence on the validation set (when loss on validation set stopped decreasing).

The DNN was trained using five-fold CV to compensate for overfitting and obtain more robust performance metrics. We used Python 3.9 along with the Keras library (https://keras.io/) with a Tensorflow backend (https://www.tensorflow.org), to build and train the DNN.

The metrics used to measure the performance of the models were MAE, mean absolute percentage error (MAPE), and coefficient of determination (R^2). These metrics are defined as follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|Age_{true,i} - Age_{predicted,i}\right|,$$

where N is the total number of samples

$$MAPE = \frac{100}{N}\sum_{i=1}^{N}\frac{1}{Age_{true,i}}\left|Age_{true,i} - Age_{predicted,i}\right|,$$

where N is the total number of samples

$$R^2 = 1 - \frac{\sum_i Prediction\ error_i^2}{N \times D\left(Age_{true}\right)},\ \text{where } D \text{ is the variance}$$

The performance of our DNN model was compared against mean age assignment and an EN model trained using the ElasticNet method from the sklearn. linear_model module for Python. The EN was trained on the same training set and features as the DNN. The EN's hyperparameters were optimized using the following grid: {'alpha': [1e-5, 1e-4, 1e-3, 1e-2, 1e-1, 0.0, 1.0, 10.0, 100.0], 'l1_ratio': np.arange (0, 1, 0.01)}.

Association with health conditions

Age prediction errors were compared between people with a history of cancer, stroke, heart attack, liver disease, or lung disease from the discovery set and healthy participants from the test set. To correct for biases associated with cohort differences, we matched disease-associated subsamples with equally sized subsamples of the test set for age, sex, and living area.

The three meta variables were normalized to generate matched healthy cohorts with the nearest neighbor approach. The test set was sampled with replacement for each condition. The Mann-Whitney U-test was carried out for each condition to establish the significance of different aging rates.

The nearest neighbor search was implemented with NearestNeighbors from scikit-learn 0.24.2, and the U-test was implemented with mannwhitneyu from scipy 1.7.0 for Python 3.9.

Association with psychosocial factors

To establish the effect of multiple variables on the pace of aging, we used EN regression. The predictions of the deep learning aging clock were regressed as a linear combination of chronological age, sex, marital status, living area, smoking status, and eight psychological

attributes (Table 3). All variables, except for chronological age, were binarized.

The coefficients of the final EN regression represent the effect of each associated parameter on the pace of aging. Standard deviations of the coefficients were calculated based on ten iterations with different random seeds. The EN model was trained using the 6,511 training samples and verified with 2,786 samples from the test set.

Abbreviations

BMI: Body mass index; BUN: Blood urea nitrogen; CHARLS: China Health and Retirement Longitudinal Study; CV: Cross-validation; DNN: Deep neural network; EN: Elastic net; HbA1C: Glycated hemoglobin; HDL: High-density lipoprotein; LDL: Low-density lipoprotein; MAE: Mean absolute error; MAPE: Mean absolute percentage error; MCV: Mean corpuscular volume; NHANES: National Health and Nutrition Examination Survey; PTSD: Post-traumatic stress disorder; WBC: White blood cells.

AUTHOR CONTRIBUTIONS

AC – Writing (reviewing and editing), validation, data curation, AZ – Conceptualization, resources, supervision, writing (reviewing and editing), DK – Data curation, formal analysis, FG – Project administration, formal analysis, writing (original draft), methodology, visualization, HF – Conceptualization, resources, writing (reviewing and editing), supervision, KK – Software, methodology, formal analysis, MF – Writing (reviewing and editing), visualization.

CONFLICTS OF INTEREST

AZ, DK, FG, KK are affiliated with Deep Longevity, a subsidiary of a publicly-traded company, Endurance RP (HK:0575), which holds intellectual property in the area of deep aging clocks.

