ORIGINAL RESEARCH

# Biological age as a health index for mortality and major age-related disease incidence in Koreans: National Health Insurance Service – Health screening 11-year follow-up study

Young Gon Kang[1]
Eunkyung Suh[2]
Jae-woo Lee[3]
Dong Wook Kim[4]
Kyung Hee Cho[5]
Chul-Young Bae[1]

[1]Department of R&D, MediAge Research Center, Seongnam, Republic of South Korea; [2]Department of Family Medicine, College of Medicine, CHA University, Chaum, Seoul, Republic of South Korea; [3]Department of Family Medicine, College of Medicine, Chungbuk National University, Cheongju, Republic of South Korea; [4]Department of Policy Research Affairs, National Health Insurance Service Ilsan Hospital, Goyang, Republic of South Korea; [5]Department of Family Medicine, National Health Insurance Service Ilsan Hospital, Goyang, Republic of South Korea

Correspondence: Young Gon Kang
MediAge Research Center, 660, Daewangpangyo-ro, Bundang-gu, Seongnam-si, Gyeonggi-do, 13494, Republic of South Korea
Tel +82 2 555 6438
Fax +82 2 555 6439
Email androkyg@naver.com

**Purpose:** A comprehensive health index is needed to measure an individual's overall health and aging status and predict the risk of death and age-related disease incidence, and evaluate the effect of a health management program. The purpose of this study is to demonstrate the validity of estimated biological age (BA) in relation to all-cause mortality and age-related disease incidence based on National Sample Cohort database.

**Patients and methods:** This study was based on National Sample Cohort database of the National Health Insurance Service – Eligibility database and the National Health Insurance Service – Medical and Health Examination database of the year 2002 through 2013. BA model was developed based on the National Health Insurance Service – National Sample Cohort (NHIS – NSC) database and Cox proportional hazard analysis was done for mortality and major age-related disease incidence.

**Results:** For every 1 year increase of the calculated BA and chronological age difference, the hazard ratio for mortality significantly increased by 1.6% (1.5% in men and 2.0% in women) and also for hypertension, diabetes mellitus, heart disease, stroke, and cancer incidence by 2.5%, 4.2%, 1.3%, 1.6%, and 0.4%, respectively ($p<0.001$).

**Conclusion:** Estimated BA by the developed BA model based on NHIS – NSC database is expected to be used not only as an index for assessing health and aging status and predicting mortality and major age-related disease incidence, but can also be applied to various health care fields.

**Keywords:** health, aging, biological age, comprehensive index

## Introduction

Age is known as a general index by which we can evaluate the health and aging status of an individual, and parameters that show certain patterns according to age are called biomarkers of aging.[1] But chronological age (CA) that we usually call "age" is determined by the simple flow of time, so it has limitations in accurate evaluation of one's physiological function, health, or aging status.[2] Thus, biological age (BA) estimated by the biomarkers, which clearly reflect the health status with aging has been utilized as an index for evaluating one's physiologic function, health status, and aging state.[3]

BA is a term that quantifies and digitizes the aging state based on the characteristic of aging that body function or physiological function shows various changes with age.[3,4] The difference of body or physiological function change among individuals as we age show as different BAs among the same age group.

Recently, the importance of utilizing BA in health care field is being emphasized for compensating limitations of the dichotomy categorization of disease diagnosis.[5,6] The proposed utilization and importance of BA in the aspect of health care are as follows.[7–10] First, it can be used as a continuous index for detecting subtle changes in the health and aging status. Second, it can be used as an easy communication index in which anyone can comprehend one's health and aging status as "new age" of BA. Third, it can also provide us with a management index to evaluate an individual's health management program.

Along with the rise in the elderly population and overall interest in health promotion, the importance and the need for a health index by which we can easily evaluate one's health status and utilize in health care management is being emphasized not only in the medical field, but also in the related ubiquitous health care fields.[7] Many studies have shown that BA developed for various purposes can be utilized as an evaluation index of mortality, morbidity, incidence of disease, and disability.[3,11–17] Still, long-term studies presenting validity of BA as a health and aging index are very limited and scarce.

The national health insurance system in Korea is a single social insurance that includes all Korean population as recipients and 97% of the population is enlisted. In 1995, national health and medical examination was started with the purpose of preventing chronic disease and decreasing mortality rate by health promotion.[18,19] But despite the various efforts for disease prevention and health promotion, evaluation of overall health status and the effect of lifestyle modification are unsatisfactory. Thus, in this study, we developed the health index based on the National Health Insurance Service – National Sample Cohort (NHIS – NSC) database. Using the developed BA model, we analyzed mortality and major age-related disease incidence. We hope to provide the basis for utilizing BA as a comprehensive health index in various health care fields.

## Patients and methods
### Constitution of data
This study was based on NHIS – NSC database, eligibility database of the NHIS, and Medical and Health Examination database of the year 2002 through 2013. Sample cohort database was made by sampling the data of 2.2% of the National Health Insurance members of the year 2002–2003 according to sex, age, and income by stratified sampling method and was followed up to the year 2013.

