

Gerontology , DOI: 10.1159/000534015

Received: November 29, 2022

Accepted: September 5, 2023

Published online: September 19, 2023

## **Associations between differential aging and lifestyle, environment, current, and future health conditions: Findings from Canadian Longitudinal Study on Aging**

Song Y, Liu YS, Talarico F, Zhang Y, Hayward J, Wang M, Stroulia E, Dixon RA, Greiner R, Li X, Greenshaw A, Jie S, Cao B

ISSN: 0304-324X (Print), eISSN: 1423-0003 (Online)

<https://www.karger.com/GER>

Gerontology

### Disclaimer:

Accepted, unedited article not yet assigned to an issue. The statements, opinions and data contained in this publication are solely those of the individual authors and contributors and not of the publisher and the editor(s). The publisher and the editor(s) disclaim responsibility for any injury to persons or property resulting from any ideas, methods, instructions or products referred to the content.

### Copyright:

© 2023 S. Karger AG, Basel

***Associations between differential aging and lifestyle, environment, current, and future health conditions: Findings from Canadian Longitudinal Study on Aging***

Yipeng Song<sup>a</sup>, Yang S. Liu<sup>a</sup>, Fernanda Talarico<sup>a</sup>, Yanbo Zhang<sup>a</sup>, Jake Hayward<sup>b</sup>, Mengzhe Wang<sup>c</sup>, Eleni Stroulia<sup>d</sup>, Roger A. Dixon<sup>e</sup>, Russell Greiner<sup>a,e</sup>, Xinmin Li<sup>a</sup>, Andrew Greenshaw<sup>a</sup>, Sui Jie<sup>f</sup>, Bo Cao<sup>a,\*</sup>

<sup>a</sup> University of Alberta, Department of Psychiatry, Edmonton, Canada

<sup>b</sup> University of Alberta, Department of Emergency Medicine, Edmonton, Canada

<sup>c</sup> Ministry of Health (Alberta), Edmonton, Canada

<sup>d</sup> University of Alberta, Department of Computer Science, Edmonton, Canada

<sup>e</sup> University of Alberta, Department of Psychology, Edmonton, Canada

<sup>f</sup> University of Aberdeen, The School of Psychology, Aberdeen, Scotland, UK

Short Title: Differential aging and lifestyle, environment, current and future health conditions.

Corresponding Author: Dr. Bo Cao, Department of Psychiatry, Faculty of Medicine & Dentistry, University of Alberta, 11361 87th Avenue, Edmonton, Alberta, T6G 2E1, Canada.

(780) 407-6504, [cloudbocao@gmail.com](mailto:cloudbocao@gmail.com).

Number of Tables: 1.

Number of Figures: 3.

Word count: 3850.

Keywords: Aging, Canadian Longitudinal Study of Aging, Biological age, Healthy longevity, Machine learning.

## Abstract

*Introduction:* An aging population is a pressing challenge for the healthcare system. Insights into promoting healthy longevity can be gained by quantifying the biological aging process and understanding the roles of modifiable lifestyle and environmental factors and chronic disease conditions.

*Methods:* We developed a biological age (BioAge) index by applying multiple state-of-art machine learning models based on easily accessible blood test data from the Canadian Longitudinal Study of Aging (CLSA). The BioAge gap, which is the difference between BioAge index and chronological age, was used to quantify the differential aging, i.e., the difference between biological and chronological age, of the CLSA participants. We further investigated the associations between the BioAge gap and lifestyle, environmental factors, and current and future health conditions. *Results:* BioAge gap had strong associations with existing adverse health conditions (e.g., cancers, cardiovascular diseases, diabetes, kidney diseases) and future disease onset (e.g., Parkinson's disease, diabetes, and kidney diseases). We identified that frequent consumption of processed meat, pork, beef, and chicken, poor outcomes in nutritional risk screening, cigarette smoking, exposure to passive smoking, are associated with positive BioAge gap ("older" than expected in BioAge index). We also identified several modifiable factors, including eating fruits, legumes, vegetables, related to negative BioAge gap ("younger" than expected in BioAge index).

*Discussion/Conclusions:* Our study shows that a BioAge index based on easily accessible blood tests has the potential to quantify the biological aging process that is associated with current and future adverse health events. The identified risk and protective factors for differential aging indicated by BioAge gap are informative for guiding policy making to promote healthy longevity.

## Introduction

The global population is aging rapidly. In 2015, people aged 65 and over accounted for 8.5% of the global population, and this percentage is projected to increase to 16.7% in 2050.[1] For developed countries, the proportions are higher; for example, in Canada, 16.2% were over 65 years in 2018, and this is projected to reach 23.4% in 2040.[2] An aging population places strain on healthcare systems as rates of chronic disease increase. For example, in Canada, older adult aged 65 and over, although representing only 16.2% of the general population, accounted for 45.7% of the total health expenditures in 2019.[2] Understanding aging, and its interactions with lifestyle, environmental factors and chronic disease, is a prerequisite for implementing effective interventions to promote healthy longevity. Chronological age is a common choice for representing and quantifying aging processes. However, the full complexity of biological aging cannot be encapsulated solely by chronological age, as people with the same chronological age can have significantly different functionality and health conditions.

Currently, researchers are searching for biomarkers to help gauge the full complexity of the aging process.[3] Among them, various biological age (BioAge) predictors, also called BioAge indices, have been developed.[3–6] Aging leaves many footprints on the biological system, including telomere shortening, gradual changes in gene expression, and DNA methylation levels.[4] A commonly used approach to developing a BioAge index is to construct machine learning (ML) models that predict chronological age based on biological measurements collected from a large population. Then, given the unique biological measurements of an individual, the model can estimate their BioAge.[4] Such a method may capture differences in BioAge for people of the same chronological age. If trained on a representative population, the model estimates of BioAge should reflect the typical (i.e., expected) aging. The BioAge gap, which is defined as the gap between estimated BioAge index and chronological age, is a measure of the differential aging (the difference between BioAge and chronological age).[4] A positive BioAge gap indicates individuals are biologically “older” than would be expected given their chronological age, which may result from accelerated accumulation of multiple morbidities in aging. A negative BioAge gap suggests individuals are biologically “younger” than would be expected given their chronological age, which may result from fewer morbidities or delayed biological aging.