FUNDING

REFERENCES

1. Mamoshina P, Kochetov K, Putin E, Cortese F, Aliper A, Lee WS, Ahn SM, Uhn L, Skjodt N, Kovalchuk O, Scheibye-Knudsen M, Zhavoronkov A. Population Specific Biomarkers of Human Aging: A Big Data Study Using South Korean, Canadian, and Eastern European Patient Populations. J Gerontol A Biol Sci Med Sci. 2018; 73:1482–90.
https://doi.org/10.1093/gerona/gly005
PMID:29340580

2. Galkin F, Mamoshina P, Kochetov K, Sidorenko D, Zhavoronkov A. DeepMAge: A Methylation Aging Clock Developed with Deep Learning. Aging Dis. 2021; 12:1252–62.
https://doi.org/10.14336/AD.2020.1202
PMID:34341706

3. Aliper AM, Galkin F, Zavoronkovs A. Deep transcriptomic markers of human biological aging and methods of determining a biological aging clock. US20190034581A1. 2018.
https://patents.google.com/patent/US20190034581A1/en.

4. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol. 2013; 14:R115.
https://doi.org/10.1186/gb-2013-14-10-r115
PMID:24138928

5. Galkin F, Mamoshina P, Aliper A, de Magalhães JP, Gladyshev VN, Zhavoronkov A. Biohorology and biomarkers of aging: Current state-of-the-art, challenges and opportunities. Ageing Res Rev. 2020; 60:101050.
https://doi.org/10.1016/j.arr.2020.101050
PMID:32272169

6. McCrory C, Fiorito G, Hernandez B, Polidoro S, O'Halloran AM, Hever A, Ni Cheallaigh C, Lu AT, Horvath S, Vineis P, Kenny RA. GrimAge Outperforms Other Epigenetic Clocks in the Prediction of Age-Related Clinical Phenotypes and All-Cause Mortality. J Gerontol A Biol Sci Med Sci. 2021; 76:741–9.
https://doi.org/10.1093/gerona/glaa286
PMID:33211845

7. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, Hou L, Baccarelli AA, Stewart JD, Li Y, Whitsel EA, Wilson JG, Reiner AP, et al. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY). 2018; 10:573–91.
https://doi.org/10.18632/aging.101414
PMID:29676998

8. Galkin F, Parish A, Bischof E, Zhang J, Mamoshina P, Zhavoronkov A. Increased Pace of Aging in COVID-Related Mortality. Life (Basel). 2021; 11:730.
https://doi.org/10.3390/life11080730
PMID:34440474

9. Moskalev A, Anisimov V, Aliper A, Artemov A, Asadullah K, Belsky D, Baranova A, de Grey A, Dixit VD, Debonneuil E, Dobrovolskaya E, Fedichev P, Fedintsev A, et al. A review of the biomedical innovations for healthy longevity. Aging (Albany NY). 2017; 9:7–25.
https://doi.org/10.18632/aging.101163

10. Zhavoronkov A, Ivanenkov YA, Aliper A, Veselov MS, Aladinskiy VA, Aladinskaya AV, Terentiev VA, Polykovskiy DA, Kuznetsov MD, Asadulaev A, Volkov Y, Zholus A, Shayakhmetov RR, et al. Deep learning enables rapid identification of potent DDR1 kinase inhibitors. Nat Biotechnol. 2019; 37:1038–40. https://doi.org/10.1038/s41587-019-0224-x PMID:31477924

11. Mamoshina P, Vieira A, Putin E, Zhavoronkov A. Applications of Deep Learning in Biomedicine. Mol Pharm. 2016; 13:1445–54. https://doi.org/10.1021/acs.molpharmaceut.5b00982 PMID:27007977

12. Aliper AM, Galkin F, Zavoronkovs A. Aging markers of human microbiome and microbiomic aging clock. US20200075127A1. 2020. https://patents.google.com/patent/US20200075127A1/en.

13. Aliper AM, Zavoronkovs A, Ozerov I, Bozdaganyan ME, Artemov AV. Method of treating senescence with multi-stage longevity therapeutics. US11260078B2. 2022. https://patents.google.com/patent/US11260078B2/en.

14. Horvath S. Method to estimate the age of tissues and cell types based on epigenetic markers. EP3049535B1. 2021. https://patents.google.com/patent/EP3049535B1/en.

15. Zhang K, Hannum G, Ideker T, Friend SH, Guinney J. Methods for predicting age and identifying agents that induce or inhibit premature aging. US20150259742A1. 2015. https://patents.google.com/patent/US20150259742A1/en.