## Subjects
NHIS – NSC database was categorized and analyzed for BA model development and validity assessment as follows. For BA model development, we analyzed the data of 484,724 subjects, which satisfied the inclusion criteria out of 665,594 subjects who did the NHIS – Medical and Health Examination through the years 2009–2013. For validity assessment of the developed BA model, we analyzed the data of 769,858 subjects who did the NHIS – Medical and Health Examination through the years 2002–2008. A total of 682,001 cases were analyzed excluding subjects diagnosed with age-related diseases, such as hypertension, diabetes mellitus, heart disease, stroke, and cancer before the baseline examination, and also subjects with parameters of extreme values. The average follow-up period (mean ± standard deviation [SD]) was 7.85±3.22 (7.99±3.20 for men and 7.99±3.20 for women) years for the subjects evaluated for mortality and incidence for major age-related diseases. This study was approved exempt from review by the Institutional Review Board of the National Health Insurance Service Ilsan Hospital (NHIMC 2016-20-019). The reason for exemption from Institutional Review Board was because our study analyzed data with which the identity of the subjects could not be identified. The study data were composed of those whose identities were not identifiable and therefore met the criteria for exemption from Institutional Review Board.

## Parameters
A total of 18 test results of the NHIS – Health Screening were included to develop BA model. Physical measurements included height (HT), weight (WT), body mass index (BMI), waist circumference (WC), systolic blood pressure (SBP), and diastolic blood pressure. Blood tests included fasting blood sugar (FBS), total cholesterol (TC), triglyceride (TG), high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol, hemoglobin (Hgb), serum creatinine, estimated glomerular filtration rate (eGFR), aspartate aminotransferase (AST), alanine aminotransferase (ALT), gamma glutamyl transpeptidase (r-GTP), and urine test including urine protein.

Data of the date and cause of death in Korea were derived from the Korean Statistics Information Service database. Extrinsic causes of death, such as suicide or accident, were excluded from the study.

The incidence of major age-related diseases, such as hypertension, diabetes mellitus, heart disease, stroke, and cancer, was evaluated using the National Health Insurance

Review and Assessment Service associated data classified according to the International Statistical Classification of Diseases and Related Health Problems (10th Edition, diabetes mellitus: E10–E14, hypertension: I10–I15, heart disease: I21–I25, stroke: I63).

## Inclusion criteria

In this study, we determined the inclusion criteria considering the means and standard deviations of the collected parameters and the normal range established by the American Medical Association (AMA). The inclusion criteria in the supplements are presented within mean ± 3–4 SD with some exceptions depending on the distribution of study parameters and disease criteria of AMA (Table 1).

## Statistical analysis

Principal component analysis (PCA) was done for the development of BA model. Correlation analysis of age and the measured parameters were done for the selection of biomarkers of aging and assessment of redundancy. BA score (BAS) was developed using PCA based on the selected biomarkers. Individual BAS was transformed into terms of years (BA)

for comparison with CA. The transformed BA makes under-estimations of BA in the above-average group and makes overestimations in the below-average group. The final BA was calculated after correction for this systemic error.

Cox proportional hazard analysis was done for mortality and major age-related disease incidence based on BA according to the developed BA model. Statistical analyses were performed with SAS software for Windows (v.9.1; Cary, NC, USA). Significance was confirmed at a 2-sided 0.05 level, unless otherwise specified.

## Results

### General characteristics of the subjects

A total of 484,724 cases were analyzed for the development of BA model. Among the 484,724 subjects, there were 260,315 men and 224,409 women. The average age of the subjects was 50.75±14.11 years, 50.12±13.92 years in men and 51.48±14.29 years in women. Table 2 shows the average

**Table 1** Inclusion criteria for parameters

| Parameters | Inclusion criteria | | |
|---|---|---|---|
| Age (year) | 20.00 | ≤ | |
| HT (cm) | NA | | |
| WT (kg) | NA | | |
| BMI (kg/m²) | 15.00 | ~ | 30.00 |
| WC (cm) | 60.00 | ~ | 105.00 |
| SBP (mmHg) | 80.00 | ~ | 160.00 |
| DBP (mmHg) | 50.00 | ~ | 100.00 |
| FBS (mmol/L) | 2.78 | ~ | 7.77 |
| TC (mmol/L) | 1.30 | ~ | 6.73 |
| TG (mmol/L) | 0.57 | ~ | 4.52 |
| HDL-C (mmol/L) | 0.52 | ~ | 2.33 |
| LDL-C (mmol/L) | | < | 4.92 |
| Hgb (g/L) | 100.00 | ~ | 180.00 |
| u-Prot (negative, trace, + ~ ++++) | Negative | ~ | ++ |
| Cr (μmol/L) | 35.36 | ~ | 176.80 |
| AST (IU/L) | | < | 60.00 |
| ALT (IU/L) | | < | 80.00 |
| r-GTP (IU/L) | | < | 120.00 |

**Notes:** SI conversion factors: to convert FBS to mmol/L, multiply by 0.0555; TC, HDL-C and LDL-C to mmol/L, multiply by 0.0259; triglyceride to mmol/L, multiply by 0.0113; Hgb to g/L, multiply by 10.0; creatinine to μmol/L, multiply by 88.4.
**Abbreviations:** ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; Cr, creatinine; DBP, diastolic blood pressure; FBS, fasting blood sugar; HDL-C, high-density lipoprotein cholesterol; Hgb, hemoglobin; HT, height; LDL-C, low-density lipoprotein cholesterol; NA, not available; r-GTP, gamma glutamyl transpeptidase; SBP, systolic blood pressure; TC, total cholesterol; TG, triglyceride; u-Prot, urine protein; WC, waist circumference; WT, weight.