Many types of biological measures have been used to quantify aging, including metabolomics, proteomics, epigenomic, and transcriptomics.[3,7] Several recent studies have also used easily accessible clinical laboratory biochemical and haematology tests to develop BioAge indices in humans. Putin et al.[8] created a BioAge index using an ensemble of 21 deep neural networks on 46 clinical blood test measures. Mamoshina et al.[9] developed another aging index based on 20 selected laboratory biomarkers using a deep neural network, and the model was used to quantify the accelerated aging effects among smokers. At present, no study has fully explored the utility of laboratory measures based BioAge indices in characterizing the interactions between differential aging and lifestyles, environmental factors, current disease, and prospective disease. Understanding these associations and identifying risk factors for differential aging could guide effective public-health interventions to promote healthy longevity. Further, a BioAge gap could theoretically be used to predict and prevent future diseases and promote healthier aging.

The Canadian Longitudinal Study on Aging (CLSA) is one of the largest aging research projects in the world[10], collecting comprehensive information on lifestyle, home, and environmental factors, disease status, and standard clinical laboratory biochemistry and haematology measures for about 30,000 participants aged 45 years to 87 years old. We aimed to develop a BioAge index using the clinical laboratory measures in CLSA data[10] to quantify biological aging at the individual level. We also aimed to investigate the relationships between differential aging indicated by the BioAge gap and lifestyle, home and environmental factors, and current and prospective chronic diseases based on the comprehensive measures in CLSA data. Addressing the role of associated risk factors occurring in differential aging may provide a more sensitive index of biological aging.

## Materials and Methods

### *Data*

CLSA is a Canada-wide longitudinal study on adult development and aging.[10] Over 50,000 Canadians aged 45 and over participated in the study with a planned follow up of at least 20 years. Residents living on First Nations reserves or in long-term care institutions, full-time members of the Canadian Armed Forces, and those cognitively impaired or who cannot respond in English or French are excluded from the CLSA study.[10] For this study, we used the Comprehensive Cohort of CLSA study, comprised of 30,097 participants for whom data were collected through both in-home interviews and CLSA data collection sites visit. 4,673 subjects without laboratory measures at the baseline were excluded from the analysis. Currently, the baseline data collected from 2012 to 2015 and the follow-up phase data collected from 2015 to 2018 are available. We retrieved the chronological age, clinical laboratory measures (detailed information can be found in Table S1), socioeconomic status, lifestyle, home and neighbourhood environmental factors, and disease conditions at the baseline, and 10 primary chronic disease conditions (including all the neurological and psychological disorders and some common chronic diseases, e.g., cancer, heart disease, and kidney disease) at the follow-up phase.

### *Aging index development*

To construct the BioAge index, the 31 laboratory measures present in the CLSA data and sex were used as the predictive variables, while the chronological age at the time of blood sample collection was used as the target outcome. See Table S1 for the summary statistics of these measures and chronological age. The data was initially split into a training set (80% of the data) and a test set (20%). Various ML models were chosen and trained on the training set. The model that exhibited the highest performance in terms of R-Squared error on the test set was then selected to generate a biological aging index through the use of 10-fold cross-validation on the complete dataset. Details about the data pre-processing and modelling process can be found in the supplement.

### *Post hoc analysis of important variables for building aging index*

After the models were selected and evaluated, the importance of each predictive variable was derived to assess their contribution in estimating the BioAge. Almost all the available variables in the model will have nonzero feature importance measures. However, we want to focus on interpreting the most important ones rather than all the variables. We implemented a permutation-based method to heuristically select the variables with higher than chance contributions in estimating BioAge for interpretation. See supplementary Figure S1 for details.

### *Post hoc association analysis between variables and BioAge gap*

Having estimated the BioAge index for our CLSA cohort, we computed the BioAge gap as the difference between an individual's estimated BioAge index based on CV and their chronological age. Details about the calculation of the BioAge gap (selecting the best performed ML models in estimating BioAge, predicting the BioAge in a CV framework to avoid overfitting) can be found in the supplement. After that, we used a linear model to explore the association between the estimated BioAge gap, i.e., the difference between predicted age and chronological age, and a phenotype of interest (e.g., diabetes) while controlling for common confounding factors, including the subject's chronological age, sex, and socioeconomic status factors (education level and household income). As shown by Lange et al.[11], the dependence of the BioAge gap on the chronological age can be corrected by including chronological age as the covariate. See supplement for details of these socioeconomic status factors. In the above linear model, the coefficient of the phenotype represents its partial association with the outcome after adjusting the effect of the confounding factors. For a categorical phenotype variable, one level was taken as the reference level, while for a numerical phenotype variable, its raw value is scaled to have a mean 0 and standard deviation 1. In addition, the corresponding samples were removed for rare levels of categorical variables (frequency less than

1%) to avoid false discoveries in the related association analysis. The BioAge gap is an estimated value of uncertainty. Without accounting for the uncertainty, the association analysis will underestimate the variation, thus leading to false discovery. We applied bootstrapping to account for the uncertainty in the BioAge gap. See supplement for details. In addition, as chronological age is not measured with precision, e.g., both 70.9 years and 70.1 years are taken as 70 years old, we will only interpret the significant variables with an effect size larger than or close to one year.

## Results

The best performing model explained 34.6% of the variation in the chronological age of the hold-out test set (Table S2, Fig. S2).

### *Important variables for building the biological aging index*

The contribution of each variable to building the BioAge index is shown in Fig. S3. The post hoc analysis identified lymphocytes, red blood cells, mean corpuscular volume, red blood cell distribution width, Hemoglobin A1C, 25-hydroxyvitamin D, albumin, alanine aminotransferase, ferritin, and low-density lipoprotein (LDL), as the most important features contributing to the BioAge index. How these variables affect the BioAge index in the ML model is shown in Fig. S4.

