

Hearing age gap as a marker of auditory ageing and its association with mortality: a nationally representative cohort study

Xin Qi

Chinese People's Liberation Army General Hospital

Shuo Liu

Chinese People's Liberation Army General Hospital

Qiuchen Zheng

Hainan 301 Hospital

Beicheng Li

Hainan 301 Hospital

Pu Dai

Chinese People's Liberation Army General Hospital

Yu Su

suyu029@qq.com

Hainan 301 Hospital

Research Article

Keywords: ageing, hearing loss, machine learning, mortality, pure-tone audiometry, hearing age gap

Posted Date: October 20th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10712456/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Hearing age gap as a marker of auditory ageing and its association with mortality: a nationally representative cohort study

Xin Qi^{1,2,3,4}, Shuo Liu^{3,4}, Qiuchen Zheng^{1,2}, Beicheng Li^{1,2}, Pu Dai^{3,4*}, Yu Su^{1,2,3,4*}

1 *Department of Otolaryngology Head and Neck Surgery, Hainan Hospital of the Chinese People's Liberation Army General Hospital, Sanya 572013, China*

2 *Hainan Clinical Research Center for Otolaryngology Head and Neck Diseases, Sanya 572013, China*

3 *Department of Otolaryngology Head and Neck Surgery, Chinese People's Liberation Army General Hospital, Beijing 100853, China*

4 *National Clinical Research Center for Otolaryngologic Diseases, Beijing 100853, China*

Corresponding author: Pu Dai, MD, PhD

E-mail: Daipu301@vip.sina.com

Address: Department of Otolaryngology Head and Neck Surgery, Chinese People's Liberation Army General Hospital, Beijing 100853, China

Corresponding author: Yu Su, MD, PhD

E-mail: suyu029@qq.com

Address: Department of Otolaryngology Head and Neck Surgery, Hainan Hospital of the Chinese People's Liberation Army General Hospital, Sanya 572013, China

Abstract

Background: The audiogram changes systematically with age, yet clinical and epidemiological practice summarises hearing by a pure-tone average and has not used its frequency-specific pattern for prognosis. We derived a hearing age from the audiogram using machine learning. We then examined whether its gap from chronological age predicts mortality.

Methods: We trained a gradient boosting model to predict chronological age from 14 pure-tone thresholds in 13,331 adults from the National Health and Nutrition Examination Survey (2001 to 2018) with linked mortality. The hearing age gap was defined as hearing age minus chronological age. Survey-weighted Cox models estimated its association with all-cause mortality, and the frozen model was applied without

recalibration to an independent hospital-based cohort of 2,386 adults in southern China.

Results: The model estimated chronological age with a mean absolute error of 8.5 years. The gap was nearly uncorrelated with the pure-tone average, and each additional year of gap was associated with higher mortality after full adjustment (hazard ratio 1.024, 95% CI 1.018 to 1.030), with a dose-dependent increase across gap quartiles. A penalised Cox model retained the gap at every penalty strength while discarding the pure-tone average. The gap improved risk ranking but not absolute risk prediction and was weakly correlated with a blood-based measure of biological ageing. In the external cohort, the model transported well (mean absolute error 8.3 years, correlation 0.67), and each additional year of gap was associated with higher mortality (hazard ratio 1.022, 95% CI 1.006 to 1.038), although the association was attenuated after mutual adjustment with the pure-tone average.

Conclusion: The hearing age gap captures the configuration of the audiogram rather than its severity, behaves as a marker of auditory ageing, and is associated with mortality independently of conventional audiometric measures.

Keywords: ageing; hearing loss; machine learning; mortality; pure-tone audiometry; hearing age gap

Introduction

Hearing loss is the most common sensory deficit of adult life. More than 1.5 billion people lived with some degree of hearing loss in 2021, and the World Health Organization projects 2.5 billion affected by 2050 [1]. Age-related hearing loss develops gradually from mid-life onwards and affects both peripheral and central auditory function [2]. In the United States, roughly one in eight adults has measurable bilateral hearing loss on pure-tone audiometry, and prevalence rises steeply with age [3].

The consequences of hearing loss extend well beyond communication. Longitudinal studies have linked hearing loss to incident dementia [4], and the Lancet Commission on dementia prevention lists hearing loss as the largest potentially modifiable risk factor in mid-life [5]. A meta-analysis of 36 cohort studies with more than six million participants found that hearing loss is associated with a 21% higher risk of all-cause mortality, with stronger associations for audiometrically measured loss than for self-reported loss [6].

Regular hearing aid use is in turn associated with lower mortality in adults with hearing loss [7]. Speech-in-noise impairment shows a similar pattern for incident dementia in the UK Biobank [8].

Despite this evidence, clinical practice reduces the audiogram to summary measures such as the pure-tone average. These measures capture the degree of loss but ignore its configuration and, more importantly, they cannot separate the hearing expected at a given age from the hearing a person actually has. Geroscience has addressed an analogous problem through the concept of biological age. Statistical and machine learning models can estimate how old a body system appears, and the difference between that estimate and chronological age carries prognostic information [9]. Phenotypic age acceleration derived from routine blood tests predicts morbidity and mortality in NHANES [10]. A retinal age gap derived from fundus photographs predicts all-cause mortality in the UK Biobank [11], and a brain age gap derived from neuroimaging shows the same pattern [12]. In the auditory domain, Allen and Eddins proposed an average auditory age computed by inverting published age-norm equations [13]. That measure was formula-based, was never validated against health outcomes, and has not been taken up in epidemiology. Machine learning has been applied to audiometric data, but only in the reverse direction, to predict hearing thresholds from demographic and clinical factors [14]. Whether the audiogram itself can be used to estimate how old the auditory system appears, and whether the resulting age gap predicts death, is unknown.

We therefore trained a machine learning model to estimate chronological age from routine pure-tone audiograms in a nationally representative sample of US adults. We defined the hearing age gap as the difference between this estimate and chronological age, and we examined its association with all-cause and cause-specific mortality over up to 19 years of follow-up. We asked whether the configuration of the audiogram, and not only its average, tracks mortality, and we quantified what the gap adds, and what it does not add, beyond the conventional pure-tone average.

Methods

Data sources and study population

We used data from eight two-year cycles of the National Health and Nutrition Examination Survey (NHANES, 2001 to 2018). NHANES is a repeated cross-sectional survey of the civilian non-institutionalised US population that uses a stratified multistage probability design [15]. Audiometry was administered in a mobile examination centre to eligible participants. We excluded the 2013 to 2014 cycle because it did not include adult audiometry. Participants aged 20 to 79 years at examination were eligible for the present analysis. We excluded participants without a complete bilateral audiogram at 0.5, 1, 2, 3, 4, 6 and 8 kHz, participants ineligible for mortality linkage, and participants with zero follow-up time (**Figure 1**). Mortality status was ascertained through linkage to the National Death Index with follow-up to 31 December 2019 [16]. NHANES protocols were approved by the National Center for Health Statistics ethics review board, and all participants provided written informed consent. The present analysis used publicly available de-identified data and was exempt from additional review.

We pooled data from the seven NHANES cycles conducted between 2001 and 2018 that included audiometry. The number of participants enrolled in each cycle was 11,039 in 2001-2002, 10,122 in 2003-2004, 10,348 in 2005-2006, 10,537 in 2009-2010, 9,756 in 2011-2012, 9,971 in 2015-2016 and 9,254 in 2017-2018, for a total of 71,027 interviewed participants. Audiometry was offered to the full adult examination sample in 2001-2004 and 2011-2018 but only to age-restricted subsamples in 2005-2006 and 2009-2010, which accounts for the smaller contribution of these two cycles.

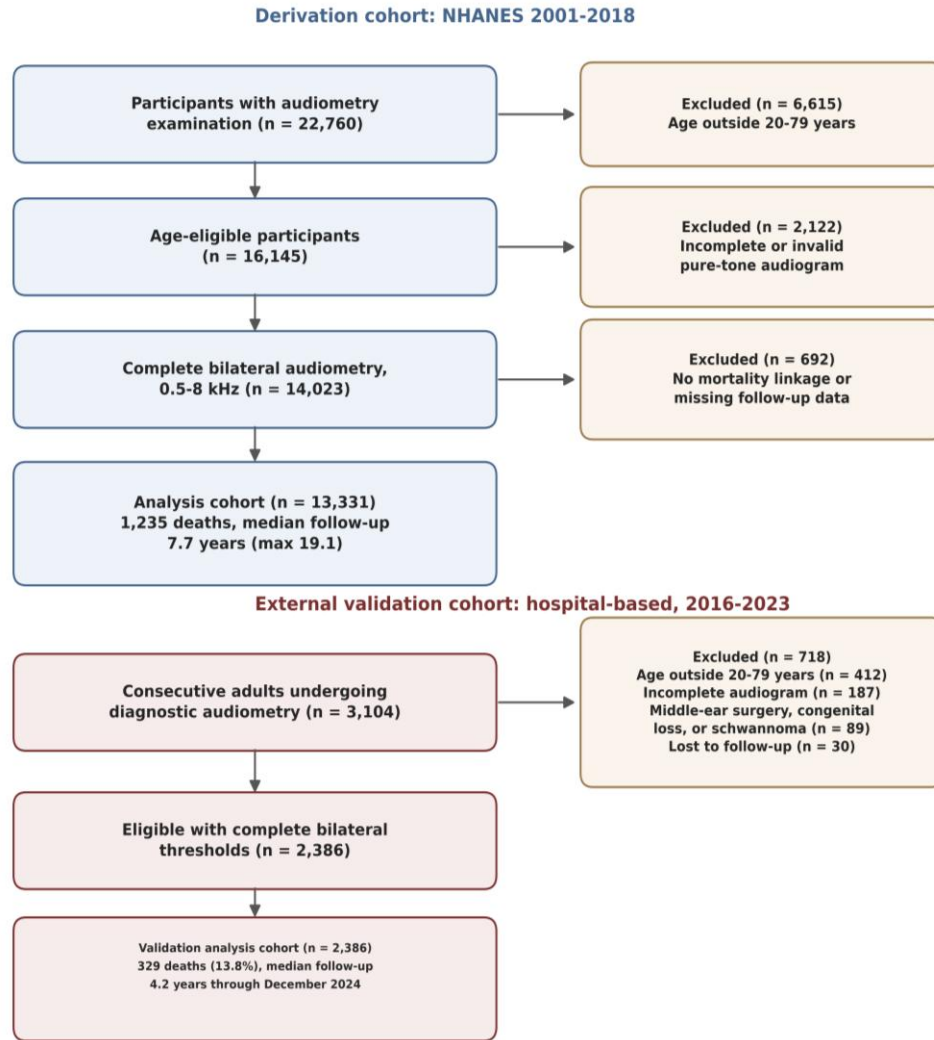


Figure 1. Flow diagram of participant selection. (A) Derivation cohort: National Health and Nutrition Examination Survey (NHANES) 2001-2018 with mortality linkage through 2019. (B) External validation cohort: consecutive adults undergoing diagnostic audiometry at Hainan Hospital of Chinese PLA General Hospital between 2016 and 2023, followed for all-cause mortality through December 2024.