**Table 2** General characteristics of the biological age model subjects

| Parameters | Sexes combined (n=484,724) | Men (n=260,315) | Women (n=224,409) |
|---|---|---|---|
| Age (years) | 50.75±14.11 | 50.12±13.92 | 51.48±14.29 |
| HT (cm) | 163.25±9.12 | 169.37±6.39 | 156.15±6.19 |
| WT (kg) | 62.76±10.33 | 68.36±9.16 | 56.26±7.38 |
| BMI (kg/m²) | 23.47±2.72 | 23.79±2.60 | 23.09±2.81 |
| WC (cm) | 80.02±8.19 | 83.14±7.08 | 76.39±7.89 |
| SBP (mmHg) | 121.18±13.18 | 123.00±12.29 | 119.05±13.85 |
| DBP (mmHg) | 75.34±8.81 | 76.74±8.40 | 73.72±9.00 |
| FBS (mmol/L) | 5.24±0.71 | 5.30±0.74 | 5.16±0.67 |
| TC (mmol/L) | 4.96±0.80 | 4.92±0.80 | 5.01±0.80 |
| TG (mmol/L) | 1.42±0.72 | 1.53±0.77 | 1.28±0.64 |
| HDL-C (mmol/L) | 1.39±0.32 | 1.33±0.30 | 1.46±0.32 |
| LDL-C (mmol/L) | 2.92±0.75 | 2.89±0.75 | 2.95±0.75 |
| Hgb (g/L) | 139.25±14.74 | 148.40±11.63 | 128.63±10.14 |
| u-Prot (negative, trace, + ~ ++++) | 1.05±0.27 | 1.05±0.27 | 1.05±0.27 |
| Cr (μmol/L) | 79.62±17.61 | 88.92±15.32 | 68.83±13.44 |
| eGFR (mL/min/1.73 m²) | 102.27±29.35 | 87.69±19.76 | 119.20±29.57 |
| AST (IU/L) | 23.51±7.14 | 24.64±7.29 | 22.19±6.72 |
| ALT (IU/L) | 22.11±11.19 | 25.00±11.92 | 18.77±9.21 |
| r-GTP (IU/L) | 29.22±21.98 | 36.93±24.60 | 20.28±13.89 |

**Notes:** SI conversion factors: To convert FBS to mmol/L, multiply by 0.0555; TC, HDL-C and LDL-C to mmol/L, multiply by 0.0259; triglyceride to mmol/L, multiply by 0.0113; Hgb to g/L, multiply by 10.0; creatinine to μmol/L, multiply by 88.4.
**Abbreviations:** ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; Cr, creatinine; DBP, diastolic blood pressure; eGFR, estimated glomerular filtration rate, $186.3 \times Cr \, (mg/dL)^{-1.154} \times Age^{-0.203} \times (0.742$ in female); FBS, fasting blood sugar; HDL-C, high-density lipoprotein cholesterol; Hgb, hemoglobin; HT, height; LDL-C, low-density lipoprotein cholesterol; NA, not available; r-GTP, gamma glutamyl transpeptidase; SBP, systolic blood pressure; TC, total cholesterol; TG, triglyceride; u-Prot, urine protein; WC, waist circumference; WT, weight.

(±SD) results of the parameters used in the development of BA model.

## Development of BA model

### Correlation analysis and assessment of redundancy

The parameters that showed negative correlation with age were HT, WT, BMI, HDL-C, Hgb, eGFR, and ALT in men and HT, WT, HDL-C, Hgb, and eGFR, in women. Other parameters showed positive correlation with age. Parameters that were selected for model development were HT, WC, SBP, FBS, Hgb, eGFR, and AST in men and HT, WC, SBP, FBS, TC, TG, HDL-C, eGFR, AST, and r-GTP in women (Table 3).

### Principal component analysis

PCA, including age, was done. A significant positive correlation (0.844 in men and 0.799 in women) was seen and we defined these as principle components. To see the correlation between other parameters and principle components, excluding the influence of age, we performed reanalysis after excluding

**Table 3** Pearson correlation coefficients between age and parameters

| | Sexes combined (n=484,724) | Men (n=260,315) | Women (n=224,409) |
|---|---|---|---|
| HT (cm) | −0.387** | −0.478** | −0.549** |
| WT (kg) | −0.175** | −0.276** | −0.050** |
| BMI (kg/m$^2$) | 0.129** | −0.014** | 0.293** |
| WC (cm) | 0.254** | 0.169** | 0.435** |
| SBP (mmHg) | 0.272** | 0.175** | 0.391** |
| DBP (mmHg) | 0.154** | 0.076** | 0.258** |
| FBS (mmol/L) | 0.222** | 0.212** | 0.249** |
| TC (mmol/L) | 0.126** | 0.030** | 0.229** |
| TG (mmol/L) | 0.114** | 0.006** | 0.285** |
| HDL-C (mmol/L) | −0.125** | −0.068** | −0.214** |
| LDL-C (mmol/L) | 0.134** | 0.052** | 0.223** |
| Hgb (g/L) | −0.180** | −0.321** | −0.042** |
| u-Prot (negative, trace, + ~ ++++) | 0.027** | 0.040** | 0.012** |
| Cr (μmol/L) | 0.007** | 0.005** | 0.088** |
| eGFR (mL/min/1.73 m$^2$) | −0.207** | −0.243** | −0.309** |
| AST (IU/L) | 0.186** | 0.116** | 0.297** |
| ALT (IU/L) | 0.023** | −0.073** | 0.201** |
| r-GTP (IU/L) | 0.023** | 0.006** | 0.125** |