### *Lifestyle factors associations with BioAge gap*

After FDR correction, 90 of the 409 lifestyle factors in CLSA data were identified to be significantly associated with the BioAge gap. These significant lifestyle factors, the associated effect sizes and p-values, are shown in Table S3. The lifestyle factors are categorized into dietary and other lifestyle factors for results demonstration.

#### *Dietary lifestyle factors*

The detected associations between dietary lifestyle factors and BioAge gap can be summarized as follows. 1) Eating meat, i.e., beef, pork, chicken or turkey, regularly was associated with a positive BioAge gap. (Fig. 1A, 1B) 2) Daily consumption of processed meat, i.e., sausages, hot dogs, ham, smoked meat, bacon, was associated with a positive BioAge gap. (Fig. 1C) 3) The habit of eating fruits, legumes and vegetables was associated with a negative BioAge gap. (Fig. 1D, 1E, 1F).

#### *Other lifestyle factors*

The other lifestyle factors associated with the differential aging indicated by the BioAge gap can be summarized as: 1) Cigarette smoking status was related to a positive BioAge gap (Fig. 2A). Current smokers, on average, were  $1.479 \pm 0.133$  (p-value < .001) years older measured by the BioAge index compared to non-smokers, and previous smoking status was also significantly associated with a positive BioAge gap. However, the effect size ( $0.236 \pm 0.076$  p-value =.003) of previous smoking status is not large enough for accurate interpretation. Many other smoking-related variables, e.g., number of cigarettes smoked per day, type of smokers, etc., were significant and had large effect sizes. In addition, to compare the smoking results with previous research[9,12], we also explored the association among different age groups. As shown in Fig. S5, compared to non-smokers, current smoking status is significantly associated with an over one-year positive BioAge gap in all the age groups except for those over 80 years old. Furthermore, the association between previous smoking status and BioAge gap is not significant among all the age groups. Furthermore, alcohol consumption was associated with a positive BioAge gap, but the effect size was less than one year. 2) People participating in strenuous activities, such as paddleball, and tennis, were predicted to be  $1.40 \pm 0.451$  (p-value =.002) years younger compared to the people not participating in such activities (Fig. 2C). However, all other 24 strenuous sports included in the CLSA data, e.g., basketball, bicycling, hiking, etc., were not significant. In addition, participating in light activities, e.g., badminton, billiards, boating, fishing, etc., were not significant. 3) People who used a wheelchair or motorized carts for transportation were at risk of a one-year positive BioAge gap (Fig. 2D).

### *Association of the home and the neighbourhood environmental factors with the BioAge gap*

After FDR correction, 6 of the 28 home and the neighbourhood environmental factors were significantly associated with the BioAge gap. These significant environmental factors, the associated effect sizes and p-values, are shown in Table S4. The main finding is that passive smoking exposure at home was associated with a positive BioAge gap. The higher the frequency of passive smoking, the higher the positive age gap (Fig. 2B).

### *Association between the BioAge gap and current disease conditions*

After FDR correction, 79 of the 379 disease-related variables were significantly associated with a BioAge gap. These significant disease-related variables, the associated effect sizes and p-values, are shown in Table S5. Some of the features with significant associations and large effect sizes can be summarized as: 1) Reported chronic conditions were associated with a positive BioAge gap (Fig. 3A). 2) Various cancers, i.e., kidney cancer, other lymphoid, hematopoietic and related tissue cancer, lung cancer, and non-Hodgkin lymphoma, were associated with a positive BioAge gap (Fig. 3C). 3) Cardiovascular diseases, i.e., congestive heart failure, aneurysm, peripheral vascular disease, angina, stroke, heart attack, or myocardial infarction, were associated with a positive BioAge gap (Fig. 3B). 4) Other diseases and conditions, e.g., kidney disease or failure, multiple sclerosis, diabetes, emphysema, chronic bronchitis, chronic obstructive pulmonary disease (COPD), or chronic changes in lungs due to smoking, under-active thyroid gland, pneumonia in the last year, shuffling gait and poor balance, were also associated with positive BioAge gap (Fig. 3D, 3E, 3F).

### *Association between BioAge gap and prospective chronic disease conditions*

We selected several major chronic physical and mental disorders to see if the onset of these diseases in the follow-up phase was associated with the derived baseline BioAge gap. The association analysis results of all the selected disease conditions are summarized in Table S6. After FDR correction, 5 out of 10 future disease conditions were significant. The most significant ones are as follows: Parkinson's disease onset at the follow-up phase was associated with  $2.21 \pm 0.782$  years ( $p$ -value = 0.0054) positive BioAge gap at the baseline. Being told by a doctor to have memory problem onset at follow-up was associated with  $0.789 \pm 0.394$  years ( $p$ -value = 0.0475) positive BioAge gap at baseline and the onset of kidney disease at the follow-up phase was associated with  $1.77 \pm 0.370$  years ( $p$ -value < 0.001) positive BioAge gap at baseline. Also, the onset of diabetes at the follow-up phase was related to  $1.10 \pm 0.184$  years ( $p$ -value < 0.001) positive BioAge gap at baseline.

### *Modifiable risk factors*

As modifiable risk factors, those factors with the potential to be intervened or prevented, are crucial for interventions, we also listed out the identified modifiable risk factors with large effect size in Table 1.

## **Discussion/Conclusion**

In this study, we showed how the biological aging can be quantified with a BioAge index estimated from commonly available laboratory measures and its interaction with lifestyle, home and the neighbourhood environmental factors, and current and future disease conditions. We successfully identified a series of modifiable risks, protective lifestyle and environmental factors for the differential aging gauged by the BioAge gap. Furthermore, we demonstrated the utility of the BioAge gap in informing adverse health status by its strong association with various current and future disease conditions. To our best knowledge, this is the first study to fully explore the associations between human aging, quantified by a laboratory measure-based BioAge index, and comprehensive lifestyle, environmental, and disease conditions.