Pure-tone audiometry

Air-conduction thresholds were measured by trained examiners using an automated protocol in a sound-treated room. Thresholds at 0.5, 1, 2, 3, 4, 6 and 8 kHz in each ear were used as model inputs, yielding 14 threshold features per participant. Hearing aids were removed before testing. Conventional hearing status was summarised as the speech-frequency pure-tone average (PTA4), the mean threshold at 0.5, 1, 2 and 4 kHz in the better-hearing ear, in accordance with standard practice.

Derivation of hearing age and the hearing age gap

Gradient boosting was selected after a five-fold cross-validated comparison with linear regression, LASSO, random forest and XGBoost using the same 14 thresholds as inputs. The three ensemble learners (gradient boosting, random forest and XGBoost) performed similarly (MAE 8.5-8.7 years), and all clearly outperformed the two linear models (MAE 9.8 years), consistent with a substantially non-linear association between hearing thresholds and age (**Table S1**). We therefore trained a gradient-boosted regression model (scikit-learn) to predict chronological age from these threshold features. For the final model, predictions were generated with a separate five-fold cross-validation so that every participant received an out-of-fold estimate, which prevents overfitting and information leakage. Model performance was summarised by the mean absolute error and the Pearson correlation between predicted and chronological age (**Figure 2**). Hearing age was defined as the out-of-fold predicted age. The hearing age gap was defined as hearing age minus chronological age, in years. A positive gap indicates that the auditory system appears older than expected, and a negative gap indicates that it appears younger. This formulation follows the age-gap paradigm established for phenotypic, retinal and brain age [9-12]. The concept extends the formula-based average auditory age of Allen and Eddins [13] by replacing fixed age-norm equations with a data-driven model and by validating the gap against hard outcomes.

Mortality outcomes

The primary outcome was all-cause mortality. Secondary outcomes were cardiovascular mortality (underlying cause coded as diseases of the heart) and cancer mortality (malignant neoplasms) from the linked mortality file. Follow-up time was measured from the examination date to death or to 31 December 2019.

Statistical analysis

Cox proportional hazards models estimated hazard ratios for the hearing age gap per one-year increase. Model 1 adjusted for chronological age and sex. Model 2 additionally adjusted for race and ethnicity, education, smoking status, body mass index, self-reported diabetes and self-reported hypertension. A more extensive Model 3 further adjusted for

lifetime noise exposure, ototoxic medication use, cardiovascular medication use and systemic inflammation. Lifetime noise exposure was coded positive when a participant reported occupational exposure to loud noise (OCQ340 or OCQ420 in 2001-2004, AUQ330 in 2011-2012 and 2017-2018; AUQ331 in 2015-2016), recreational exposure to loud noise or music (AUQ230 in 2001-2004; AUQ370 in 2011-2018), or firearm noise exposure (AUQ210 in 2001-2004; AUQ300 in 2011-2018). These items were not administered to adults aged 20-69 years in the 2005-2010 cycles, so noise exposure was unavailable for those participants. Prescription medication data identified loop diuretics, aminoglycosides and platinum agents as ototoxic medications, and antihypertensive, lipid-lowering, antiplatelet and anticoagulant agents as cardiovascular medications. Systemic inflammation was summarised as log₁₀-transformed C-reactive protein, which was measured in all cycles except 2011-2012. All models incorporated NHANES examination weights, scaled for the pooled cycles, with robust standard errors clustered on primary sampling units. To examine whether the gap carried information beyond conventional audiometry, we fitted models with the gap and PTA4 entered together, and a parallel model with a high-frequency pure-tone average at 6 and 8 kHz. The gap was also analysed in quartiles, with the lowest quartile as reference (Figure 3). Sensitivity analyses excluded deaths within the first two years of follow-up to address reverse causation. Discrimination was summarised by the concordance index for nested models. Incremental prognostic value over a model containing age, sex and PTA4 was quantified at 10 years by the continuous net reclassification improvement and the integrated discrimination improvement, with inverse probability of censoring weighting for incomplete follow-up and bootstrap confidence intervals [17]. Analyses used Python 3.12 with scikit-learn [18] and lifelines [19]. Two-sided p values below 0.05 were considered statistically significant. To place the hearing age gap in the context of multisystem biological ageing, phenotypic age was computed from nine blood-based biomarkers and chronological age using the formulation of Levine and colleagues [10]: albumin, creatinine, fasting glucose, log-transformed C-reactive protein, lymphocyte percentage, mean corpuscular volume, red cell distribution width, alkaline phosphatase and white blood cell count. Phenotypic age acceleration was defined as the residual of phenotypic age regressed on chronological age. Fasting glucose was measured only in the morning

fasting subsample, and C-reactive protein was not measured in the 2011-2012 cycle. This analysis therefore included 4,271 participants with complete biomarker data. The association between the hearing age gap and phenotypic age acceleration was summarised by the Pearson correlation coefficient, computed with both measures residualised on chronological age. As a pre-specified sensitivity analysis, we also performed data-driven feature selection from the full candidate pool. An elastic-net Cox proportional hazards model (L1 ratio 0.9) was fitted to 31 standardised candidate variables spanning demographic, lifestyle and clinical characteristics, medication use, PTA4, the hearing age gap and all 14 frequency-specific thresholds, across a grid of penalty strengths. Variables retained at a given penalty were refitted in an unpenalised survey-weighted Cox model to obtain unbiased hazard ratios. Noise exposure, C-reactive protein and the other Model 3 covariates were missing by design in specific survey cycles rather than through individual non-response, so complete-case analysis under this structural missingness is not expected to bias the estimated associations.

External validation cohort

We externally validated the hearing age model in the Hainan Hospital of Chinese PLA General Hospital-based cohort of 2,386 adults who underwent diagnostic pure-tone audiometry at the Department of Otolaryngology between January 2016 and December 2023 and were followed for all-cause mortality through December 2024. Eligibility criteria mirrored the primary analysis: age 20 to 79 years, complete bilateral air-conduction thresholds at 0.5 to 8 kHz, and no history of middle-ear surgery, congenital hearing loss, or vestibular schwannoma. The frozen NHANES model was applied to the 14 thresholds without recalibration or refitting, and the hearing age gap was computed as in the primary analysis. Cox models used the same covariate sets, except that race was omitted because the cohort was ethnically homogeneous. Unweighted models were used because the cohort is a consecutive clinical sample without sampling weights.

Results

The analysis included 13,331 adults with complete audiometry and mortality linkage (**Figure 1**). Mean age at examination was 47.2 years (standard deviation 16.6), and

49.0% were men (**Table 1**). The weighted prevalence of hearing loss (PTA4 above 25 dB HL) was 10.6%. During a median follow-up of 7.7 years (maximum 19.1), 1,235 participants died.

The hearing age model estimated chronological age with a mean absolute error of 8.5 years and a correlation of 0.77 between predicted and chronological age (**Figure 2, Figure S1**). The hearing age gap had a mean of 0.0 years and a standard deviation of 10.7 years (**Figure S2**), indicating wide inter-individual variation in auditory ageing relative to chronological age. About half of the participants (49.9%) had ears that appeared older than their age.

Table 1. Characteristics of the study population (n = 13,331)

Characteristic	Value
Age at examination, years, mean (SD)	47.2 (16.6)
Men, n (%)	6,531 (49.0)
Race and ethnicity, n (%)	
Mexican American	2,120 (15.9)
Other Hispanic	1,198 (9.0)
Non-Hispanic White	5,342 (40.1)
Non-Hispanic Black	3,054 (22.9)
Non-Hispanic Asian	1,122 (8.4)
Other or multiracial	495 (3.7)
PTA4 in the better ear, dB HL, median (IQR)	10.0 (5.0 to 17.5)
Hearing loss (PTA4 > 25 dB HL), n (%)	1,800 (13.5)
Hearing age gap, years, mean (SD)	0.0 (10.7)
Follow-up, years, median (range)	7.7 (0.1 to 19.1)
Deaths during follow-up, n (%)	1,235 (9.3)

PTA4, pure-tone average at 0.5, 1, 2 and 4 kHz; IQR, interquartile range; SD, standard deviation.

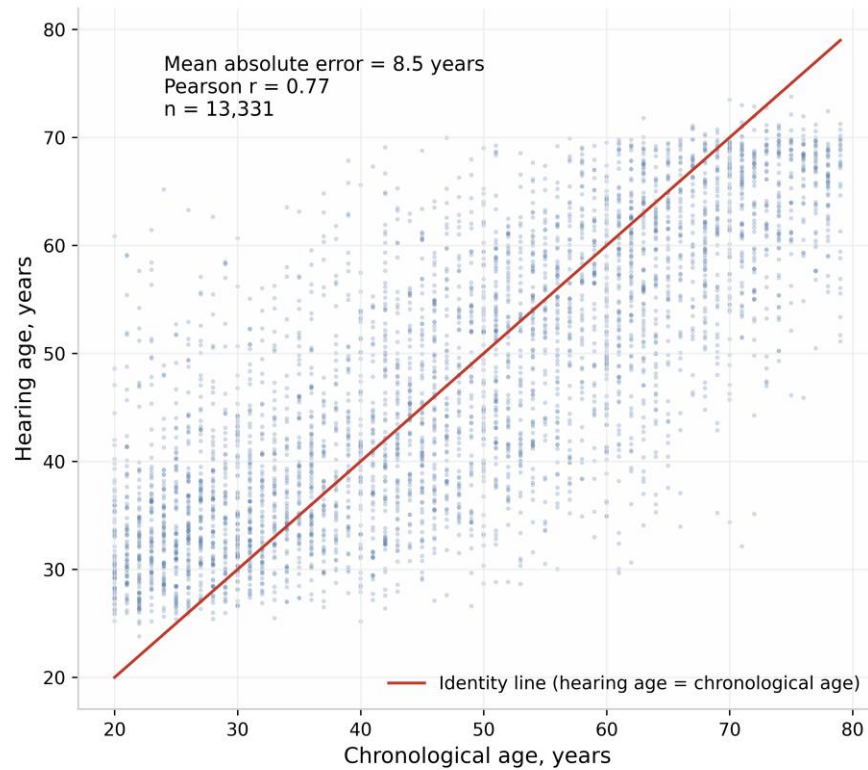


Figure 2. Hearing Age Versus Chronological Age in 13,331 Adults. Each point is one participant; 4,000 randomly sampled points are shown. The red line marks identity.