16. Galkin F, Kochetov KS, Mamoshina P, Zavoronkovs A. Methylation data signatures of aging and methods of determining a methylation aging clock. US20220005552A1. 2022. https://patents.google.com/patent/US20220005552A1/en.

17. ALIPER A, Zavoronkovs A, Zhebrak A, Polykovsky D, Kuznetsov M, Ivanenkov Y, Veselov M, Aladinskiy V, Putin E, Volkov Y, Asadulaev A. Subset conditioning using variational autoencoder with a learnable tensor train induced prior. WO2020060603A1. 2020. https://patents.google.com/patent/WO2020060603A1/en.

18. Colman RJ, Anderson RM, Johnson SC, Kastman EK, Kosmatka KJ, Beasley TM, Allison DB, Cruzen C, Simmons HA, Kemnitz JW, Weindruch R. Caloric restriction delays disease onset and mortality in rhesus monkeys. Science. 2009; 325:201–4. https://doi.org/10.1126/science.1173635 PMID:19590001

19. Fitzgerald KN, Hodges R, Hanes D, Stack E, Cheishvili D, Szyf M, Henkel J, Twedt MW, Giannopoulou D, Herdell J, Logan S, Bradley R. Potential reversal of epigenetic age using a diet and lifestyle intervention: a pilot randomized clinical trial. Aging (Albany NY). 2021; 13:9419–32. https://doi.org/10.18632/aging.202913 PMID:33844651

20. Quach A, Levine ME, Tanaka T, Lu AT, Chen BH, Ferrucci L, Ritz B, Bandinelli S, Neuhouser ML, Beasley JM, Snetselaar L, Wallace RB, Tsao PS, et al. Epigenetic clock analysis of diet, exercise, education, and lifestyle factors. Aging (Albany NY). 2017; 9:419–46. https://doi.org/10.18632/aging.101168 PMID:28198702

21. Arora I, Sharma M, Sun LY, Tollefsbol TO. The Epigenetic Link between Polyphenols, Aging and Age-Related Diseases. Genes (Basel). 2020; 11:1094. https://doi.org/10.3390/genes11091094 PMID:32962067

22. Scott AJ, Ellison M, Sinclair DA. The economic value of targeting aging. Nat Aging. 2021; 1:616–23. https://doi.org/10.1038/s43587-021-00080-0

23. Malone JC, Liu SR, Vaillant GE, Rentz DM, Waldinger RJ. Midlife Eriksonian psychosocial development: Setting the stage for late-life cognitive and emotional health. Dev Psychol. 2016; 52:496–508. https://doi.org/10.1037/a0039875 PMID:26551530

24. Kemph JP. Erik H. Erikson. Identity, youth and crisis. New York: W. W. Norton Company, 1968. Syst Res. 1969; 14:154–9. https://doi.org/10.1002/bs.3830140209

25. Strachey J. The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume VII (1901-1905): A Case of Hysteria, Three Essays on Sexuality and Other Works. The Standard Edition of the Complete Psychological Works of Sigmund Freud. 1953; 123–246.

26. Jiang D, Fung HH. Social and emotional theories of aging. Work across the lifespan. San Diego, CA, USA: Elsevier Academic Press. 2019; 135–53. https://doi.org/10.1016/B978-0-12-812756-8.00006-2

27. Carstensen LL. Socioemotional Selectivity Theory: The Role of Perceived Endings in Human Motivation. Gerontologist. 2021; 61:1188–96. https://doi.org/10.1093/geront/gnab116 PMID:34718558

28. Mitina M, Young S, Zhavoronkov A. Psychological aging, depression, and well-being. Aging (Albany NY). 2020; 12:18765–77.

https://doi.org/10.18632/aging.103880
PMID:32950973

29. Fung HH. Aging in culture. Gerontologist. 2013; 53:369–77.
https://doi.org/10.1093/geront/gnt024
PMID:23585454

30. Boks MP, van Mierlo HC, Rutten BP, Radstake TR, De Witte L, Geuze E, Horvath S, Schalkwyk LC, Vinkers CH, Broen JC, Vermetten E. Longitudinal changes of telomere length and epigenetic age related to traumatic stress and post-traumatic stress disorder. Psychoneuroendocrinology. 2015; 51:506–12.
https://doi.org/10.1016/j.psyneuen.2014.07.011
PMID:25129579