The variables contributing to the BioAge index are largely consistent with previous studies. For example, lymphocytes[13], red blood cell counts[14], Alanine Aminotransferase[15] and Ferritin[16]

have been identified to be associated with aging. Also, 25-hydroxyvitamin D[17] and Albumin[18] were shown to be related to the aging and also to poor health outcomes. Hemoglobin A1c, the biomarker for diagnosing prediabetic and diabetic, was observed to have a higher level among older nondiabetic individuals.[19] In addition, the increasing of LDL level with age was identified as a result of reduced LDL clearance ability within aging population.[20]

We identified multiple medical conditions associated with the positive BioAge gap (“older” than expected in BioAge) in the aging population (Table 1). An array of existing chronic diseases, i.e., cancers, cardiovascular diseases, diabetes, kidney diseases, chronic respiratory diseases, and under-active thyroid gland, were identified to be associated with a positive BioAge gap, which reflects the subject is potentially biologically “older” than expected based on her/his chronological age. Previous research has identified that positive BioAge gap gauged by methylation-based aging indices are risk factors for the incidence of cardiovascular conditions, i.e., fatal coronary heart disease, peripheral arterial disease, and heart failure.[21] Our results provide more evidence to support the association between positive BioAge gap and cardiovascular diseases. The associations observed between positive BioAge gap and various cancers also support previous findings that cancer and chemotherapy are potential aging accelerators.[22] In addition, the identified associations between positive BioAge gap, kidney diseases, and diabetes are also supported by previous research. Jeroen et al.[23] summarized the underlying mechanisms of chronic kidney disease and premature aging. And type 2 diabetes was identified to associated with aging acceleration, increased telomere shortening and mitochondrial DNA depletion.[24] Increased prevalence of thyroid disorder and decreased thyroid function has been observed in elderly population[25], our finding of the association between the under-active thyroid gland and positive BioAge gap suggests that aging acceleration may exist in those with hypothyroidism. The associations between positive BioAge gap, recent pneumonia, and chronic lung diseases suggest aging acceleration patterns may exist in those with lung disease. The results were partly supported by previous research that COPD, a chronic lung disease, has been considered a condition of accelerated lung aging[26].

We also found that positive differential aging at baseline, i.e., a positive BioAge gap, was associated with the onset of kidney disease, diabetes and Parkinson’s disease about three years into the future, suggesting that the biological aging index might predict future disease. The association with Parkinson’s disease is particularly notable for Parkinson’s disease is not typically detectable in routine blood tests[27], which are the basis for our aging index. One potential explanation for the strong association with Parkinson’s disease is that dopamine neurons degenerate more quickly in Parkinson’s disease, leading to accelerated or exaggerated aging. [28]

Except for the medical outcomes, we also identified multiple modifiable risk factors, those factors with the potential to be intervened or prevented, associated with the differential aging in the aging population (Table 1). We find that current and previous smoking behaviours are associated with positive differential aging. However, the effect size of current smoking is about 1.480 years, while the effect size of previous smoking is only moderate. The previous research on the association between smoking and aging acceleration is inconsistent. For example, Z Linli <sup>41</sup> identified that active regular smoking behaviour is associated with 1.190 years of brain age acceleration. Mamoshina et al.[9] found that active smokers have a higher aging rate based on laboratory measures. Associations between smoking status and methylation-based BioAge gaps were found by Horvath et al.[29]. Our results support these findings, showing that smoking may contribute to differential aging. In addition, unlike the results in Mamoshina et al. [9], which showed that only subjects younger than 40 years have signs of accelerated aging, we identified associations between current smoking and a positive BioAge gap across all the age groups except for those over 80 years (Figure S5). For smokers over 80 years old, previous research has shown no increase in mortality risk compared to smokers from other age groups [30]. Notably, we find an association between second-hand smoking exposure and aging acceleration. Overall, our results imply that quitting smoking and reducing second-hand smoking exposure could promote healthy longevity and slow down biological aging.



Previous research has shown that accelerated aging, as indicated by telomere length and DNA methylation contents, is associated with dietary phosphate intake from red meat consumption. Our results support this finding. We further identified that eating chicken or processed meat are related to “older” than expected BioAge. Healthy eating, on the other hand, e.g., frequent consumption of legumes, vegetables, and fruits, is known to lower the risk of cardiovascular disease and cancer and are recommended in dietary guidelines [31]. We find support for these dietary recommendations showing their associations with “younger” than expected BioAge. Playing tennis was particularly protective against negative BioAge gap, supporting the idea that physical activities promote health status in older adults.[32] Overall, our results imply that healthy eating and vigorous sport benefit the elderly by slowing down their biological aging.

Our study has several limitations. First, the positive BioAge gap only indicates that the subject’s BioAge *may* “older” than expected. To validate this aging acceleration effect, we need to show that the positive BioAge gap increases over time in a longitudinal study. However, while CLSA is a longitudinal study, the laboratory measures are currently only available at baseline. Therefore, it is not possible to derive the BioAge gap at follow-up. Second, the observed associations cannot be taken as casualties and our results should be interpreted with caution.

In summary, a BioAge index, estimated with standard lab tests, shows potential as a biomarker for biological aging. The strong associations between the BioAge gap and many current and future disease may lead to potential clinical utilization of this BioAge as a measure of differential aging. We identify several actionable risk factors (current smoking, regular consumption of red meat, chicken, and processed meat, and poor outcomes in nutritional risk screening) and protective factors (quitting smoking, playing tennis, regular consumption of legumes, fruits, and vegetables) that might inform public health policy-making to promote healthy longevity.