**Notes:** **$p<0.001$. SI conversion factors: To convert FBS to mmol/L, multiply by 0.0555; TC, HDL-C and LDL-C to mmol/L, multiply by 0.0259; triglyceride to mmol/L, multiply by 0.0113; Hgb to g/L, multiply by 10.0; creatinine to μmol/L, multiply by 88.4.
**Abbreviations:** ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; Cr, creatinine; DBP, diastolic blood pressure; eGFR, estimated glomerular filtration rate, 186.3 × Cr (mg/dL)$^{-1.154}$ × Age$^{-0.203}$ × (0.742 in female); FBS, fasting blood sugar; HDL-C, high-density lipoprotein cholesterol; Hgb, hemoglobin; HT, height; LDL-C, low-density lipoprotein cholesterol; NA, not available; r-GTP, gamma glutamyl transpeptidase; SBP, systolic blood pressure; TC, total cholesterol; TG, triglyceride; u-Prot, urine protein; WC, waist circumference; WT, weight; WT (kg)/HT (m)$^2$.

age and confirmed the influence of age on the principle components. The principal component accounted for 23.74% in men and 26.39% in women of the total variance, and 20.79% and 23.79%, respectively, after excluding age (Table 4).

### Construction of BA

Calculated BAS from PCA was converted into BA in the unit of years using the average age and SD of the subjects. The formulas for this process are as follows (1-1 & 1-2, 2-1 & 2-2).

$$\text{BA in men} = (\text{BAS} \times 13.92) + 50.12 \text{ (1-1)}$$

$$\text{BA in women} = (\text{BAS} \times 14.29) + 51.48 \text{ (1-2)}$$

$$\begin{aligned}\text{BA in men} = &-95.12 - 0.19 \times (\text{HT}) + 0.88 \times (\text{WC}) + 0.46 \\ &\times (\text{SBP}) + 0.40 \times (\text{FBS}) + 0.57 \times (\text{Hgb}) \\ &- 0.17 \times (\text{eGFR}) + 0.62 \times (\text{AST}) \text{ (2-1)}\end{aligned}$$

$$\begin{aligned}\text{BA in women} = &17.64 - 0.45 \times (\text{HT}) + 0.49 \times (\text{WC}) \\ &+ 0.24 \times (\text{SBP}) + 0.22 \times (\text{FBS}) + 0.07 \\ &\times (\text{TC}) + 0.07 \times (\text{TG}) - 0.18 \times (\text{HDL-C}) \\ &- 0.07 \times (\text{eGFR}) + 0.44 \times (\text{AST}) + 0.20 \\ &\times (\text{r-GTP}) \text{ (2-2)}\end{aligned}$$

There can be underestimation of the calculated BA when CA is higher than average and overestimation when CA is lower than average. To reduce this systemic error, we used the correction equation shown in a previous study (3-1 & 3-2).[20]

**Table 4** Principal component analysis

| Parameters | First PCA loading | | | |
|---|---|---|---|---|
| | Men | | Women | |
| Age (years) | 0.844 | – | 0.799 | – |
| HT (cm) | −0.617 | −0.129 | −0.553 | −0.465 |
| WC (cm) | 0.282 | 0.653 | 0.631 | 0.643 |
| SBP (mmHg) | 0.373 | 0.597 | 0.561 | 0.559 |
| FBS (mmol/L) | 0.435 | 0.554 | 0.422 | 0.449 |
| TC (mmol/L) | – | – | 0.364 | 0.376 |
| TG (mmol/L) | – | – | 0.568 | 0.632 |
| HDL-C (mmol/L) | – | – | −0.345 | −0.364 |
| Hgb (g/L) | −0.408 | 0.070 | – | – |
| eGFR (mL/min/1.73 m$^2$) | −0.386 | −0.347 | −0.366 | −0.334 |
| AST (IU/L) | 0.290 | 0.472 | 0.466 | 0.488 |
| r-GTP (IU/L) | – | – | 0.377 | 0.461 |
| Eigenvalue | 1.899 | 1.455 | 2.903 | 2.379 |
| Total variance, % | 23.743 | 20.785 | 26.394 | 23.789 |

**Notes:** SI conversion factors: To convert FBS to mmol/L, multiply by 0.0555; TC, HDL-C and LDL-C to mmol/L, multiply by 0.0259; triglyceride to mmol/L, multiply by 0.0113; Hgb to g/L, multiply by 10.0; creatinine to μmol/L, multiply by 88.4.
**Abbreviations:** AST, aspartate aminotransferase; eGFR, estimated glomerular filtration rate, 186.3 × Cr (mg/dL)$^{-1.154}$ × Age$^{-0.203}$ × (0.742 in female); FBS, fasting blood sugar; HDL-C, high-density lipoprotein cholesterol; Hgb, hemoglobin; HT, height; LDL-C, low-density lipoprotein cholesterol; r-GTP, gamma glutamyl transpeptidase; SBP, systolic blood pressure; TC, total cholesterol; TG, triglyceride; u-Prot, urine protein; WC, waist circumference.