31. Mutz J, Roscoe CJ, Lewis CM. Exploring health in the UK Biobank: associations with sociodemographic characteristics, psychosocial factors, lifestyle and environmental exposures. BMC Med. 2021; 19:240.
https://doi.org/10.1186/s12916-021-02097-z
PMID:34629060

32. Han LKM, Verhoeven JE, Tyrka AR, Penninx BWJ, Wolkowitz OM, Månsson KNT, Lindqvist D, Boks MP, Révész D, Mellon SH, Picard M. Accelerating research on biological aging and mental health: Current challenges and future directions. Psychoneuroendocrinology. 2019; 106:293–311.
https://doi.org/10.1016/j.psyneuen.2019.04.004
PMID:31154264

33. Wertz J, Caspi A, Ambler A, Broadbent J, Hancox RJ, Harrington H, Hogan S, Houts RM, Leung JH, Poulton R, Purdy SC, Ramrakha S, Rasmussen LJH, et al. Association of History of Psychopathology With Accelerated Aging at Midlife. JAMA Psychiatry. 2021; 78:530–9.
https://doi.org/10.1001/jamapsychiatry.2020.4626
PMID:33595619

34. Wurm S, Warner LM, Ziegelmann JP, Wolff JK, Schüz B. How do negative self-perceptions of aging become a self-fulfilling prophecy? Psychol Aging. 2013; 28:1088–97.
https://doi.org/10.1037/a0032845
PMID:24128074

35. Hatton SN, Franz CE, Elman JA, Panizzon MS, Hagler DJ Jr, Fennema-Notestine C, Eyler LT, McEvoy LK, Lyons MJ, Dale AM, Kremen WS. Negative fateful life events in midlife and advanced predicted brain aging. Neurobiol Aging. 2018; 67:1–9.
https://doi.org/10.1016/j.neurobiolaging.2018.03.004
PMID:29609076

36. Valle Gottlieb MG, Closs VE, Junges VM, Schwanke CHA. Impact of human aging and modern lifestyle on gut microbiota. Crit Rev Food Sci Nutr. 2018; 58:1557–64.

37. Robertson T, Batty GD, Der G, Fenton C, Shiels PG, Benzeval M. Is socioeconomic status associated with biological aging as measured by telomere length? Epidemiol Rev. 2013; 35:98–111.
https://doi.org/10.1093/epirev/mxs001
PMID:23258416

38. Anderson JA, Johnston RA, Lea AJ, Campos FA, Voyles TN, Akinyi MY, Alberts SC, Archie EA, Tung J. High social status males experience accelerated epigenetic aging in wild baboons. Elife. 2021; 10:e66128.
https://doi.org/10.7554/eLife.66128
PMID:33821798

39. Zannas AS, Arloth J, Carrillo-Roa T, Iurato S, Röh S, Ressler KJ, Nemeroff CB, Smith AK, Bradley B, Heim C, Menke A, Lange JF, Brückl T, et al. Lifetime stress accelerates epigenetic aging in an urban, African American cohort: relevance of glucocorticoid signaling. Genome Biol. 2015; 16:266.
https://doi.org/10.1186/s13059-015-0828-5
PMID:26673150

40. Epel ES, Blackburn EH, Lin J, Dhabhar FS, Adler NE, Morrow JD, Cawthon RM. Accelerated telomere shortening in response to life stress. Proc Natl Acad Sci U S A. 2004; 101:17312–5.
https://doi.org/10.1073/pnas.0407162101
PMID:15574496

41. Jovanovic T, Vance LA, Cross D, Knight AK, Kilaru V, Michopoulos V, Klengel T, Smith AK. Exposure to Violence Accelerates Epigenetic Aging in Children. Sci Rep. 2017; 7:8962.
https://doi.org/10.1038/s41598-017-09235-9
PMID:28827677

42. Brody GH, Yu T, Chen E, Beach SR, Miller GE. Family-centered prevention ameliorates the longitudinal association between risky family processes and epigenetic aging. J Child Psychol Psychiatry. 2016; 57:566–74.
https://doi.org/10.1111/jcpp.12495
PMID:26680699