### **Acknowledgement**

This research was made possible using the data/biospecimens collected by the Canadian Longitudinal Study on Aging (CLSA). Funding for the Canadian Longitudinal Study on Aging (CLSA) is provided by the Government of Canada through the Canadian Institutes of Health Research (CIHR) under grant reference: LSA 94473 and the Canada Foundation for Innovation, as well as the following provinces, Newfoundland, Nova Scotia, Quebec, Ontario, Manitoba, Alberta, and British Columbia. This research has been conducted using the CLSA dataset (Comprehensive Cohort), under Application Number 1906013. The CLSA is led by Drs. Parminder Raina, Christina Wolfson and Susan Kirkland.

### **Statement of Ethics**

Ethical approval and consent were not required as this study was based on publicly available data from the Canadian Longitudinal Study on Aging (CLSA) ([www.clsa-elcv.ca](http://www.clsa-elcv.ca)) project. Researchers who meet the criteria have the access to de-identified CLSA data. To learn more about the accessibility of CLSA data sets, see <https://www.clsa-elcv.ca/data-access>.

### **Conflict of Interest Statement**

The authors have no conflicts of interest to declare.

### **Funding Sources**

This research was undertaken, in part, thanks to funding from the Canada Research Chairs program, Alberta Innovates, Mental Health Foundation, MITACS Accelerate program, Simon & Martina Sochatsky Fund for Mental Health, the Alberta Synergies in Alzheimer’s and Related Disorders (SynAD) program and University of Alberta Hospital Foundation.

### **Author Contributions**

BC participated in conceptualization, funding acquisition, and supervision. YS participated in formal analysis and writing the original draft. YSL verified the underlying data. YS, YSL, FT, YZ, JH, MW, ES, RAD, RG, XL, AG, SJ, and BC participated in the results interpretation, draft review, and editing of the manuscript. All authors have read and approved the final version of the manuscript.

#### **Data Availability Statement**

Data are available from the Canadian Longitudinal Study on Aging ([www.clsa-elcv.ca](http://www.clsa-elcv.ca)) for researchers who meet the criteria for access to de-identified CLSA data. To learn more about the accessibility of CLSA data sets, see <https://www.clsa-elcv.ca/data-access>. Further enquiries can be directed to the corresponding author.

Accepted Manuscript

Accepted Manuscript

## References

- 1 United States Census Bureau, He W, Goodkind D, Kowal P. An aging world: 2015. Census Bureau. 2016
- 2 Gliberman S. Aging and expenditures on health care. JSTOR; 2021.
- 3 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. DOI: 10.1016/j.arr.2020.101050
- 4 Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol.* 2013;14(10). DOI: 10.1186/gb-2013-14-10-r115
- 5 Earls JC, Rappaport N, Heath L, Wilmanski T, Magis AT, Schork NJ, et al. Multi-Omic Biological Age Estimation and Its Correlation With Wellness and Disease Phenotypes: A Longitudinal Study of 3,558 Individuals. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences.* 2019;74. DOI: 10.1093/gerona/glz220
- 6 Skirbekk VF, Staudinger UM, Cohen JE. How to Measure Population Aging? the Answer Is Less than Obvious: A Review. *Gerontology.* 2019;65(2). DOI: 10.1159/000494025
- 7 Ng TP, Zhong X, Gao Q, Gwee X, Chua DQL, Larbi A. Socio-Environmental, Lifestyle, Behavioural, and Psychological Determinants of Biological Ageing: The Singapore Longitudinal Ageing Study. *Gerontology.* 2020;66(6). DOI: 10.1159/000511211
- 8 Putin E, Mamoshina P, Aliper A, Korzinkin M, Moskalev A, Kolosov A, et al. Deep biomarkers of human aging: Application of deep neural networks to biomarker development. *Ageing.* 2016;8(5). DOI: 10.18632/aging.100968
- 9 Mamoshina P, Kochetov K, Cortese F, Kovalchuk A, Aliper A, Putin E, et al. Blood biochemistry analysis to detect smoking status and quantify accelerated aging in smokers. *Sci Rep.* 2019;9(1). DOI: 10.1038/s41598-018-35704-w
- 10 Raina PS, Wolfson C, Kirkland SA, Griffith LE, Oremus M, Patterson C, et al. The Canadian Longitudinal Study on Aging (CLSA). *Canadian Journal on Aging.* 2009;28(3). DOI: 10.1017/S0714980809990055
- 11 de Lange AMG, Cole JH. Commentary: Correction procedures in brain-age prediction. *Neuroimage Clin.* 2020;26. DOI: 10.1016/j.nicl.2020.102229
- 12 Gao X, Zhang Y, Breitling LP, Brenner H. Relationship of tobacco smoking and smoking-related DNA methylation with epigenetic age acceleration. *Oncotarget.* 2016;7(30). DOI: 10.18632/oncotarget.9795
- 13 Carrard G, Dieu M, Raes M, Toussaint O, Friguet B. Impact of ageing on proteasome structure and function in human lymphocytes. *Int J Biochem Cell Biol.* 2003;35(5):728–39.
- 14 Kubota K, Shirakura T, Orui T, Muratani M, Maki T, Tamura J, et al. Changes in the blood cell counts with aging. *Nihon Ronen Igakkai Zasshi.* 1991;28(4):509–14.
- 15 Dong MH, Bettencourt R, Barrett-Connor E, Lomboa R. Alanine aminotransferase decreases with age: The Rancho Bernardo study. *PLoS One.* 2010;5(12). DOI: 10.1371/journal.pone.0014254
- 16 CASALE G, BONORA C, MIGLIAVACCA A, ZURITA IE, de NICOLA P. Serum ferritin and ageing. *Age Ageing.* 1981;10(2):119–22.
- 17 Schöttker B, Hagen L, Zhang Y, Gaò X, Holleczeck B, Gao X, et al. Serum 25-hydroxyvitamin D levels as an aging marker: Strong associations with age and all-cause mortality independent from

telomere length, epigenetic age acceleration, and 8-isoprostane levels. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*. 2019;74(1). DOI: 10.1093/gerona/gly253