Model interpretation using Shapley additive explanations (SHAP) [20] showed that high-frequency thresholds drove the estimated hearing age. The 8 kHz thresholds in the two ears were the largest contributors, followed by the 4 kHz thresholds, while low-frequency thresholds contributed little (**Figure 3**). This pattern mirrors the high-frequency predominance of age-related hearing loss and supports the physiological plausibility of the model.

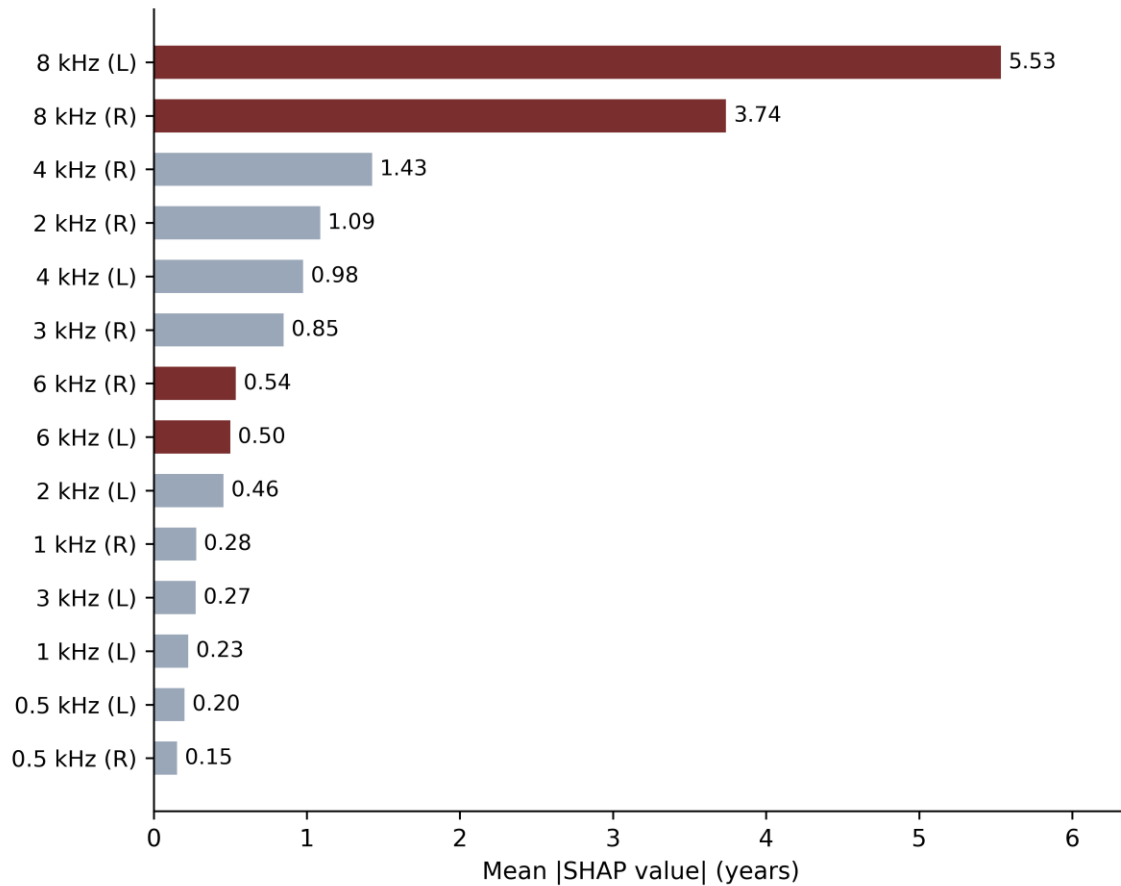


Figure 3. Feature Contributions to the Hearing Age Estimate. Ranked by mean absolute SHAP value in a random sample of 3,000 participants. High-frequency thresholds (8 and 6 kHz) are shown in red.

Each additional year of hearing age gap was associated with higher all-cause mortality. The conventional PTA4 was likewise associated with mortality (per decibel, **Table 2**). The gap was essentially uncorrelated with PTA4 (Pearson $r = -0.01$, Figure S3), which indicates that it captures the configuration of the audiogram rather than its overall level. When both measures were entered together, the gap remained associated with mortality (**Table 2**). Mortality rose monotonically across gap quartiles: compared with the youngest hearing age quartile, Model 2 hazard ratios were 1.19, 1.54 and 1.81 for quartiles 2 to 4 (**Figure 4, Figure S4, Table 2**). In the fully extended Model 3, which also adjusted for lifetime noise exposure, ototoxic medication use, cardiovascular medication use and C-reactive protein, the estimate was unchanged (**Table 2, Figure S5**; 7,902 participants with complete data). The high-frequency pure-tone average at 6 and 8 kHz was likewise associated with mortality (**Table 2**), and adjustment for it attenuated

the gap only modestly (**Figure S5**), consistent with the SHAP finding that high-frequency thresholds carry most of the age signal (**Figure 3, Figure S6**). Participants included in Model 3 were somewhat younger than those excluded because of cycle-dependent item administration, had slightly lower PTA4 and experienced lower mortality (**Table S2**).

In the data-driven sensitivity analysis, the hearing age gap was selected by the elastic-net Cox model at every penalty strength examined from the full pool of 31 candidate variables (**Figure S7**). PTA4 was eliminated from a penalty strength of 1.2 upwards and the right-ear 8 kHz threshold from 3.0 upwards, whereas the gap was retained at every penalty examined, including the strongest (8.0), at which 24 of the 31 variables still had non-zero coefficients (**Table S3**). In the unpenalised refit, with all 14 frequency-specific thresholds and PTA4 simultaneously in the model, each additional year of gap remained associated with higher mortality (HR 1.019, 95% CI 1.003 to 1.035, $p = 0.018$), whereas PTA4 was null in the same model (HR per SD 0.987, 95% CI 0.882 to 1.104, $p = 0.82$) (**Table S4**). The gap was likewise selected in an extended 40-variable pool that additionally included routine laboratory measures ($n = 3,747$). When the algorithm was free to choose among all available variables, it therefore retained the gap, which corroborates the pre-specified modelling strategy.

For cause-specific mortality, the association with cardiovascular death was directionally consistent but imprecise, with a confidence interval that included no association, and the association with cancer death was weak and uncertain. The cancer analysis was based on only 64 deaths and is therefore exploratory. Excluding deaths within the first two years of follow-up left the primary estimate unchanged, and the estimate was also unchanged when deaths within the first five years were excluded (HR 1.023, 95% CI 1.019 to 1.028; **Table 2**). The concordance index for 10-year mortality was 0.792 for age and sex alone, 0.796 with the addition of either the gap or PTA4, and 0.797 with both (**Figure S8**). Incremental value over age, sex and PTA4 was therefore limited for absolute risk prediction. The continuous net reclassification improvement was 0.247 (95% CI 0.15 to 0.35; **Figure S9**), which shows that the gap meaningfully reorders individual risk ranking, but the integrated discrimination improvement was 0.0006 (95% CI -0.0005 to 0.003) and reclassification across a 5% ten-year mortality threshold was

negligible (categorical net reclassification improvement 0.003). The integrated discrimination improvement is known to be conservative when the base model already discriminates well, and values of this magnitude are typical when established risk markers are added to models containing age. The continuous net reclassification improvement, which is independent of the event rate, captures the improvement in risk ordering more directly. The gap thus refines how individuals are ranked by risk more than it changes their predicted absolute risk.

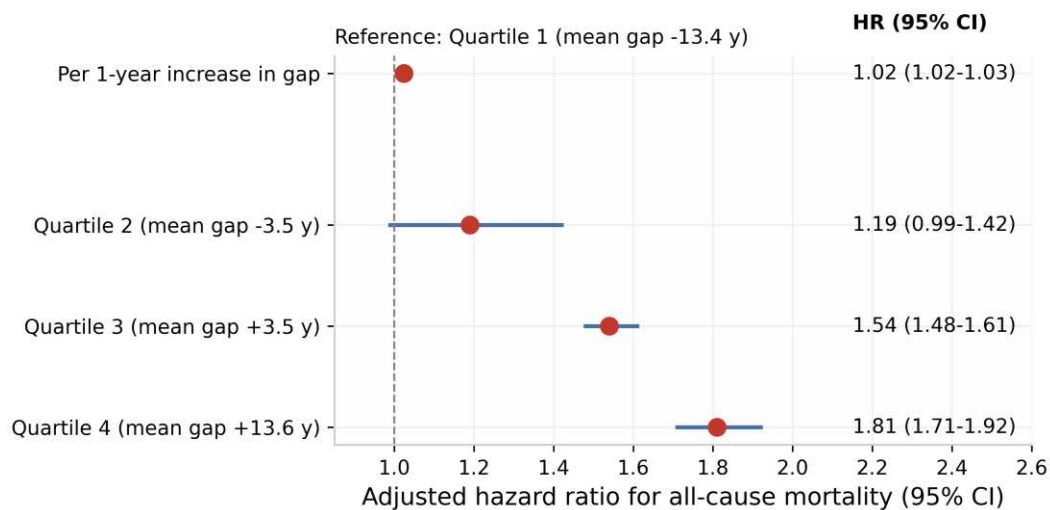


Figure 4. Mortality by Hearing Age Gap Quartile. Adjusted hazard ratios for all-cause mortality by hearing age gap. Model 2 adjustment (age, sex, race and ethnicity, education, smoking, body mass index, diabetes, hypertension). Quartile 1 (youngest-appearing ears) is the reference category.

Table 2. Association between hearing age gap and all-cause mortality

Analysis	Hazard ratio (95% CI)	p value
Hearing age gap (per +1 year)		
Model 1	1.032 (1.026-1.038)	<0.001
Model 2	1.024 (1.018-1.030)	<0.001
Model 3	1.020 (1.010-1.030)	<0.001

Comparator measures		
PTA4 (per +1 dB), Model 2	1.013 (1.007-1.019)	<0.001
High-frequency PTA, 6-8 kHz (per +10 dB), Model 2	1.09 (1.06-1.13)	<0.001
Hearing age gap (per +1 year), adjusted for PTA4	1.016 (1.013-1.019)	<0.001
Hearing age gap quartiles (Model 2)		
Quartile 2 vs Quartile 1	1.19 (0.99-1.42)	0.06
Quartile 3 vs Quartile 1	1.54 (1.48-1.61)	<0.001
Quartile 4 vs Quartile 1	1.81 (1.71-1.92)	<0.001
Secondary and sensitivity analyses		
Cardiovascular mortality, gap per +1 year	1.025 (0.993-1.058)	0.13
Cancer mortality, gap per +1 year	1.044 (1.001-1.090)	0.045
Gap per +1 year, excluding deaths within 2 years	1.024 (1.018-1.030)	<0.001
Gap per +1 year, excluding deaths within 5 years	1.023 (1.019-1.028)	<0.001

Model 1 adjusts for age and sex. Model 2 additionally adjusts for race and ethnicity, education, smoking, body mass index, diabetes and hypertension. Model 3 further adjusts for lifetime noise exposure, ototoxic medication use, cardiovascular medication use and log10 C-reactive protein in participants with complete data (n = 7,902).