The correlation coefficients between corrected BA and CA calculated from the previously corrected equation in men and women were 0.730 and 0.792, respectively.

$$\text{Corrected BA in men} = -127.6 - 0.19 \times (HT) + 0.88 \times (WC) + 0.46 \times (SBP) + 0.40 \times (FBS) + 0.57 \times (Hgb) - 0.17 \times (eGFR) + 0.62 \times (AST) + 0.65 \times (Age) \quad (3\text{-}1)$$

$$\text{Corrected BA in women} = -0.01 - 0.45 \times (HT) + 0.49 \times (WC) + 0.24 \times (SBP) + 0.22 \times (FBS) + 0.07 \times (TC) + 0.07 \times (TG) - 0.18 \times (HDL\text{-}C) - 0.07 \times (eGFR) + 0.44 \times (AST) + 0.20 \times (r\text{-}GTP) + 0.36 \times (Age) \quad (3\text{-}2)$$

## Mortality and major age-related disease incidence

### Frequency analysis on mortality and major age-related disease incidence

After analysis of a total of 749,861 cases, mortality was 16,704 (4.1%) cases out of 410,584 cases in men and 7,563 (2.2%) cases out of 339,277 cases in women. Major age-related disease incidence is shown in Table 5.

### Influence of BA difference on mortality

For every 1-year change in the calculated BA and CA difference, mortality changed by 1.6%. Also, the hazard ratios of mortality for BA difference grew weaker with increasing CA in men and women. The hazard ratios for mortality by gender and CA are shown in Table 6.

### Influence of BA difference on major age-related disease incidence

For every 1-year change in the calculated BA and CA difference, major age-related disease incidence changed by 2.5% for hypertension, 4.2% for diabetes mellitus, 1.3% for heart disease, 1.6% for stroke, and 0.4% for cancer. Also, the hazard ratios of disease incidence for BA difference grew weaker with increasing CA in men and women except for cancer incidence in women. The hazard ratios for major age-related disease incidence by gender and CA are shown in Table 7.

## Discussion

Our results showed that for every 1-year increase in the calculated BA and CA difference, hazard ratio for mortality

**Table 5** Frequency of mortality and major age-related diseases

| | Sexes combined | | | | | | Men | | | | | | Women | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Event | | | | Total | | Event | | | | Total | | Event | | | | Total | |
| | Yes | | No | | | | Yes | | No | | | | Yes | | No | | | |
| | n | % | n | % | n | % | n | % | n | % | n | % | n | % | n | % | n | % |
| Death | 23,191 | 3.4 | 658,810 | 96.6 | 682,001 | 100.0 | 15,978 | 4.2 | 361,315 | 95.8 | 377,293 | 100.0 | 7,213 | 2.4 | 297,495 | 97.6 | 304,708 | 100.0 |
| Hypertension | 116,887 | 19.9 | 469,226 | 80.1 | 586,113 | 100.0 | 69,553 | 21.1 | 259,839 | 78.9 | 329,392 | 100.0 | 47,334 | 18.4 | 209,387 | 81.6 | 256,721 | 100.0 |
| Diabetes mellitus | 46,018 | 7.3 | 585,361 | 92.7 | 631,379 | 100.0 | 28,497 | 8.1 | 321,425 | 91.9 | 349,922 | 100.0 | 17,521 | 8.2 | 263,936 | 93.8 | 281,457 | 100.0 |
| Heart diseases | 39,829 | 5.9 | 635,471 | 94.1 | 675,300 | 100.0 | 23,220 | 6.2 | 350,727 | 93.8 | 373,947 | 100.0 | 16,609 | 5.5 | 284,744 | 94.5 | 301,353 | 100.0 |
| Stroke | 6,119 | 0.9 | 670,840 | 99.1 | 676,959 | 100.0 | 3,722 | 1.0 | 371,369 | 99.0 | 375,091 | 100.0 | 2,397 | 0.8 | 299,471 | 99.2 | 301,868 | 100.0 |
| (Malignant) cancer | 55,448 | 8.2 | 616,694 | 91.8 | 672,142 | 100.0 | 33,783 | 9.1 | 338,304 | 90.9 | 372,087 | 100.0 | 21,665 | 7.2 | 278,390 | 92.8 | 300,055 | 100.0 |