43. Pavanello S, Campisi M, Tona F, Lin CD, Iliceto S. Exploring Epigenetic Age in Response to Intensive Relaxing Training: A Pilot Study to Slow Down Biological Age. Int J Environ Res Public Health. 2019; 16:3074.
https://doi.org/10.3390/ijerph16173074
PMID:31450859

44. Harvanek ZM, Fogelman N, Xu K, Sinha R. Psychological and biological resilience modulates the effects of stress on epigenetic aging. Transl Psychiatry. 2021; 11:601.

https://doi.org/10.1038/s41398-021-01735-7
PMID:34839356

45. Manczak EM, Scott SR, Millwood SN. Accelerated epigenetic aging at birth interacts with parenting hostility to predict child temperament and subsequent psychological symptoms. Dev Psychopathol. 2021. [Epub ahead of print]. https://doi.org/10.1017/S0954579421000614 PMID:34620266

46. Zhavoronkov A, Kochetov K, Diamandis P, Mitina M. PsychoAge and SubjAge: development of deep markers of psychological and subjective age using artificial intelligence. Aging (Albany NY). 2020; 12:23548–77. https://doi.org/10.18632/aging.202344 PMID:33303702

47. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol. 2014; 43:61–8. https://doi.org/10.1093/ije/dys203 PMID:23243115

48. Li M, Shi Z. A Prospective Association of Nut Consumption with Cognitive Function in Chinese Adults aged 55+ _ China Health and Nutrition Survey. J Nutr Health Aging. 2019; 23:211–6. https://doi.org/10.1007/s12603-018-1122-5 PMID:30697633

49. Pan X, Chee KH. The power of weak ties in preserving cognitive function: a longitudinal study of older Chinese adults. Aging Ment Health. 2020; 24:1046–53. https://doi.org/10.1080/13607863.2019.1597015 PMID:30955345

50. Zhang Z, Li LW, Xu H, Liu J. Does widowhood affect cognitive function among Chinese older adults? SSM Popul Health. 2018; 7:100329. https://doi.org/10.1016/j.ssmph.2018.100329 PMID:30581964

51. You X, Zhang Y, Zeng J, Wang C, Sun H, Ma Q, Ma Y, Xu Y. Disparity of the Chinese elderly's health-related quality of life between urban and rural areas: a mediation analysis. BMJ Open. 2019; 9:e024080. https://doi.org/10.1136/bmjopen-2018-024080 PMID:30782725

52. Kong T, Yang P. Finding the vulnerable among China's elderly: identifying measures for effective policy targeting. J Aging Soc Policy. 2019; 31:271–90. https://doi.org/10.1080/08959420.2018.1485391 PMID:29979945

53. Wang Q, Tapia Granados JA. Economic growth and mental health in 21st century China. Soc Sci Med. 2019; 220:387–95. https://doi.org/10.1016/j.socscimed.2018.11.031 PMID:30529490

54. Jiang N. Adult Children's Education and Later-Life Health of Parents in China: The Intergenerational Effects of Human Capital Investment. Soc Indic Res. 2019; 145:257–78.

55. Li J, Shi L, Liang H, Ma C, Xu L, Qin W. Health care utilization and affordability among older people following China's 2009 health reform -- evidence from CHARLS pilot study. Int J Equity Health. 2019; 18:62. https://doi.org/10.1186/s12939-019-0969-3 PMID:31053074

56. Lin B, Lin Z, Zhang YY, Liu W. The Impact of New Rural Pension Scheme on the Rural Elderly in China: Evidences of Regional Differences in Formal and Informal Labor Supply. 2018 Conference, July 28-August 2, 2018, Vancouver, British Columbia 275997, International Association of Agricultural Economists. 2018. https://doi.org/10.22004/ag.econ.275997

57. Singh V. Ageing Society and Labor Policy in China: Analyzing Policy Challenges and Options. Chin Stud. 2018; 7:242–50. https://doi.org/10.4236/chnstd.2018.73021

58. Liu Y, Roberto KA, Savla JS. Social determinants of Health, Health Disparities, and Caregiving for Aging Parents in China. Innovation in Aging. 2017; 1:814. https://doi.org/10.1093/geroni/igx004.2938