A restricted cubic spline model showed that mortality rose smoothly across the range of the gap, with no evidence of a threshold or of departure from linearity (**Figure 5**).

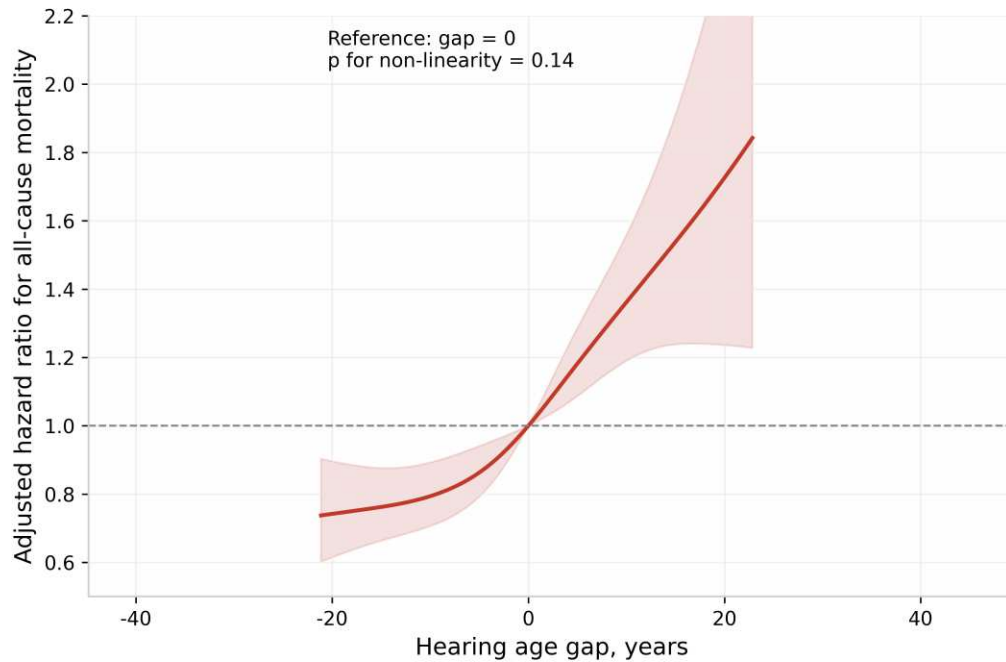


Figure 5. Dose-response Association of the Hearing Age Gap with Mortality. Adjusted hazard ratio for all-cause mortality across the hearing age gap, from a restricted cubic spline model with knots at the 5th, 35th, 65th and 95th percentiles, adjusted for the Model 2 covariates. The shaded band is the 95% confidence interval, the reference is a gap of 0 years, and the histogram shows the distribution of the gap.

In the subsample with complete blood biomarkers ($n = 4,271$), phenotypic age acceleration was positively correlated with the hearing age gap ($r = 0.09$, $p < 0.001$). With both measures residualised on chronological age, the correlation remained significant (**Figure 6**). The two measures share only about 1% of their variance, which indicates that the hearing age gap captures an organ-specific dimension of biological ageing that is largely distinct from blood-based systemic ageing.

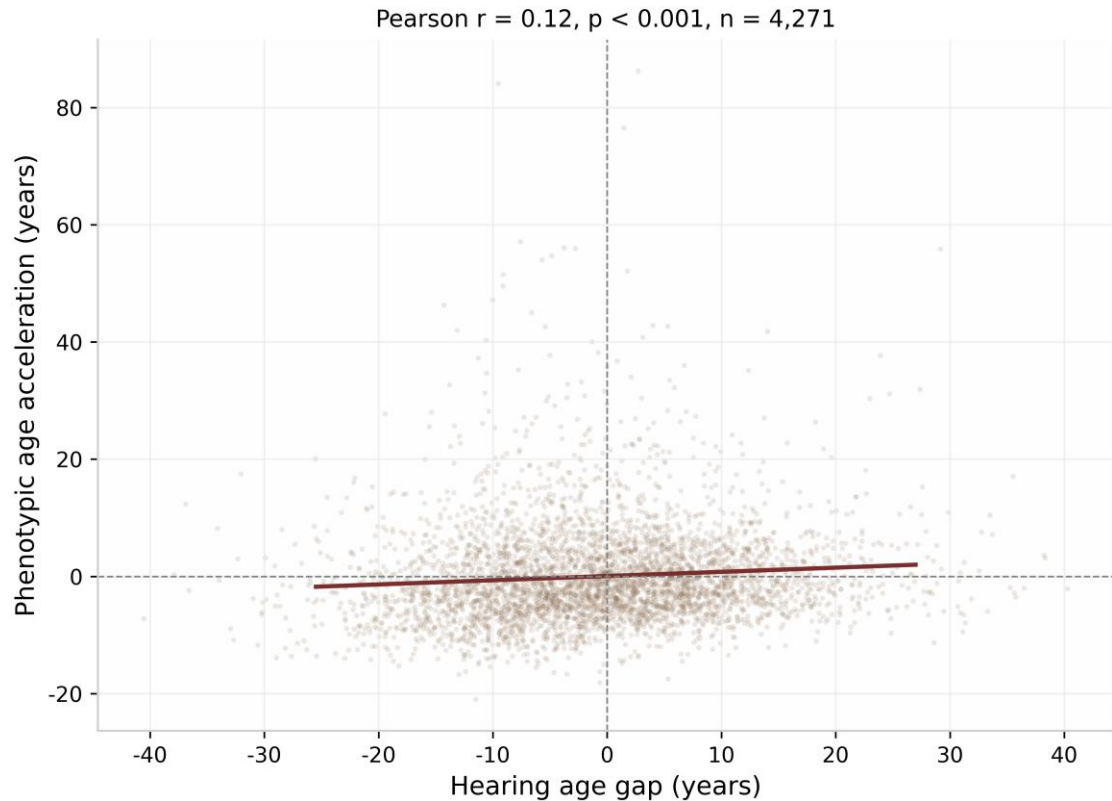


Figure 6. Hearing Age Gap Versus Phenotypic Age Acceleration in 4,271 Adults with Complete Blood Biomarkers. Both measures are residualised on chronological age. The solid line is the least-squares fit.

External validation

The hospital-based validation cohort comprised 2,386 adults (median age 56 years, 51% men) with a median follow-up of 4.2 years, during which 329 participants died (13.8%) (Table 3). Applied without recalibration, the hearing age model estimated chronological age with a mean absolute error of 8.3 years and a correlation of $r = 0.67$, slightly below its internal performance (8.5 years, $r = 0.77$). Calibration showed a modest systematic downward shift (calibration-in-the-large -2.6 years) and compression toward the mean (calibration slope 0.70, 95% CI 0.67 to 0.73), both visible in Figure 7A. The gap distribution was shifted downward relative to NHANES (mean -2.6 years, SD 10.0) and was moderately correlated with PTA4 ($r = 0.33$) (Figure 7A).

Each additional year of hearing age gap was associated with higher all-cause mortality in the minimally adjusted model (HR 1.024, 95% CI 1.009 to 1.040) and in the

fully adjusted model (HR 1.022, 95% CI 1.006 to 1.038). The estimate was essentially unchanged after additional adjustment for noise exposure and ototoxic and cardiovascular medication use (Table 4). Mortality was higher in the upper gap quartiles than in the lowest: adjusted hazard ratios were 1.29, 1.55 and 1.50 for quartiles 2 to 4 (Figure 7B and 7C). After mutual adjustment for PTA4, the gap was attenuated and no longer significant (HR 1.018, 95% CI 0.994 to 1.043), and PTA4 was likewise attenuated (HR per 10 dB 1.04, 95% CI 0.87 to 1.25), consistent with the moderate correlation between the two measures in this cohort (Table 4). This pattern differs from the primary cohort, where the gap remained significant after mutual adjustment. However, the confidence interval in the validation cohort (0.994 to 1.043) includes the primary estimate of 1.016, and the gap-by-PTA4 interaction was null (p for interaction = 0.53), so effect modification by PTA4 does not explain the attenuation. The fully adjusted model including the gap achieved a C-index of 0.70 in the validation cohort (Table 4).

Table 3. Characteristics of the external validation cohort (n = 2,386)

Characteristic	Value
Age at examination, years, median (IQR)	56 (48 to 66)
Men, n (%)	1,221 (51.2)
Educational level, %	
Less than high school	16.3
High school graduate	22.2
Some college	28.5
College graduate	22.3
Postgraduate	10.7
Body mass index, kg/m ² , mean (SD)	24.4 (3.4)
Ever smoked, %	39.0
Diabetes, %	14.5
Hypertension, %	33.2
Lifetime noise exposure, %	23.1
Ototoxic medication use, %	2.9
Cardiovascular medication use, %	42.4
C-reactive protein, mg/dl, median (IQR)	1.4 (1.0 to 2.1)
PTA4 in the better ear, dB HL, median (IQR)	16.2 (9.4 to 24.4)
Hearing loss (PTA4 > 25 dB HL), %	23.6
Hearing age gap, years, mean (SD)	-2.6 (10.0)
Follow-up, years, median (IQR)	4.2 (2.5 to 6.3)
Deaths during follow-up, n (%)	329 (13.8)

IQR, interquartile range; PTA4, pure-tone average at 0.5, 1, 2 and 4 kHz (better ear); SD, standard deviation. Values are unweighted.

Table 4. External validation: association between the hearing age gap and all-cause mortality in the hospital-based cohort (n = 2,386; 329 deaths)

Analysis	HR (95% CI)	P value
Hearing age gap (per +1 year)		
Model 1 (age, sex, education)	1.024 (1.009-1.040)	0.002
Model 2 (+ smoking, BMI, diabetes, hypertension)	1.022 (1.006-1.038)	0.006
Model 3 (Model 2 + noise exposure, ototoxic and cardiovascular medications)	1.022 (1.006-1.039)	0.008
Model 2 + PTA4 (mutual adjustment)	1.018 (0.994-1.043)	0.14
Gap × PTA4 interaction (per year × dB)	1.000 (0.998-1.001)	0.53
Comparator measure		
PTA4 per 10 dB (Model 2)	1.04 (0.87-1.25)	0.66
Hearing age gap quartiles (Model 2)		
Q2 vs Q1	1.29 (0.96-1.75)	0.09
Q3 vs Q1	1.55 (1.12-2.14)	0.008
Q4 vs Q1	1.50 (1.00-2.24)	0.05
Model discrimination (C-index)		
Age + sex only	0.68	
Model 2 covariates only	0.70	
Model 2 + gap	0.70	
Model 2 + PTA4	0.70	
Model 2 + gap + PTA4	0.70	

HR, hazard ratio; PTA4, pure-tone average at 0.5, 1, 2 and 4 kHz (better ear). The hearing age model was trained on NHANES and applied to the validation cohort without recalibration. C-index values were computed on the full validation cohort (n = 2,386, 329 events) for comparability across models.