**Table 6** Hazard ratios for all-cause mortality by BA difference

| Age at baseline health checkup | Sexes combined* | Men** | Women** |
|---|---|---|---|
| All ages (years) | 1.016 (1.015–1.016) | 1.015 (1.014–1.016) | 1.020 (1.018–1.021) |
| <50 | 1.027 (1.026–1.029) | 1.027 (1.025–1.028) | 1.034 (1.029–1.039) |
| 50–59 | 1.019 (1.018–1.021) | 1.018 (1.017–1.020) | 1.025 (1.021–1.029) |
| 60–69 | 1.014 (1.013–1.015) | 1.013 (1.012–1.014) | 1.020 (1.018–1.023) |
| 70+ | 1.010 (1.008–1.011) | 1.008 (1.006–1.009) | 1.016 (1.014–1.019) |

**Notes:** *Adjusted for age and sex. **Adjusted for age. Hazard ratios are for a 1-year increase in the difference between calculated BA and CA. All hazard ratios are $p<0.001$.
**Abbreviations:** BA, biological age; CA, chronological age.

significantly increased by 1.6% (1.5% in men and 2.0% in women) and also for hypertension, diabetes mellitus, heart disease, stroke, and cancer incidence by 2.5%, 4.2%, 1.3%, 1.6%, and 0.4%, respectively. These results are similar to other previous studies.[14,21–23]

Also, subgroups were stratified by age of baseline health checkup for evaluation of the risk of mortality and major age-related disease incidence with increasing chronologic age (Tables 6 and 7). The hazard ratios for BA difference of mortality and disease incidence were weaker in the older group compared with the younger group in both men and

women except for cancer incidence in women. First, this may have been due to the high variability of BA in the same age in the older group than the younger group. Second, this may have been due to much higher mortality and major age-related disease incidence of the older group. Third, this may have been due to the shorter time interval of the mortality and major age-related disease incidence event among the older group.[14,24]

It is known that the physiological changes associated with aging increases the susceptibility to disease and mortality, and accelerated aging advances disease onset, morbidity,

**Table 7** Hazard ratios for major age-related disease incidence by BA difference

| | Age at baseline health checkup | Sexes combined* | Men** | Women** |
|---|---|---|---|---|
| Hypertension | All ages (years) | 1.025 (1.025–1.026) | 1.024 (1.024–1.024) | 1.031 (1.030–1.032) |
| | <50 | 1.030 (1.030–1.030) | 1.028 (1.028–1.028) | 1.040 (1.039–1.040) |
| | 50–59 | 1.023 (1.022–1.024) | 1.022 (1.021–1.022) | 1.029 (1.027–1.030) |
| | 60–69 | 1.019 (1.018–1.020) | 1.018 (1.017–1.018) | 1.023 (1.022–1.024) |
| | 70+ | 1.016 (1.015–1.018) | 1.015 (1.013–1.017) | 1.018 (1.015–1.020) |
| Diabetes mellitus | All ages (years) | 1.042 (1.042–1.043) | 1.040 (1.039–1.040) | 1.050 (1.049–1.050) |
| | <50 | 1.047 (1.046–1.047) | 1.044 (1.044–1.045) | 1.058 (1.057–1.059) |
| | 50–59 | 1.041 (1.040–1.042) | 1.039 (1.038–1.039) | 1.048 (1.047–1.049) |
| | 60–69 | 1.034 (1.033–1.035) | 1.031 (1.030–1.032) | 1.041 (1.039–1.042) |
| | 70+ | 1.032 (1.030–1.034) | 1.029 (1.026–1.031) | 1.038 (1.035–1.041) |
| Heart diseases | All ages (years) | 1.013 (1.013–1.014) | 1.012 (1.011–1.013) | 1.017 (1.016–1.018) |
| | <50 | 1.017 (1.016–1.018) | 1.016 (1.015–1.017) | 1.025 (1.022–1.027) |
| | 50–59 | 1.013 (1.012–1.014) | 1.012 (1.010–1.013) | 1.017 (1.016–1.019) |
| | 60–69 | 1.008 (1.007–1.009) | 1.007 (1.006–1.008) | 1.011 (1.009–1.013) |
| | 70+ | 1.007 (1.005–1.009) | 1.006 (1.004–1.009) | 1.009 (1.005–1.012) |
| Stroke | All ages (years) | 1.016 (1.014–1.017) | 1.015 (1.013–1.016) | 1.019 (1.016–1.021) |
| | <50 | 1.020 (1.017–1.023) | 1.019 (1.016–1.022) | 1.027 (1.020–1.034) |
| | 50–59 | 1.018 (1.016–1.021) | 1.018 (1.015–1.021) | 1.019 (1.013–1.025) |
| | 60–69 | 1.012 (1.010–1.015) | 1.010 (1.008–1.013) | 1.018 (1.014–1.022) |
| | 70+ | 1.008 (1.005–1.012) | 1.008 (1.004–1.013) | 1.008 (1.001–1.015)[a] |
| (Malignant) cancer | All ages (years) | 1.004 (1.004–1.005) | 1.004 (1.004–1.005) | 1.004 (1.003–1.005) |
| | <50 | 1.006 (1.005–1.007) | 1.006 (1.005–1.007) | 1.003 (1.001–1.005)[a] |
| | 50–59 | 1.004 (1.003–1.005) | 1.004 (1.003–1.005) | 1.003 (1.001–1.005)[a] |
| | 60–69 | 1.002 (1.001–1.003) | 1.002 (1.001–1.003)[a] | 1.004 (1.002–1.006) |
| | 70+ | 1.002 (1.000–1.003)[b] | 1.000 (0.998–1.002)[b] | 1.006 (1.003–1.010)[a] |