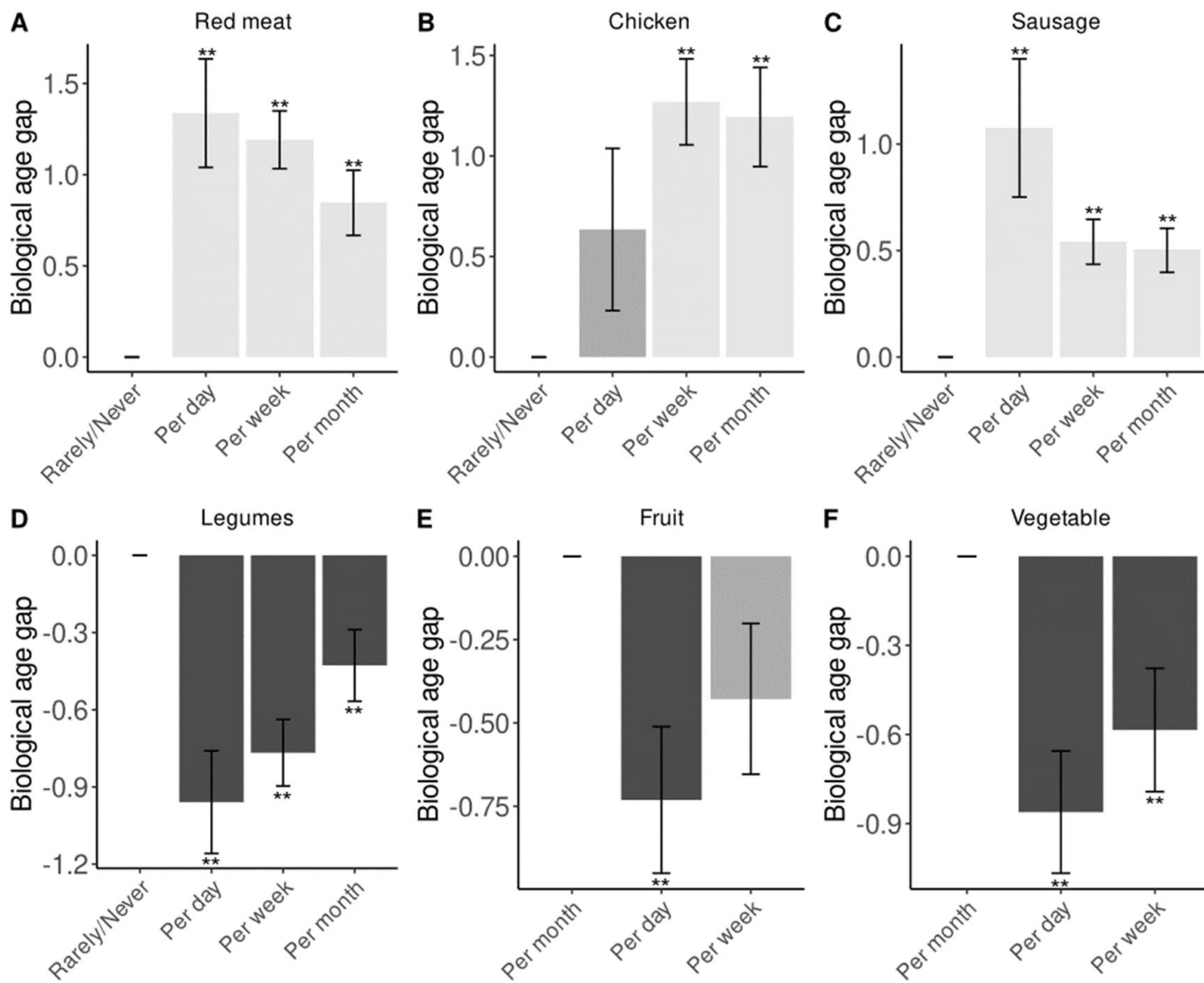
- 18 Steptoe A, Zaninotto P. Lower socioeconomic status and the acceleration of aging: An outcome-wide analysis. *Proc Natl Acad Sci U S A*. 2020;117(26). DOI: 10.1073/pnas.1915741117
- 19 Pani LN, Korenda L, Meigs JB, Driver C, Chamany S, Fox CS, et al. Effect of aging on A1C levels in individuals without diabetes: Evidence from the Framingham Offspring Study and the National Health and Nutrition Examination Survey 2001-2004. *Diabetes Care*. 2008;31(10). DOI: 10.2337/dc08-0577
- 20 Ericsson S, Berglund L, Frostegård J, Einarsson K, Angelin B. The influence of age on low density lipoprotein metabolism: Effects of cholestyramine treatment in young and old healthy male subjects. *J Intern Med*. 1997;242(4). DOI: 10.1046/j.1365-2796.1997.00238.x
- 21 Roetker NS, Pankow JS, Bressler J, Morrison AC, Boerwinkle E. Prospective study of epigenetic age acceleration and incidence of cardiovascular disease outcomes in the ARIC study (Atherosclerosis Risk in Communities). *Circ Genom Precis Med*. 2018;11(3). DOI: 10.1161/CIRCGEN.117.001937
- 22 Muhandiramge J, Orchard S, Haydon A, Zalcborg J. The acceleration of ageing in older patients with cancer. *J Geriatr Oncol*. 2021;12(3). DOI: 10.1016/j.jgo.2020.09.010
- 23 Kooman JP, Kotanko P, Schols AMWJ, Shiels PG, Stenvinkel P. Chronic kidney disease and premature ageing. *Nat Rev Nephrol*. 2014;10(12). DOI: 10.1038/nrneph.2014.185
- 24 Monickaraj F, Aravind S, Gokulakrishnan K, Sathishkumar C, Prabu P, Prabu D, et al. Accelerated aging as evidenced by increased telomere shortening and mitochondrial DNA depletion in patients with type 2 diabetes. *Mol Cell Biochem*. 2012;365(1-2). DOI: 10.1007/s11010-012-1276-0
- 25 Gesing A, Lewiński A, Karbownik-Lewińska M. The thyroid gland and the process of aging: What is new? *Thyroid Res*. 2012;5(1). DOI: 10.1186/1756-6614-5-16
- 26 MacNee W. Is chronic obstructive pulmonary disease an accelerated aging disease? *Annals of the American Thoracic Society*. 2016. DOI: 10.1513/AnnalsATS.201602-124AW
- 27 Li T, Le W. Biomarkers for Parkinson's disease: How good are they? *Neurosci Bull*. 2020;36(2). DOI: 10.1007/s12264-019-00433-1
- 28 Collier TJ, Kanaan NM, Kordower JH. Ageing as a primary risk factor for Parkinson's disease: Evidence from studies of non-human primates. *Nat Rev Neurosci*. 2011;12(6). DOI: 10.1038/nrn3039
- 29 Horvath S, Erhart W, Brosch M, Ammerpohl O, von Schönfels W, Ahrens M, et al. Obesity accelerates epigenetic aging of human liver. *Proc Natl Acad Sci U S A*. 2014;111(43). DOI: 10.1073/pnas.1412759111
- 30 Levine M, Crimmins E. Not all smokers die young: A model for hidden heterogeneity within the human population. *PLoS One*. 2014;9(2). DOI: 10.1371/journal.pone.0087403
- 31 USDA. Dietary guidelines for Americans, 2020-2025. 9th Edition. *Am J Clin Nutr*. 2020;34(1).
- 32 McPhee JS, French DP, Jackson D, Nazroo J, Pendleton N, Degens H. Physical activity in older age: Perspectives for healthy ageing and frailty. *Biogerontology*. 2016;17(3). DOI: 10.1007/s10522-016-9641-0