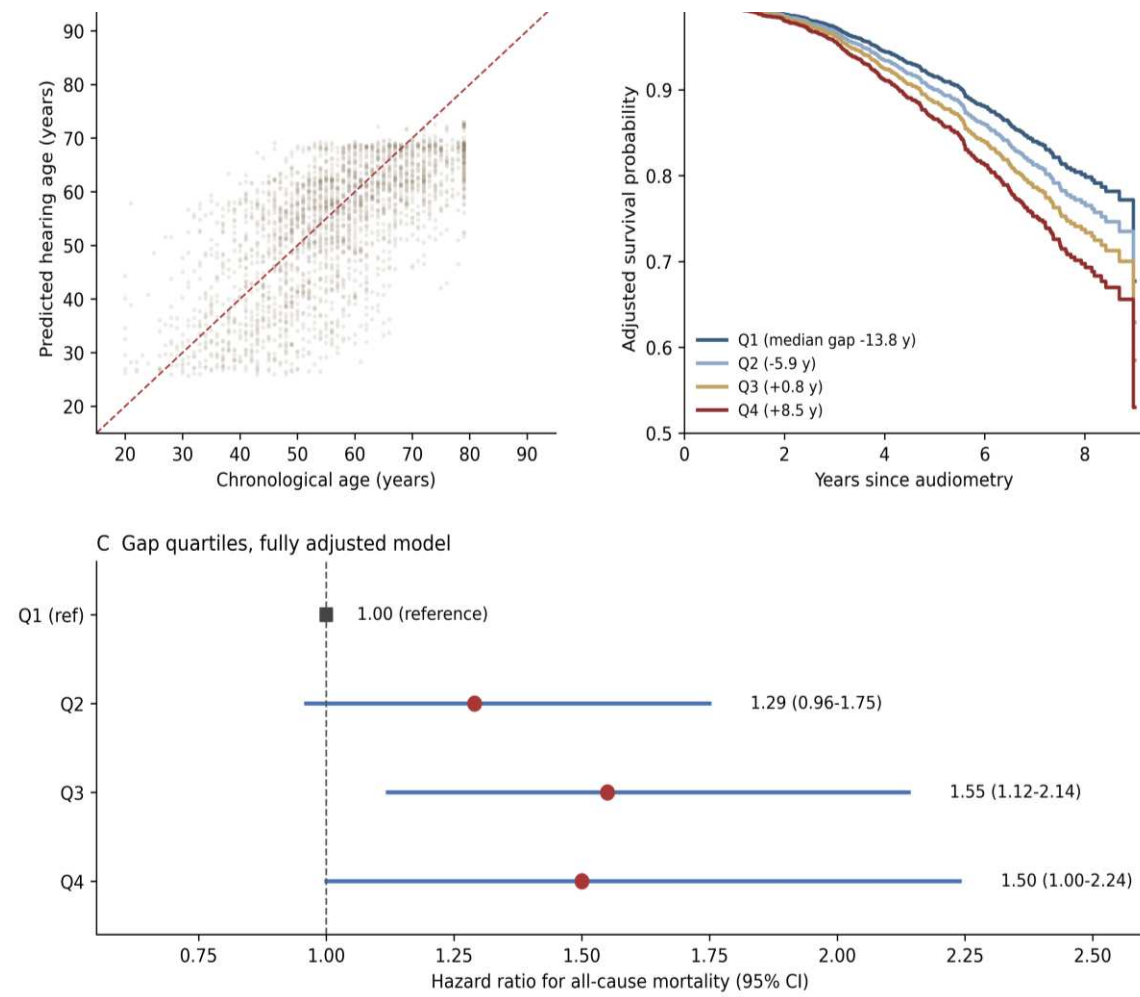


Figure 7. External Validation of the Hearing Age Model in the Hainan Hospital of Chinese PLA General Hospital cohort. (A) Predicted hearing age versus chronological age (MAE = 8.3 y, $r = 0.67$). The dashed line marks identity. (B) Adjusted survival curves by hearing age gap quartile, estimated from a Cox model adjusted for age, sex, education, smoking, body mass index, diabetes and hypertension, with covariates held at cohort means and the gap set to the quartile median. (C) Adjusted hazard ratios for all-cause mortality by gap quartile in the fully adjusted model.

Discussion

In this study we derived a hearing age from routine audiograms and showed that the gap between hearing age and chronological age is associated with mortality. Each additional year of gap carried a 2.4% higher rate of death after full adjustment, and mortality rose monotonically across gap quartiles, from 1.19 in the second quartile to 1.81 in the highest. One may reasonably ask whether the pure-tone average alone would suffice. Our data suggest that it would not, and the reason lies in what each measure sees. The pure-tone average reads the audiogram like an average score: it captures how much

hearing is lost but not how the loss is distributed across frequencies. The gap reads the shape of the audiogram. High-frequency thresholds, especially at 8 kHz, carried most of the age signal in the model. This pattern mirrors the pathology of presbycusis, in which hair cell loss and striaal degeneration begin in the basal turn of the cochlea, the region encoding high frequencies, and spread with age [2, 13, 21]. Because the cochlea does not regenerate, thresholds at these frequencies accumulate a lifetime of noise exposure, ototoxic medication, microvascular injury and genetic susceptibility. Noise alone, the most common preventable cochlear insult, has both auditory and systemic health consequences [22]. Presbycusis itself is heterogeneous: metabolic, sensory and neural components coexist in proportions that no single average can represent [13]. The audiogram therefore behaves like a cumulative exposure record, and the gap is a way of reading it.

The strongest evidence that the gap is not a repackaged average comes from the feature selection analysis. When a penalised Cox model was forced to choose among all audiometric and clinical variables, it retained the gap at every penalty strength while discarding the pure-tone average. When all fourteen thresholds and the pure-tone average were entered together without penalisation, the gap remained associated with mortality whereas the pure-tone average became null. The gap was likewise retained when routine laboratory measures were added to the candidate pool. The two measures were also nearly uncorrelated, so most adults with a normal average can still carry a large gap. In practical terms, the pure-tone average can identify people who already have hearing loss, but it has little to say about the majority of adults whose average is still normal. The gap stratifies exactly that majority. The dose-response curve was smooth, with no threshold below which the gap was harmless.

The gap appears to measure how fast the ear itself has aged. It was only weakly correlated with phenotypic age acceleration, a blood-based measure of systemic biological ageing, and the two shared about 1% of their variance. This near-independence agrees with recent work showing that organs age at different rates within the same person [23, 24]. Blood-based measures capture yet another, more dynamic dimension of the same process [25]. A hearing-based marker may therefore complement, rather than duplicate, systemic ageing clocks. The approach also extends a family of organ age gaps

derived from retinal photographs and brain imaging [11, 12], but it differs in two practical ways: the audiogram is far cheaper and already embedded in routine care, and its biological substrate in the cochlea is comparatively direct. This perspective also suggests a testable next question: whether the gap predicts outcomes of the auditory and nervous systems, such as incident dementia and functional decline, more strongly than it predicts mortality [26, 27].

The gap did not meaningfully improve absolute risk prediction. The concordance index rose by 0.004 and the integrated discrimination improvement was near zero, while the continuous net reclassification improvement was 0.25. These results are compatible: the gap reorders individuals by risk more than it moves them across any absolute risk threshold. Measurement error also deserves mention. The age-prediction error of the model (8.5 years) attenuates the estimated association with mortality, so the reported hazard ratios are likely conservative [28, 29]. Established markers rarely improve discrimination once age is in the model [30], and we therefore present the gap as a marker of auditory ageing rather than as a risk calculator. In practical terms, a clinician who ranked patients by the gap alone would place genuinely higher-risk individuals above lower-risk ones far more often than chance, even though no individual's predicted probability would change much. The same prediction error explains an apparent paradox in the validation cohort: age prediction transported almost unchanged, while discrimination of the survival model fell. The first measures the hearing age model, and the second measures a mortality model applied to a different population.

The external validation adds an important boundary to these claims. The hearing age model itself travelled well: prediction error and correlation in the hospital cohort were slightly below internal performance (mean absolute error 8.3 versus 8.5 years, correlation 0.67 versus 0.77). Calibration was shifted downward by about 2.6 years and the calibration slope was below one, which is expected when a model trained in a general population is applied to a clinic sample [31]. Patients referred for diagnostic audiometry in such a setting are older and have more hearing loss than community-dwelling adults, so the cohort was moderately enriched for hearing pathology by design.

In that cohort, the gap was attenuated and no longer significant after mutual adjustment for the pure-tone average. Notably, the pure-tone average was itself

attenuated there (hazard ratio 1.04 per 10 dB, versus about 1.13 per 10 dB in the primary cohort), so the attenuation reflects the moderate correlation between the two measures in this cohort rather than one clearly absorbing the other. Three observations temper any conclusion that the two cohorts conflict. The confidence interval in the validation cohort included the primary estimate, so the estimates are statistically compatible. The interaction between the gap and the pure-tone average was null, which argues against genuine effect modification by hearing loss severity. The validation analysis rested on 329 deaths from a single clinical cohort, so its precision and generalisability remain limited. A more parsimonious explanation lies in the correlation structure itself. In the primary cohort the two measures were nearly uncorrelated, so each association could be estimated independently. In this clinic sample they were moderately correlated, and mutual adjustment split the signal so that neither remained significant on its own. The independent contribution of the gap is therefore best demonstrated where it is orthogonal to the pure-tone average, which in our data was the community setting. Whether the stronger correlation is a general feature of clinic populations or specific to this cohort should be examined in further validation samples. We applied the frozen model without recalibration because our aim was to test transportability rather than local performance [32]. A simple recalibration of intercept and slope would be expected to restore calibration if the model were deployed locally.