**Notes:** *Adjusted for age and sex. **Adjusted for age. Hazard ratios are for a 1-year increase in the difference between calculated BA and CA. All hazard ratios are $p<0.001$ except,[a,b] where [a] $p<0.05$ and [b] $p>0.05$.
**Abbreviations:** BA, biological age; CA, chronological age.

and mortality.[25] Also, BA, which can objectively evaluate one's health and aging status, is thought to be a prediction marker, which can predict death, disease, and disability.[24] The best way to predict one's health and aging status is by presenting the aging rate through long-term repetitive BA measurement. But, due to limited data, it was difficult to show the aging rate in our study.[26] Thus, BA difference compared with people of the same age, which reflects relative aging status from birth to the present time, is frequently used instead. This study showed mortality and major age-related disease incidence in Koreans according to BA difference. This can provide the basis to be utilized in various health care fields as an objective evaluation index for assessing health and aging status and as a management index for healthy lifestyle modification in Koreans.

After the concept of BA was first stated in 1935, allostatic load or cumulative risk scores were shown by quantifying and digitizing clinical parameters in similar ways.[27–32] But these measures are used with categorical variables converted from continuous variables when included in a specific range. Thus there is limitation of not being able to use the advantages of the continuous variables.[28] In other words, the clinical biomarkers related to the conventionally allostatic load or cumulative risk scores are affected by the high risk category, and so progressive decrease in physiologic function might not be reflected. But BA based on continuous measures can reflect the progressive decrease in physiologic function with aging.[33,34]

The strengths of our study can be described as follows. First, the developed BA model based on the NHIS – NSC database can be used as a health index that can comprehensively evaluate the health and aging status of the Korean population. Thus, the mortality and major age-related disease incidence results may be generalized for application and interpretation. Second, the hazard ratios of mortality and major age-related disease incidence for BA difference may be presented as basis of the validity of BA model. Third, the developed BA model may be utilized as a prediction marker for mortality and major age-related disease. Finally, BA as a total index of health and aging state is presented as a continuous concept of "age" so it may be easily used for comparison and evaluation, and also be used in various health care fields as a management index that anyone can easily understand.

Despite the strengths of our study mentioned previously, several limitations have to be considered. First, this study was not planned for the purpose of BA model development and utility evaluation. Second, the data used in the analysis

were composed of limited parameters for the purpose of prevention and management of common chronic diseases, so the limitations in the selection of biomarkers and utilization have to be considered. Third, despite the many important age-related diseases, results for only 5 limited diseases were presented. Fourth, important lifestyle habits, such as physical activity status that may have a strong influence on individual health condition and aging status, were not accounted for in the suggested BA model development and utility evaluation. Finally, the limitation of reliance on a single-point measurement should be considered. It is difficult to evaluate intra- or inter-individual difference of health and aging status with a single-point BA measurement. Therefore, further studies on the aging rates, which can be estimated from repeated measurements of BA within a given time period, are considered necessary to better evaluate health condition and aging status.

## Conclusion

The purpose of this study was to demonstrate the validity of estimated BA in relation to all-cause mortality and age-related disease incidence based on NHIS – NSC database. Estimated BA can be expected to be used as an index for assessing health and aging status, and predicting mortality and major age-related disease incidence. Also, it may be used as a health index for overall health and aging state that anyone can easily communicate with in various health care fields.

## Acknowledgments

## Disclosure

The authors report no conflicts of interest in this work.

## References

1. Ingram DK, Nakamura E, Smucny D, Roth GS, Lane MA. Strategy for identifying biomarkers of aging in long-lived species. *Exp Gerontol*. 2001;36(7):1025–1034.
2. Finkel D, Whitfield K, McGue M. Genetic and environmental influences on functional age: a twin study. *J Gerontol B Psychol Sci Soc Sci*. 1995;50(2):P104–P113.
3. Borkan GA, Norris AH. Assessment of biological age using a profile of physical parameters. *J Gerontol*. 1980;35(2):177–184.
4. Jackson SH, Weale MR, Weale RA. Biological age-what is it and can it be measured? *Arch Gerontol Geriatr*. 2003;36(2):103–115.

5. Vickers AJ, Basch E, Kattan MW. Against diagnosis. *Ann Intern Med*. 2008;149(3):200–203.

6. Adams ST, Leveson SH. Clinical prediction rules. *BMJ*. 2012;344: d8312.

7. Kang YG, Suh E, Chun H, Kim SH, Kim DK, Bae CY. Models for estimating the metabolic syndrome biological age as the new index for evaluation and management of metabolic syndrome. *Clin Interv Aging*. 2017;12:253–261.

8. Bae CY, Kang YG, Suh YS, Han JH, Kim SS, Shim KW. A model for estimating body shape biological age based on clinical parameters associated with body composition. *Clin Interv Aging*. 2013;8:11–18.

9. Bae CY, Kang YG, Piao MH, et al. Models for estimating the biological age of five organs using clinical biomarkers that are commonly measured in clinical practice settings. *Maturitas*. 2013;75(3):253–260.