59. Nakagawa T, Cho J, Yeung DY. Successful Aging in East Asia: Comparison Among China, Korea, and Japan. J Gerontol B Psychol Sci Soc Sci. 2021; 76:S17–26. https://doi.org/10.1093/geronb/gbaa042 PMID:32324214

60. Asian Perspectives. Successful Aging. Springer Dordrecht. 2015. https://doi.org/10.1007/978-94-017-9331-5

61. Altmann A, Toloşi L, Sander O, Lengauer T. Permutation importance: a corrected feature importance measure. Bioinformatics. 2010; 26:1340–7. https://doi.org/10.1093/bioinformatics/btq134 PMID:20385727

62. Putin E, Mamoshina P, Aliper A, Korzinkin M, Moskalev A, Kolosov A, Ostrovskiy A, Cantor C, Vijg J, Zhavoronkov A. Deep biomarkers of human aging: Application of deep neural networks to biomarker development. Aging (Albany NY). 2016; 8:1021–33. https://doi.org/10.18632/aging.100968 PMID:27191382

63. Mamoshina P, Kochetov K, Cortese F, Kovalchuk A, Aliper A, Putin E, Scheibye-Knudsen M, Cantor CR, Skjodt NM, Kovalchuk O, Zhavoronkov A. Blood Biochemistry Analysis to Detect Smoking Status and Quantify Accelerated Aging in Smokers. Sci Rep. 2019; 9:142.

https://doi.org/10.1038/s41598-018-35704-w
PMID:30644411

64. Feitelson MA, Arzumanyan A, Kulathinal RJ, Blain SW, Holcombe RF, Mahajna J, Marino M, Martinez-Chantar ML, Nawroth R, Sanchez-Garcia I, Sharma D, Saxena NK, Singh N, et al. Sustained proliferation in cancer: Mechanisms and novel therapeutic targets. Semin Cancer Biol. 2015 (Suppl); 35:S25–54.
https://doi.org/10.1016/j.semcancer.2015.02.006
PMID:25892662

65. Shen Y, Zhang Y, Xiong S, Zhu X, Ke C. High-sensitivity C-reactive protein and cystatin C independently and jointly predict all-cause mortality among the middle-aged and elderly Chinese population. Clin Biochem. 2019; 65:7–14.
https://doi.org/10.1016/j.clinbiochem.2018.12.012
PMID:30592989

66. Li H, Wang A, Qi G, Guo J, Li X, Wang W, Xiang Y, Guo X. Cystatin C and risk of new-onset depressive symptoms among individuals with a normal creatinine-based estimated glomerular filtration rate: A prospective cohort study. Psychiatry Res. 2019; 273:75–81.
https://doi.org/10.1016/j.psychres.2019.01.009
PMID:30640054

67. Li T, Tsang VH, Fung HH, Qiu XL, Wang WC. Measuring dynamic goals for marriage: Development and validation of the Marital Goal Scale using Rasch modeling. Psychol Assess. 2020; 32:211–26.
https://doi.org/10.1037/pas0000779
PMID:31647255

68. Li T, Fung HH. The Dynamic Goal Theory of Marital Satisfaction. Rev Gen Psychol. 2011; 15:246–54.
https://doi.org/10.1037/a0024694

69. Yeung GTY, Fung HH. Social support and life satisfaction among Hong Kong Chinese older adults: family first? Eur J Ageing. 2007; 4:219–27.
https://doi.org/10.1007/s10433-007-0065-1
PMID:28794791

70. Jin Q, Pearce P, Hu H. The Study on the Satisfaction of the Elderly People Living with Their Children. Soc Indic Res. 2018; 140:1159–72.
https://doi.org/10.1007/s11205-017-1803-1

71. Zhang C, Xiao S, Shi L, Xue Y, Zheng X, Dong F, Zhang J, Xue B, Lin H, Ouyang P. Urban-Rural Differences in Patterns and Associated Factors of Multimorbidity Among Older Adults in China: A Cross-Sectional Study Based on Apriori Algorithm and Multinomial Logistic Regression. Front Public Health. 2021; 9:707062.
https://doi.org/10.3389/fpubh.2021.707062
PMID:34527650