The practical appeal of the gap is its cost. Audiograms are already collected in occupational health screening, otolaryngology clinics and population surveys, and the gap can be computed from stored thresholds at no marginal cost. A markedly elevated gap could prompt review of cardiovascular risk factors, frailty and functional status, and could help identify candidates for hearing rehabilitation. Hearing aid use has been associated with lower mortality in observational data [7], and a large randomised trial has shown that hearing intervention can slow cognitive decline in at-risk older adults [33]. Speech-in-noise difficulty, another auditory measure, has also been linked to dementia risk [8]. Whether treatment modifies the excess risk associated with an elevated gap is unknown, and the gap could serve as an enrichment criterion or intermediate endpoint in future trials. In settings such as China, where population ageing is rapid and access to comprehensive geriatric assessment is uneven, an indicator derived from an existing

audiometric record may be especially useful. Community hearing screening programmes already generate large volumes of audiograms, and linking these records to a gap estimate would create a low-cost risk flag without any additional testing burden.

The study has several strengths. Audiometry in NHANES followed a standardised automated protocol, the sample was large and nationally representative, and mortality follow-up was near complete. The gap was derived from cross-validated predictions, so no participant's hearing age was informed by his or her own outcome, and the modelling pipeline was frozen before it was applied to the external cohort.

This study has limitations. First, hearing aid use was not included in the adjustment sets. Only 254 NHANES participants reported regular hearing aid use, among whom 37 died, so a formal subgroup or interaction analysis was underpowered (p for interaction = 0.72) and is not reported. Hearing aid use is also plausibly a mediator on the pathway from hearing loss to mortality rather than a confounder, in which case adjustment would be inappropriate. Regular hearing aid use has been associated with lower mortality among adults with hearing loss [7], and adequately powered analyses of effect modification by hearing aid use remain an important direction for future work. Second, noise exposure, C-reactive protein and several other extended-model covariates were not administered in specific survey cycles, which reduced the sample available for Model 3. The resulting missingness is structural rather than individual, and baseline differences between included and excluded participants were modest. Third, the cause-specific analyses, particularly cancer mortality (64 events), were exploratory and should be interpreted with caution. Fourth, the external validation cohort was smaller and clinic-based, with a steeper PTA4-mortality gradient than the primary cohort, which likely explains the attenuation of the gap after mutual adjustment in that setting.

Conclusion

The hearing age gap captures the configuration of the audiogram, reflects an organ-specific dimension of biological ageing, and predicts mortality independently of conventional audiometric measures. It is inexpensive to compute and biologically interpretable. Its greatest promise lies in community and primary care settings, where it could convert a routine test into an early signal of systemic risk.

Declarations

Ethics approval and Consent to participate

The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Hainan Hospital of Chinese PLA General Hospital (Approval No.: 20210901-ZDYF2020-142). Informed consent was obtained from all participants involved in the study.

Consent for publication

Not applicable.

Availability of data and materials

The NHANES audiometry, demographic, laboratory, and linked mortality data analysed in this study are publicly available from the National Center for Health Statistics (<https://www.cdc.gov/nchs/nhanes/> and <https://www.cdc.gov/nchs/linked-data/mortality-files/>). The validation cohort data contain clinical information from the Hainan Hospital of Chinese PLA General Hospital and are not publicly available because of institutional privacy and ethics restrictions. De-identified data may be made available from the corresponding author upon reasonable request and with approval from the institutional ethics committee. The analysis code used to derive hearing age and the hearing age gap is available from the corresponding author upon reasonable request.

Competing Interests

The authors declare that they have no competing interests.

Funding: This research was supported by the National Natural Science Foundation of China General Program, (82271185) and Hainan Provincial Key Research and Development Program(ZDYF2020142).

Authors' contributors

XQ, SL contributed equally to this work. Conceptualisation: XQ and PD. Methodology: XQ, SL and YS. Formal analysis: XQ, SL and YS. Investigation: QCZ and BCL. Data curation: XQ and QCZ. Writing – original draft: XQ. Writing – review and editing: PD and YS. Supervision and funding acquisition: PD and YS. All authors have read and agreed to the published version of the manuscript.

Acknowledgements

The authors thank the National Center for Health Statistics of the Centers for Disease Control and Prevention for conducting the National Health and Nutrition Examination Survey and for making the audiometry and linked mortality data publicly available, and thank all participants of NHANES. The authors also thank the staff of the Department of Otolaryngology, Hainan Hospital of Chinese PLA General Hospital, for their support in assembling the validation cohort.

The study question, cohort definitions, exposure and outcome definitions, covariate selection, analyses, interpretation of results, manuscript preparation, and decision to submit were determined independently by the authors.

During the preparation of this work, the authors used Kimi to support language editing and formatting. The authors reviewed and edited all outputs, confirmed the validity of the text, and take full responsibility for the content of the publication.

Clinical trial number

Not applicable

References

1. World Health Organization. World report on hearing. Geneva: World Health Organization; 2021.
2. Gates GA, Mills JH. Presbycusis. *Lancet*. 2005;366(9491):1111-1120.
3. Agrawal Y, Platz EA, Niparko JK. Prevalence of hearing loss and differences by demographic characteristics among US adults: data from the National Health and Nutrition Examination Survey, 1999-2004. *Arch Intern Med*. 2008;168(14):1522-1530.
4. Lin FR, Metter EJ, O'Brien RJ, Resnick SM, Zonderman AB, Ferrucci L. Hearing loss and incident dementia. *Arch Neurol*. 2011;68(2):214-220.
5. Livingston G, Huntley J, Liu KY, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. *Lancet*. 2024;404(10452):572-628.
6. Jia J, Li H, Xu C. Hearing loss and its association with all-cause and cause-specific mortality: a meta-analysis of cohort studies. *PLoS One*. 2025;20(10):e0333125.
7. Choi JS, Adams ME, Crimmins EM, Lin FR, Ailshire JA. Association between hearing aid use and mortality in adults with hearing loss in the USA: a mortality follow-up study of a cross-sectional cohort. *Lancet Healthy Longev*. 2024 Jan;5(1):e66-e75.
8. Stevenson JS, Clifton L, Kuzma E, Littlejohns TJ. Speech-in-noise hearing impairment is associated with an increased risk of incident dementia in 82,039 UK Biobank participants. *Alzheimers Dement*. 2022;18(3):445-456.

9. Klemmera P, Doubal S. A new approach to the concept and computation of biological age. *Mech Ageing Dev.* 2006;127(3):240-248.
10. Liu Z, Kuo PL, Horvath S, Crimmins E, Ferrucci L, Levine M. A new aging measure captures morbidity and mortality risk across diverse subpopulations from NHANES IV: a cohort study. *PLoS Med.* 2018;15(12):e1002718.
11. Zhu Z, Shi D, Guankai P, et al. Retinal age gap as a predictive biomarker for mortality risk. *Br J Ophthalmol.* 2023;107(4):547-554.
12. Cole JH, Ritchie SJ, Bastin ME, et al. Brain age predicts mortality. *Mol Psychiatry.* 2018;23(5):1385-1392.
13. Allen PD, Eddins DA. Presbycusis phenotypes form a heterogeneous continuum when ordered by degree and configuration of hearing loss. *Hear Res.* 2010;264(1-2):10-20
14. Gathman TJ, Choi JS, Vasdev RMS, Schoepfoerster JA, Adams ME. Machine learning prediction of objective hearing loss with demographics, clinical factors, and subjective hearing status. *Otolaryngol Head Neck Surg.* 2023;169(3):504-513.
15. Zipf G, Chiappa M, Porter KS, Ostchega Y, Lewis BG, Dostal J. National health and nutrition examination survey: plan and operations, 1999-2010. *Vital Health Stat 1.* 2013;(56):1-37.
16. National Center for Health Statistics. The linkage of National Center for Health Statistics survey data to the National Death Index: 2019 linked mortality file (LMF): linkage methodology and analytic considerations. Hyattsville, MD: National Center for Health Statistics; 2022.
17. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med.* 2008;27(2):157-172.
18. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res.* 2011;12:2825-2830.
19. Davidson-Pilon C. lifelines: survival analysis in Python. *J Open Source Softw.* 2019;4(40):1317.
20. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst.* 2017;30:4765-4774.
21. Schuknecht HF, Gacek MR. Cochlear pathology in presbycusis. *Ann Otol Rhinol Laryngol Suppl.* 1993;102(1 Pt 2):1-16.
22. Basner M, Babisch W, Davis A, et al. Auditory and non-auditory effects of noise on health. *Lancet.* 2014;383(9925):1325-1332.
23. Oh HS-H, Rutledge J, Nachun D, et al. Organ aging signatures in the plasma proteome track health and disease. *Nature.* 2023;624(7990):164-172.
24. Tian YE, Cropley V, Maier AB, Lautenschlager NT, Breakspear M, Zalesky A. Heterogeneous aging across multiple organ systems and prediction of chronic disease and mortality. *Nat Med.* 2023;29(5):1221-1231.
25. Belsky DW, Caspi A, Corcoran DL, et al. DunedinPACE, a DNA methylation biomarker of the pace of aging. *eLife.* 2022;11:e73420.
26. Deal JA, Betz J, Yaffe K, et al. Hearing impairment and incident dementia and cognitive decline in older adults: the Health ABC Study. *J Gerontol A Biol Sci Med Sci.* 2017;72(5):703-709.

27. Lin FR, Ferrucci L, Metter EJ, An Y, Zonderman AB, Resnick SM. Hearing loss and cognition in the Baltimore Longitudinal Study of Aging. *Neuropsychology*. 2011;25(6):763-770.
28. MacMahon S, Peto R, Collins R, et al. Blood pressure, stroke, and coronary heart disease: part 1, prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *Lancet*. 1990;335(8692):765-774.
29. Clarke R, Shipley M, Lewington S, et al. Underestimation of risk associations due to regression dilution in long-term follow-up of prospective studies. *Am J Epidemiol*. 1999;150(4):341-353.
30. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21(1):128-138.
31. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW. Calibration: the Achilles heel of predictive analytics. *BMC Med*. 2019;17(1):230.
32. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med*. 2015;162(1):55-63.
33. Lin FR, Pike JR, Albert MS, et al. Hearing intervention versus health education control to reduce cognitive decline in older adults with hearing loss in the USA (ACHIEVE): a multicentre, randomised controlled trial. *Lancet*. 2023;402(10404):786-797.

Supplementary Tables and Figures

Table S1. Cross-validated comparison of age-prediction algorithms in the NHANES training data.

All models used the same 14 audiometric thresholds as inputs and the same five-fold cross-validation partition (n = 13,331). MAE, mean absolute error across all held-out folds; MAE SD, standard deviation of fold-level MAE; r, Pearson correlation between predicted and chronological age; mean error, mean of predicted minus chronological age. The gradient boosting model used for the final hearing age estimator is shown in bold.