## Figure Legends

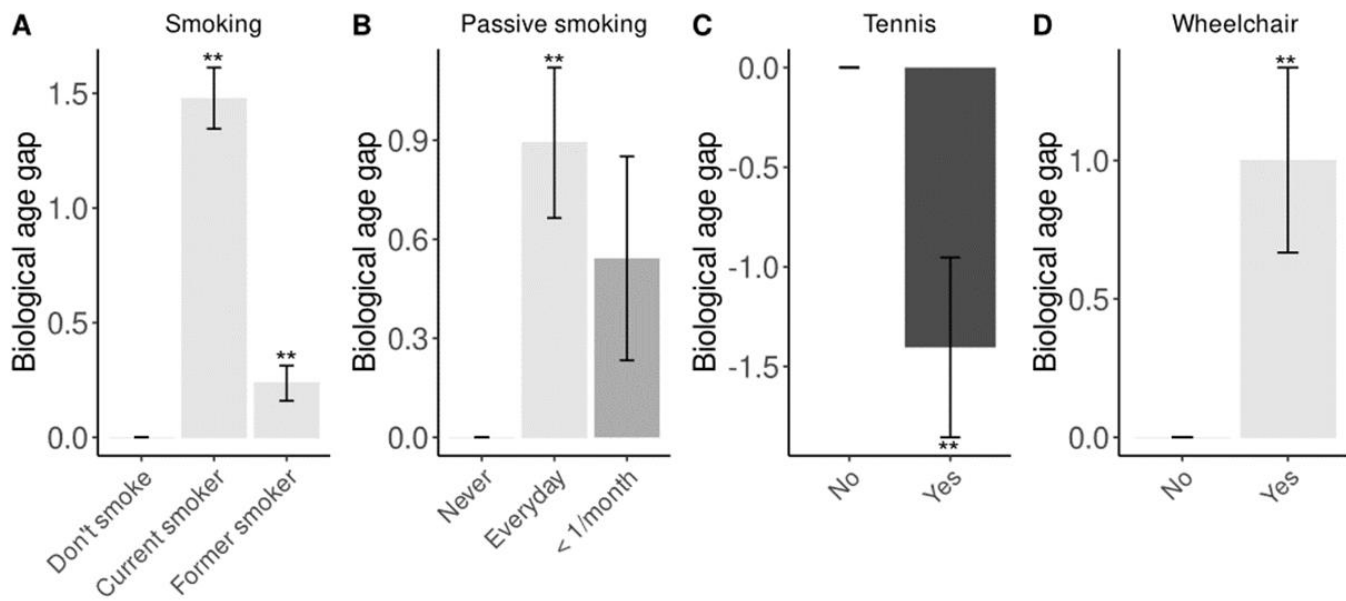
Fig. 1. Dietary factors associated with differential aging indicated by biological age gap. The level with zero height was the reference level in the association analysis. Gray white indicates significant risk factors, while black indicates significant protective factors. Red meat: the frequency of beef, pork consumption; Chicken: the frequency of chicken or turkey consumption; Sausage: the frequency of sausages, hot dogs, ham, smoked meat, bacon, consumption; Legumes: the frequency of legumes consumption, including beans, peas, lentils; Fruit: fresh, frozen or canned fruit; Vegetable: other vegetables except for carrots, potatoes or salad; \* indicates the p-value is less than .05, while \*\* indicates the p-value is less than .01.

Fig. 2. Lifestyle factors associated with the differential aging indicated by biological age gap. Smoking: smoking status; Tennis: participated in tennis; Wheelchair: used a wheelchair or motorized cart in the last month; Passive Smoking: frequency of passive smoking exposure at home.

Fig. 3. Current disease conditions associated with the differential aging indicated by biological age gap. Some diseases are combined to save space. Chronic Conditions: reported chronic conditions; Heart) heart: heart disease (including congestive heart failure); vascular: peripheral vascular disease or poor circulation in limbs; aneurysm: thoracic, abdominal, or cerebral aneurysm; angina: angina; stroke: stroke or cerebrovascular accident; heart attack: heart attack or myocardial infarction; coronary artery: coronary artery bypass surgery; Cancer) kidney: kidney cancer; lung: lung cancer; lymphoid: lymphoid, hematopoietic and related tissue; Non-Hodgkin: Non-Hodgkin lymphoma cancer; Lung) emphysema: emphysema, chronic bronchitis, COPD, or chronic changes in lungs due to smoking; pneumonia: pneumonia in the last year; Shuffling gait/ Poor balancing: shuffling gait or poor balance; Other) Diabetes: diabetes, borderline diabetes or high blood sugar; kidney: kidney disease or failure; Thyroid Gland: under-active thyroid gland.