72. Hu Y, Li P, Martikainen P. Rural-urban disparities in age trajectories of depression caseness in later life: The China Health and Retirement Longitudinal Study. PLoS One. 2019; 14:e0215907.
https://doi.org/10.1371/journal.pone.0215907
PMID:31022250

73. Hu Y, Li P, Martikainen P. Depression with age among older chinese: rural-urban difference and determinants. Innov Aging. 2018; 2:510.
https://doi.org/10.1093/geroni/igy023.1892

74. Galkin F, Kochetov K, Keller M, Zhavoronkov A, Etcoff N. Optimizing future well-being with artificial intelligence: self-organizing maps (SOMs) for the identification of islands of emotional stability. Aging (Albany NY). 2022; 14:4935–58.
https://doi.org/10.18632/aging.204061
PMID:35723468

75. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, et al. Scikit-learn: Machine learning in Python. J Mach Learn Res. 2011; 12:2825–30.
http://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf.

76. Pascanu R, Mikolov T, Bengio Y. Understanding the exploding gradient problem. ArXiv. 2012.
https://www.semanticscholar.org/paper/Understanding-the-exploding-gradient-problem-Pascanu-Mikolov/c5145b1d15fea9340840cc8bb6f0e46e8934827f.

77. Clevert DA, Unterthiner T, Hochreiter S. Fast and accurate deep network learning by exponential linear units (ELUs). 4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings. 2016.
http://arxiv.org/abs/1511.07289.

78. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res. 2014; 15:1929–58.
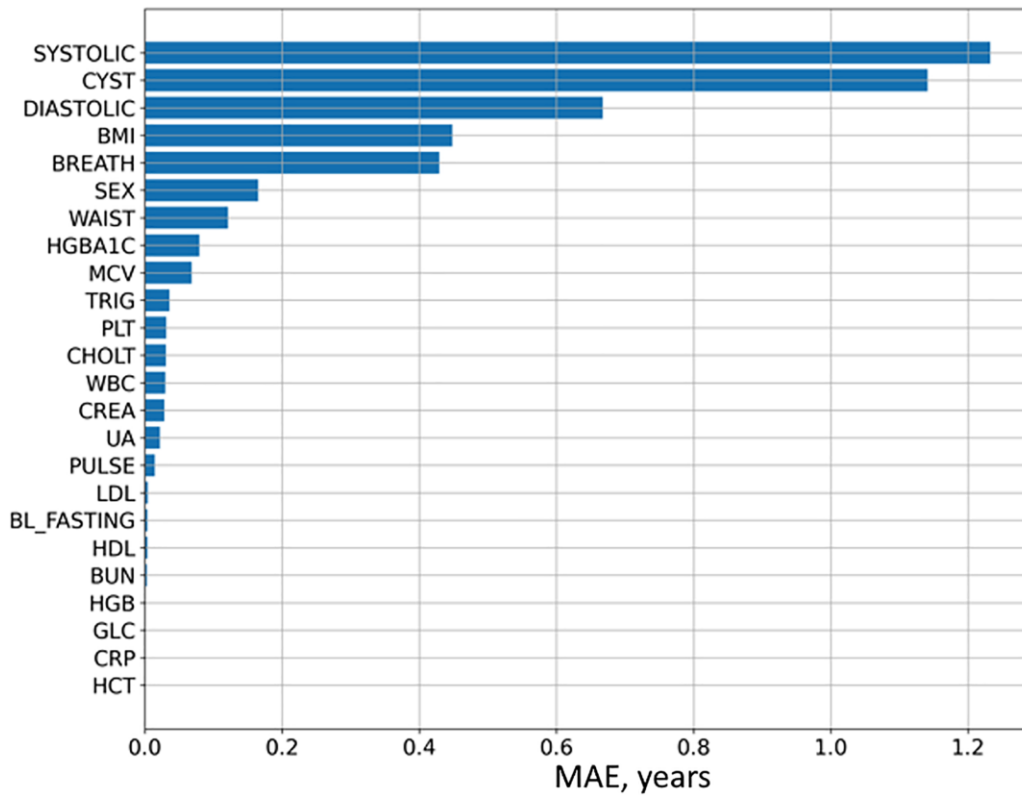
Supplementary Figure 1. Permutation feature importance analysis was carried out to determine the features our aging clock pays the most attention to during age prediction. The importance of a feature is judged based on the drop in the MAE of the model when a feature's value is shuffled in the test set.
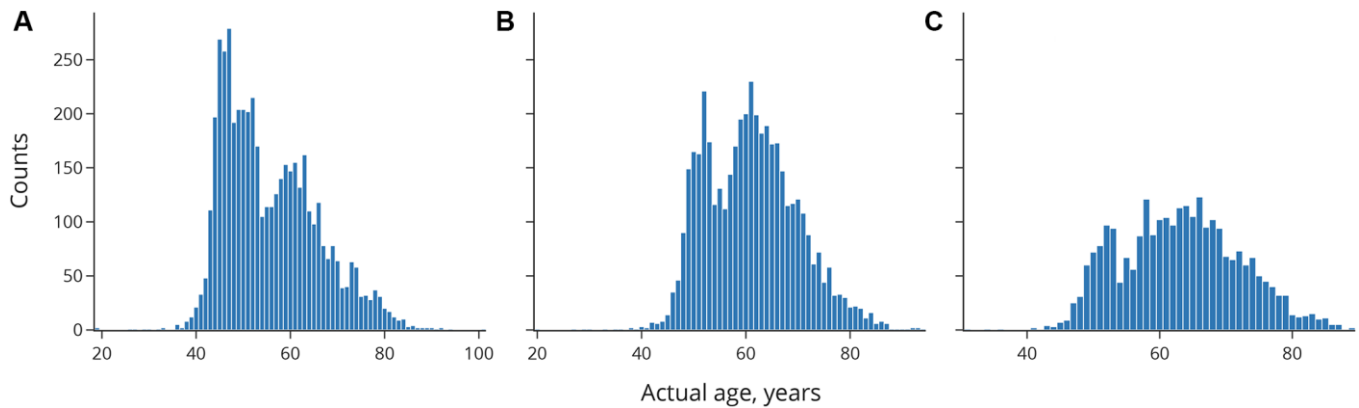
Supplementary Figure 2. Age distribution in the three cohorts used in this study. (**A**) Training set (N = 4846), mean age is 55.7 years (**B**) Test set (N = 4451), mean age is 61.0 years (**C**) Discovery set (N = 2617), mean age is 63.0 years.