Algorithm	MAE, years	MAE SD across folds	r	R-squared	Mean error, years
Linear regression	9.80	0.13	0.692	0.480	+0.00
LASSO (standardised)	9.80	0.13	0.692	0.479	+0.00
Random forest	8.58	0.19	0.763	0.582	-0.02
Gradient boosting (current)	8.51	0.20	0.766	0.586	-0.01

model)

XGBoost	8.69	0.21	0.754	0.567	-0.01
---------	------	------	-------	-------	-------

Table S2. Characteristics of participants included in versus excluded from the extended adjustment model (Model 3)

Characteristic	Included in Model 3 (n = 7,902)	Excluded from Model 3 (n = 5,429)
Age at examination, years, mean (SD)	45.6 (15.7)	49.4 (17.8)
Men, %	47.7	50.7
Race and ethnicity, %		
Mexican American	19.2	11.2
Other Hispanic	9.2	8.7
Non-Hispanic White	40.1	40.0
Non-Hispanic Black	20.5	26.4
Non-Hispanic Asian	4.0	3.2
Other or multiracial	7.0	10.5
Body mass index, kg/m ² , mean (SD)	29.2 (6.9)	29.0 (6.9)
Ever smoked, %	45.0	45.1
Diabetes, %	11.8	12.9
Hypertension, %	31.1	37.5
PTA4 in the better ear, dB HL, mean (SD)	12.7 (12.6)	14.9 (15.2)
Hearing loss (PTA4 > 25 dB HL), %	11.1	17.1
Hearing age gap, years, mean (SD)	0.5 (10.6)	-0.7 (10.8)
Survey cycle, %		
2001-2002	21.7	2.8
2003-2004	21.0	2.0
2005-2006	0.0	7.8
2009-2010	0.0	11.1
2011-2012	0.0	70.8
2015-2016	51.1	4.8
2017-2018	6.2	0.8
Follow-up, years, median (IQR)	4.7 (3.7 to 16.3)	8.0 (7.4 to 8.8)
Deaths during follow-up, n (%)	528 (6.7)	707 (13.0)

Model 3 additionally adjusts for lifetime noise exposure, ototoxic medication use, cardiovascular medication use and log10 C-reactive protein. Excluded participants were concentrated in the 2011-2012 cycle, in which C-reactive protein was not measured, and in the 2005-2010 cycles, in which noise exposure items were not administered to adults aged 20-69 years. The missingness is therefore structural, driven by survey cycle design, rather than individual non-response. Values are unweighted. IQR, interquartile range; SD, standard deviation.

Table S3. Elastic-net Cox feature selection across penalty strengths (n = 12,129; 763 deaths; 31 candidate variables; L1 ratio 0.9)

Penalty	Variables selected,	Hearing age	PTA4	8 kHz (right	8 kHz (left ear)
---------	---------------------	-------------	------	--------------	------------------

strength	n (of 31)	gap		ear)	
0.5	31	Yes	Yes	Yes	Yes
0.8	31	Yes	Yes	Yes	Yes
1.2	30	Yes	No	Yes	Yes
2.0	28	Yes	No	Yes	Yes
3.0	27	Yes	No	No	Yes
5.0	27	Yes	No	No	Yes
8.0	24	Yes	No	No	Yes

A variable is counted as selected when its penalised coefficient exceeded $1e-4$. The hearing age gap was retained at every penalty strength examined. PTA4 was eliminated from a penalty of 1.2 upwards, and the right-ear 8 kHz threshold from 3.0 upwards. The gap was likewise retained in an extended 40-variable pool that additionally included routine laboratory measures (n = 3,747; 269 deaths).

Table S4. Unpenalised survey-weighted Cox refit including all 31 candidate variables simultaneously (n = 12,129; 763 deaths)

Variable	HR per SD (95% CI)	P value
Age	2.524 (1.699 to 3.749)	<0.001
Male sex	1.129 (1.042 to 1.223)	0.003
Education	0.806 (0.709 to 0.917)	0.001
Ever smoked	1.344 (1.266 to 1.427)	<0.001
Body mass index	0.944 (0.857 to 1.040)	0.24
Diabetes	1.268 (1.219 to 1.318)	<0.001
Hypertension	1.238 (1.190 to 1.289)	<0.001
Lifetime noise exposure	1.059 (0.993 to 1.130)	0.08
Ototoxic medication use	1.079 (1.058 to 1.100)	<0.001
Cardiovascular medication use	1.013 (0.963 to 1.065)	0.62
PTA4	0.987 (0.882 to 1.104)	0.82
Hearing age gap	1.221 (1.035 to 1.440)	0.018
0.5 kHz (R)	1.154 (1.106 to 1.205)	<0.001
1 kHz (R)	0.917 (0.874 to 0.963)	<0.001
2 kHz (R)	0.940 (0.832 to 1.062)	0.32
3 kHz (R)	1.141 (1.086 to 1.199)	<0.001
4 kHz (R)	0.967 (0.948 to 0.986)	<0.001
6 kHz (R)	0.829 (0.723 to 0.950)	0.007
8 kHz (R)	1.018 (1.011 to 1.025)	<0.001
0.5 kHz (L)	1.068 (0.963 to 1.185)	0.21
1 kHz (L)	1.104 (1.053 to 1.156)	<0.001
2 kHz (L)	0.957 (0.915 to 1.000)	0.050
3 kHz (L)	0.815 (0.770 to 0.862)	<0.001
4 kHz (L)	1.165 (0.888 to 1.529)	0.27
6 kHz (L)	1.097 (0.911 to 1.319)	0.33
8 kHz (L)	1.088 (0.799 to 1.482)	0.59
Other Hispanic	1.101 (1.068 to 1.135)	<0.001
Non-Hispanic White	1.159 (1.121 to 1.198)	<0.001
Non-Hispanic Black	1.279 (1.248 to 1.311)	<0.001
Non-Hispanic Asian	1.016 (0.982 to 1.052)	0.37

Other or multiracial	0.887 (0.785 to 1.002)	0.05
----------------------	------------------------	------

All variables were standardised; hazard ratios are per 1 standard deviation. For the hearing age gap (SD 10.7 years), the estimate corresponds to HR 1.019 (95% CI 1.003 to 1.035, $p = 0.018$) per 1-year greater gap. Race and ethnicity categories are relative to Mexican American. Models used scaled examination weights with robust standard errors clustered on primary sampling units.

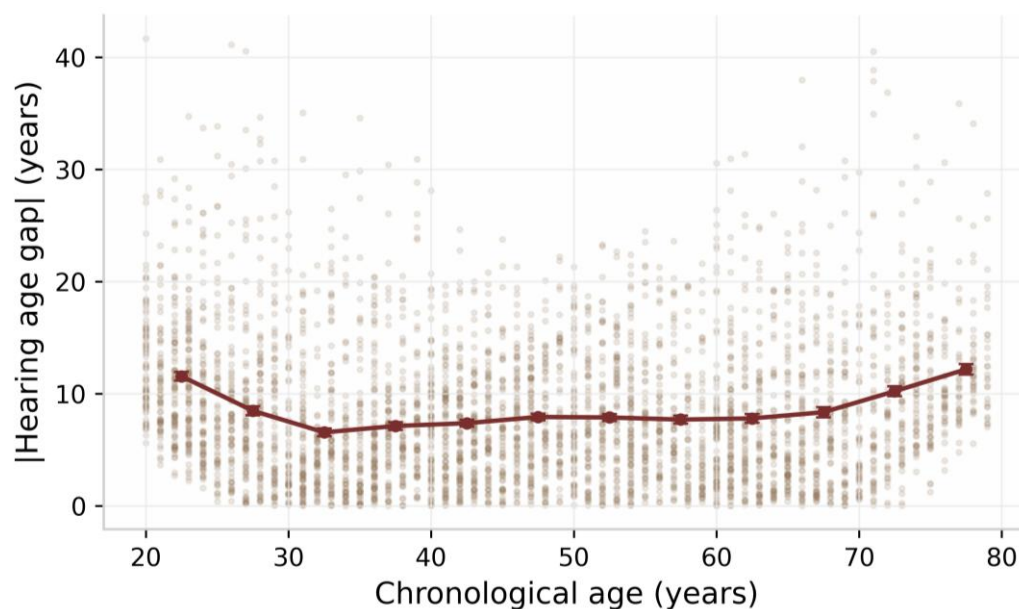


Figure S1. Absolute error of the hearing age model across chronological age. Points show 4,000 randomly sampled participants; the red line shows the mean absolute error with 95% confidence intervals in 5-year age bins. Prediction error is stable across adulthood, with a modest increase at the oldest ages.

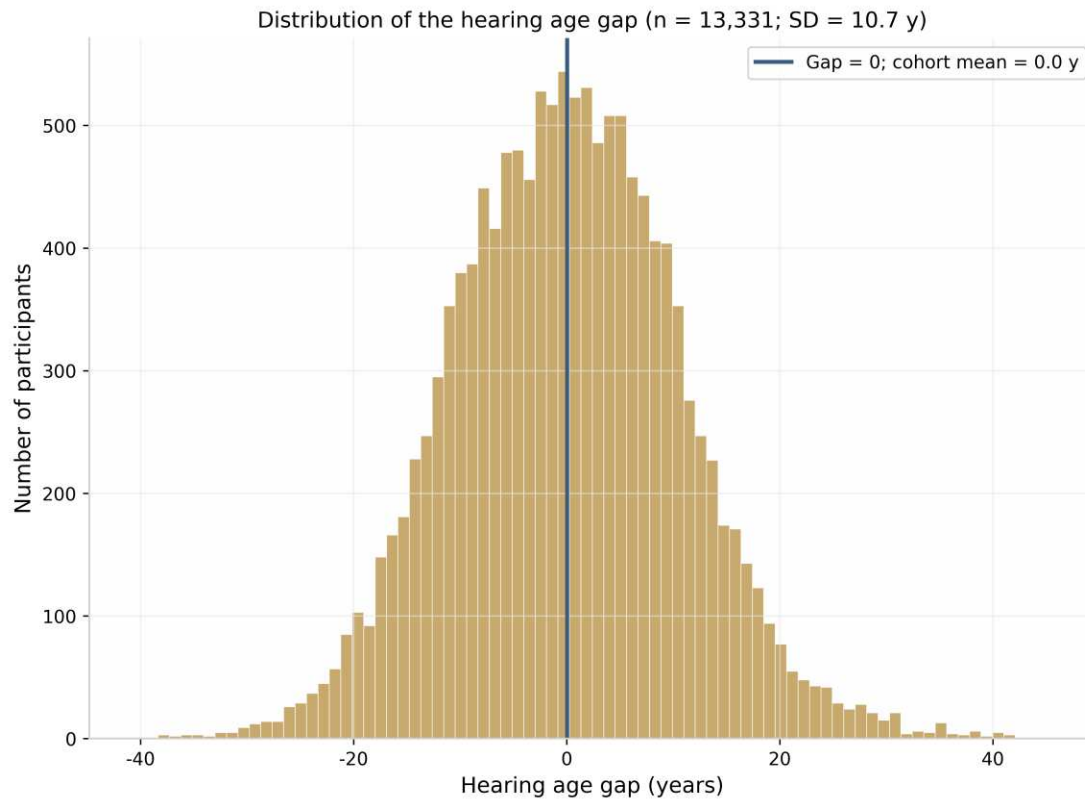


Figure S2. Distribution of the hearing age gap in the analytic cohort. The dashed red line marks a gap of zero (hearing age equal to chronological age); the blue line marks the cohort mean.