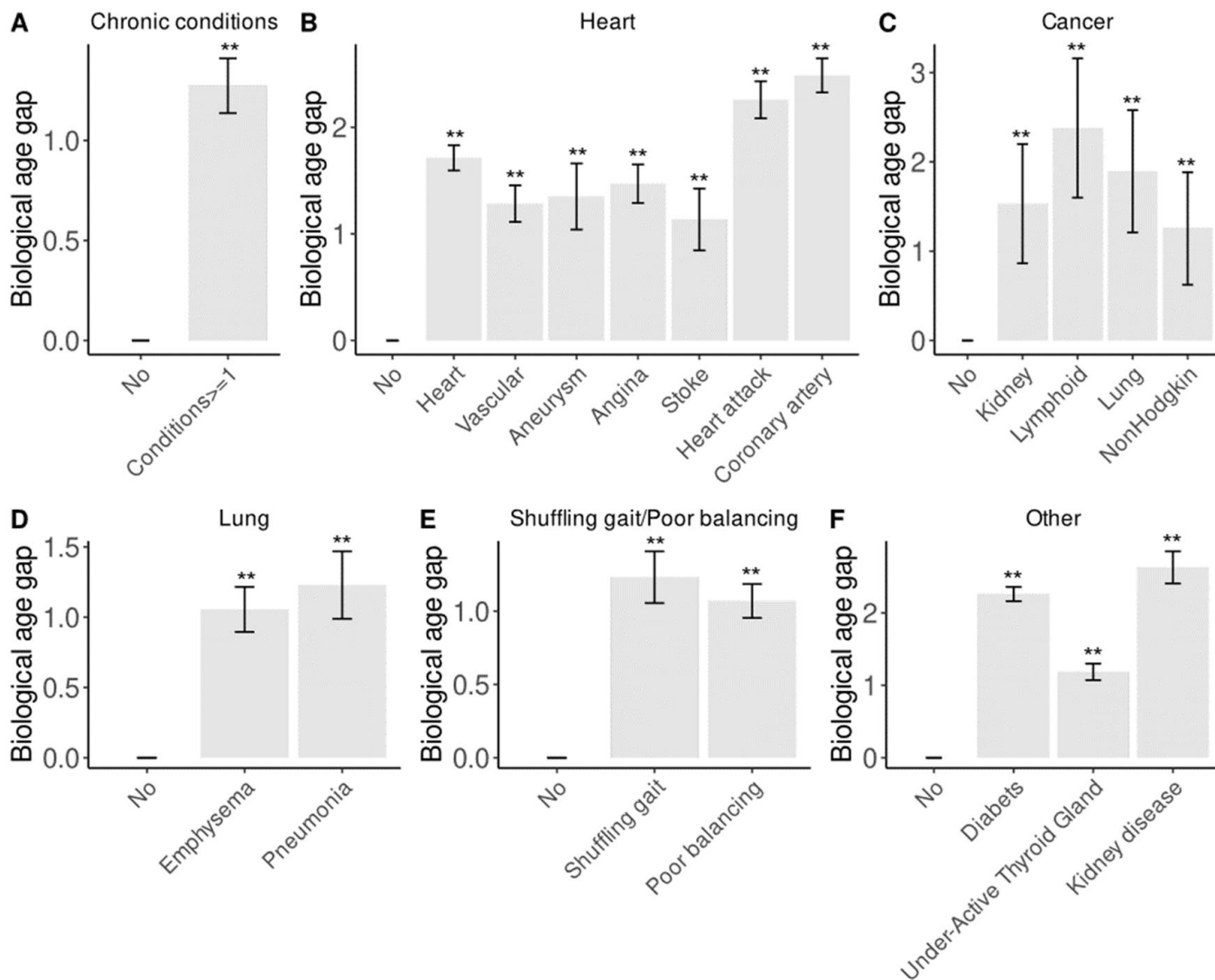


ACCEPTED



Accepted Manuscript





Accept

**Table 1.** Modifiable risk factors for differential aging.

Variable	Level	Effect size (years)	Std	P-value
Smoking status	Don't smoke (reference level)	0	NA	NA
	Currently smoker	1.48	0.133	< 0.001
	Former smoker	0.24	0.076	0.003
Passive smoking	Never (reference level)	0	NA	NA
	Everyday	0.89	0.228	< 0.001
	< 1/Month	0.54	0.309	0.082
Red meat	Rarely/never (reference level)	0	NA	NA
	Per day	1.34	0.297	< 0.001
	Per week	1.19	0.158	< 0.001
	Per month	0.85	0.178	< 0.001
White meat	Rarely/never (reference level)	0	NA	NA
	Per day	0.63	0.403	0.119
	Per week	1.27	0.214	< 0.001
	Per month	1.19	0.247	< 0.001
Processed meat	Rarely/never (reference level)	0	NA	NA

	Per day	1.08	0.324	0.001
	Per week	0.54	0.106	< 0.001
	Per month	0.50	0.103	< 0.001
Legumes	Rarely/never	0	NA	NA
	(reference level)			
	Per day	-0.96	0.199	< 0.001
	Per week	-0.77	0.129	< 0.001
	Per month	-0.43	0.139	0.003
Vegetable	Per month	0	NA	NA
	(reference level)			
	Per day	-0.86	0.206	< 0.001
	Per week	-0.58	0.208	0.006
Fruit	Per month	0	NA	NA
	(reference level)			
	Per day	-0.73	0.220	0.001
	Per week	-0.43	0.226	0.061
High	No (reference level)	0	NA	NA
Nutritional	Yes	0.51	0.079	< 0.001
Risk				
Playing	No (reference level)	0	NA	NA
tennis	Yes	-1.40	0.451	0.002