10. Bae CY, Kang YG, Kim S, et al. Development of models for predicting biological age (BA) with physical, biochemical, and hormonal parameters. *Arch Gerontol Geriatr*. 2008;47(2):253–265.

11. Uttley M, Crawford MH. Efficacy of a composite biological age score to predict ten-year survival among Kansas and Nebraska Mennonites. *Hum Biol*. 1994;66(1):121–144.

12. Brant L, Fozard J, Metter E, Balin A. *Age Differences in Biological Markers of Mortality*. Boca Raton FL: CRC Press; 1994:257–270.

13. Mitnitski AB, Graham JE, Mogilner AJ, Rockwood K. Frailty, fitness and late-life mortality in relation to chronological and biological age. *BMC Geriatr*. 2002;2(1):1.

14. Karasik D, Pavlovsky O, Batsevich V, Livshits G, Kobyliansky E. Use of the hand bones roentgenographs in the prediction of age in nine human populations. *Anthropol Anz*. 2000;58(2):199–214.

15. Bulpitt CJ, Antikainen RL, Markowe HL, Shipley MJ. Mortality according to a prior assessment of biological age. *Curr Aging Sci*. 2009; 2(3):193–199.

16. Kimura M, Mizuta C, Yamada Y, Okayama Y, Nakamura E. Constructing an index of physical fitness age for Japanese elderly based on 7-year longitudinal data: sex differences in estimated physical fitness age. *Age (Dordr)*. 2012;34(1):203–214.

17. Bai X, Han L, Liu Q, et al. Evaluation of biological aging process–a population-based study of healthy people in China. *Gerontology*. 2009; 56(2):129–140.

18. Lee J, Lee JS, Park SH, Shin SA, Kim K. Cohort profile: the national health insurance service–national sample cohort (NHIS-NSC), South Korea. *Int J Epidemiol*. 2017;46(2):e15.

19. Lee WC, Lee SY. National health screening program of Korea. *J Korean Med Assoc*. 2010;53(5):363–370.

20. Dubina TL, Mints AYA, Zhuk EV. Biological age and its estimation. III. Introduction of a correction to the multiple regression model of biological age and assessment of biological age in cross-sectional and longitudinal studies. *Exp Gerontol*. 1984;19(2):133–143.

21. Levine ME, Crimmins EM. A comparison of methods for assessing mortality risk. *Am J Hum Biol*. 2014;26(6):768–776.

22. Levine ME. Modeling the rate of senescence: can estimated biological age predict mortality more accurately than chronological age? *J Gerontol A Biol Sci Med Sci*. 2013;68(6):667–674.

23. Lin SY, Lee WJ, Chou MY, Peng LN, Chiou ST, Chen LK. Frailty Index predicts all-cause mortality for middle-aged and older Taiwanese: implications for Active-Aging Programs. *PLoS One*. 2016;11(8): e0161456.

24. Karasik D, Demissie S, Cupples LA, Kiel DP. Disentangling the genetic determinants of human aging: biological age as an alternative to the use of survival measures. *J Gerontol A Biol Sci Med Sci*. 2005; 60(5):574–587.

25. Yin D, Chen K. The essential mechanisms of aging: irreparable damage accumulation of biochemical side-reactions. *Exp Gerontol*. 2005; 40(6):455–465.

26. Dean W, Anacker PC, Kaufman RC, Weber HU. *Biological Aging Measurement: Clinical Applications*. Center for Bio-Gerontology; LA, California; 1988.

27. Karlamangla AS, Singer BH, McEwen BS, Rowe JW, Seeman TE. Allostatic load as a predictor of functional decline: MacArthur studies of successful aging. *J Clin Epidemiol*. 2002;55(7):696–710.

28. Seeman TE, McEwen BS, Rowe JW, Singer BH. Allostatic load as a marker of cumulative biological risk: MacArthur studies of successful aging. *Proc Natl Acad Sci U S A*. 2001;98(8):4770–4775.

29. McEwen BS, Stellar E. Stress and the individual: mechanisms leading to disease. *Arch Intern Med*. 1993;153(18):2093–2101.

30. Karlamangla AS, Merkin SS, Crimmins EM, Seeman TE. Socioeconomic and ethnic disparities in cardiovascular risk in the United States, 2001–2006. *Ann Epidemiol*. 2010;20(8):617–628.

31. Merkin SS, Basurto-Dávila R, Karlamangla A, et al. Neighborhoods and cumulative biological risk profiles by race/ethnicity in a national sample of US adults: NHANES III. *Ann Epidemiol*. 2009;19(3):194–201.

32. Geronimus AT, Hicken M, Keene D, Bound J. "Weathering" and age patterns of allostatic load scores among blacks and whites in the United States. *Am J Public Health*. 2006;96(5):826–833.

33. Post Hospers G, Smulders Y, Maier AB, Deeg D, Muller M. Relation between blood pressure and mortality risk in an older population: role of chronological and biological age. *J Intern Med*. 2015;277(4): 488–497.

34. Melton PE, Zlojutro M, Kimminau K, Crawford M. Biological aging and Cox hazard analysis of mortality trends in a Mennonite community from south-central Kansas. *Am J Hum Biol*. 2006;18(3):387–401.