Supplementary Tables

Supplementary Table 1. The prevalence of the conditions used in the elastic net importance analysis section.

Feature	Percent positive
Is male	45.90%
Is married	89.15%
Is widowed	9.25%
Is rural	73.93%
Current smoker	28.37%
Bothered by things is rare	55.45%
Lack of focus is rare	56.53%
Depression is rare	55.43%
Hopeful is rare	29.32%
Fearful is rare	83.32%
Restless sleep is rare	53.14%
Happiness is rare	19.05%
Loneliness is rare	74.88%
Total N	9297

The statistics are derived from the training and test sets only.

Supplementary Table 2. The mean values and standard deviations (std) of the variables used by the age predictor.

Variable name	Mean	Std	Units
White blood cells	5.98	2.06	x10E9/L
Hemoglobin	137.28	19.22	g/L
Hematocrit	41.50	5.58	%
Mean corpuscular volume	91.28	7.70	fL
Platelets	205.31	74.14	x10E9/L
Triglycerides	1.61	1.03	mmol/L
Creatinine	70.96	24.05	umol/L
Blood urea nitrogen	5.48	1.63	mmol/L
HDL cholesterol	1.32	0.29	mmol/L
LDL cholesterol	2.64	0.74	mmol/L
Total cholesterol	4.75	0.93	mmol/L
Glucose	5.70	1.87	mmol/L
Uric acid	0.29	0.08	mmol/L
Cystatin C	0.84	0.23	mg/L
C reactive protein	2.57	5.31	mg/L
Glycated hemoglobin	5.96	0.96	%
Systolic pressure	127.59	19.64	mmHg
Diastolic pressure	75.61	11.63	mmHg
Heart rate	73.98	10.53	bpm

BMI	23.95	3.63	kg/m2
Spirometry	296.03	117.63	L/min
Waist circumference	86.36	10.4	cm
Total *N*		11914	

The statistics are calculated based on the total study sample including training, test, and discovery sets.