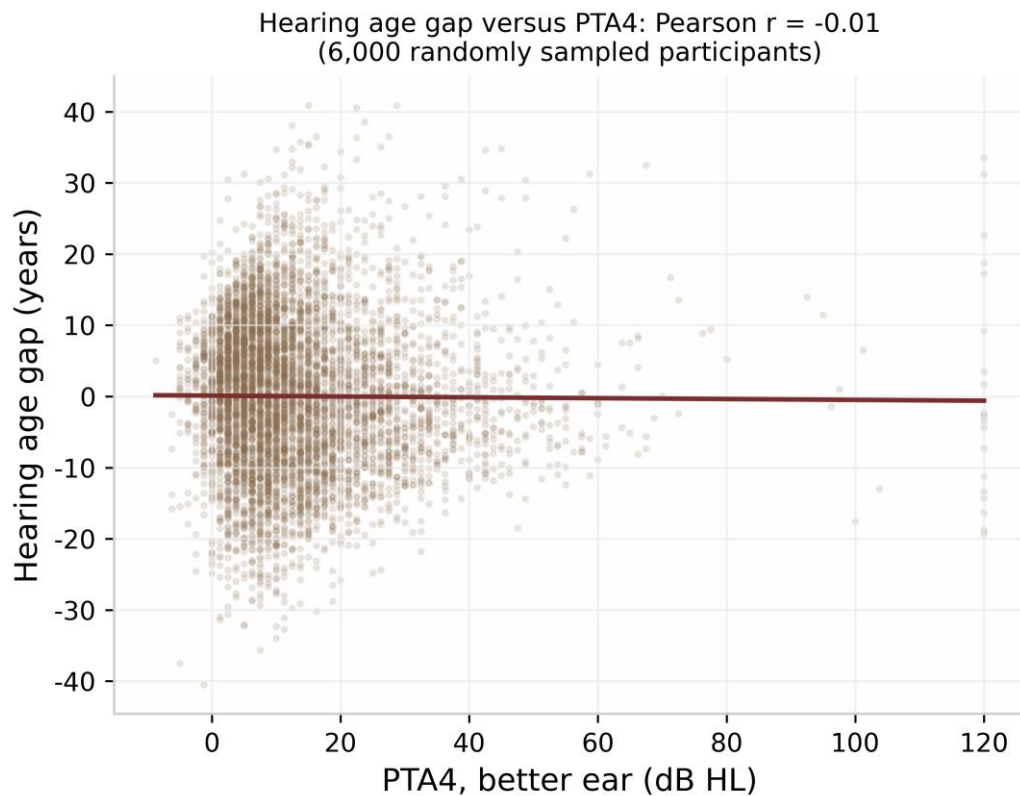


Figure S3. Hearing age gap versus the four-frequency pure-tone average (PTA4) of the better ear. The two measures were essentially uncorrelated (Pearson $r = -0.01$), indicating that the gap captures the configuration of the audiogram rather than its overall level.

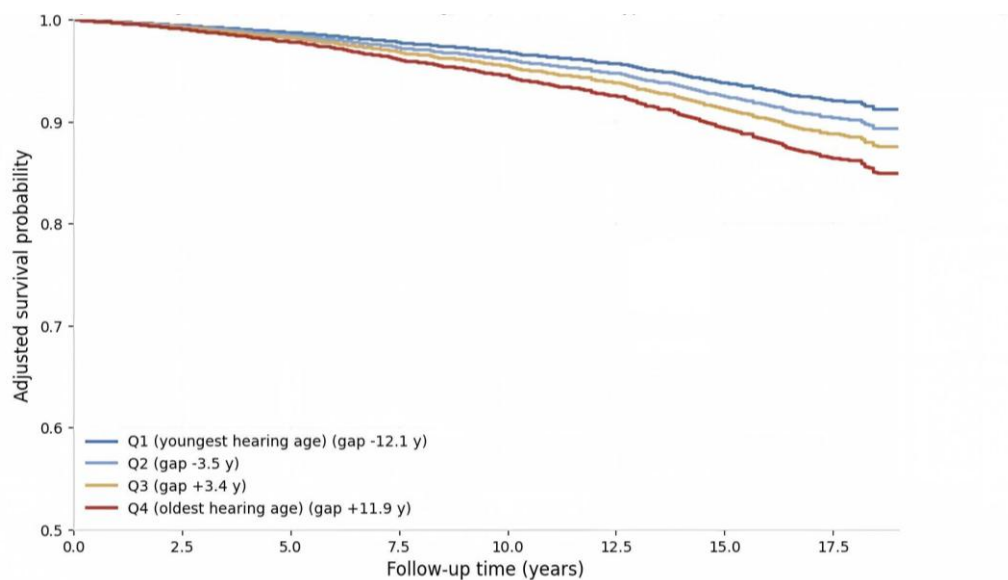


Figure S4. Adjusted survival curves by hearing age gap quartile, estimated from a survey-weighted Cox model adjusted for age, sex, race, education, smoking, body mass index, diabetes and hypertension, with covariates held at cohort means and the gap set to the quartile median.

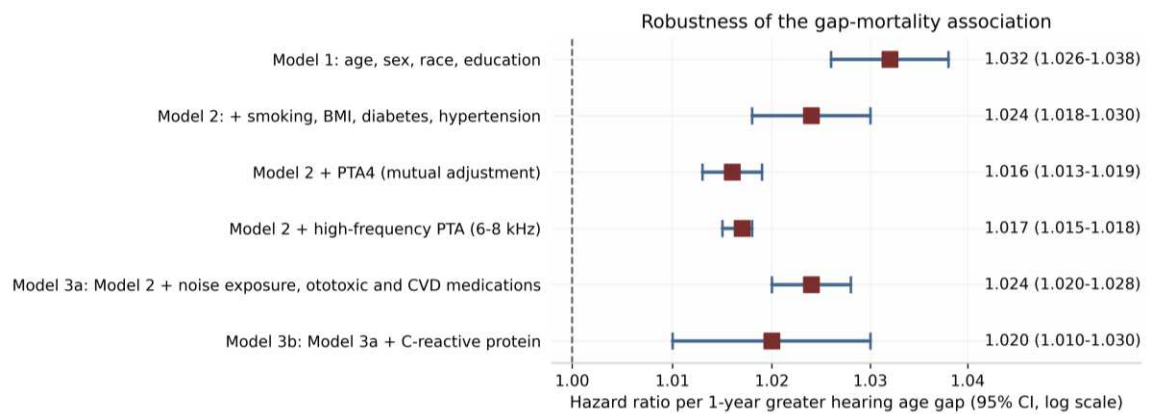


Figure S5. Hazard ratios for all-cause mortality per 1-year greater hearing age gap across progressively richer adjustment sets, including mutual adjustment for PTA4 and high-frequency PTA. The association remained statistically significant in all specifications.

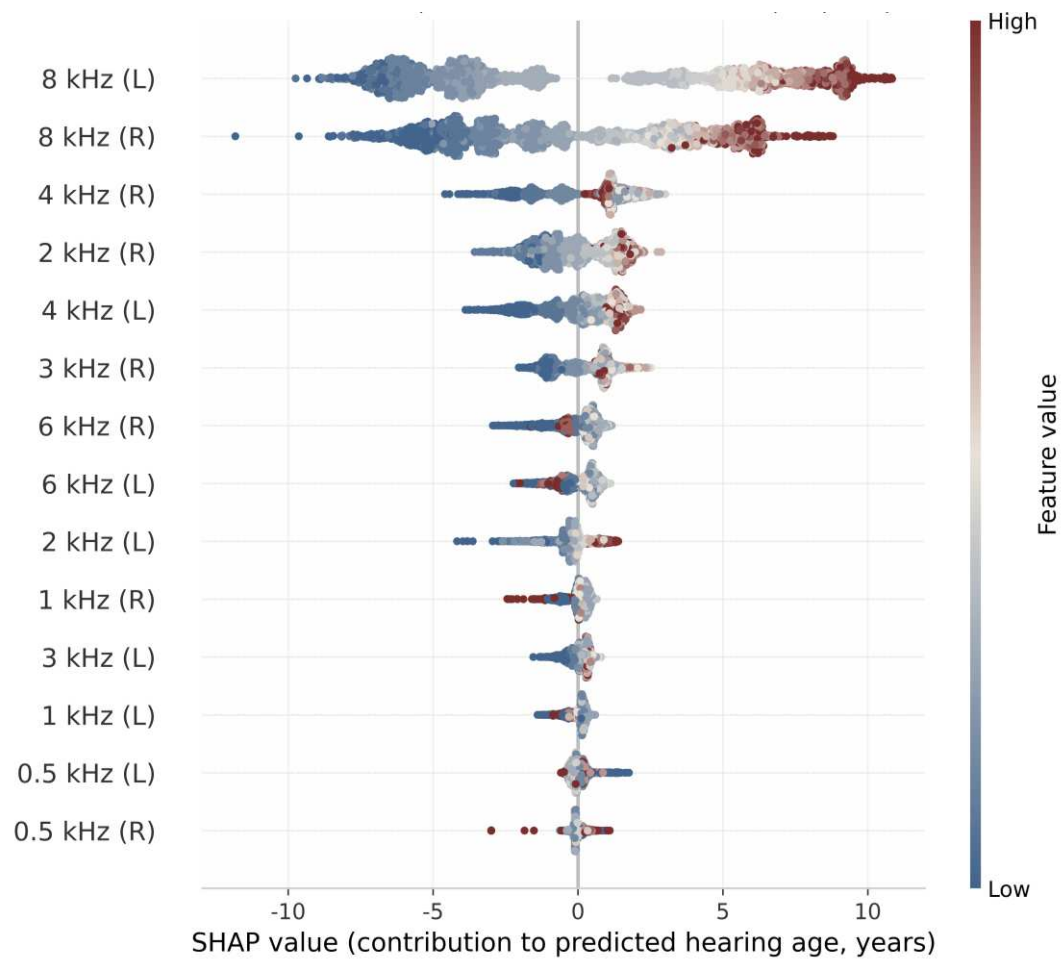


Figure S6. SHAP value distribution for each audiometric feature in the hearing age model, computed on 3,000 randomly sampled participants. Each point is one participant; color indicates the threshold value (red = poorer hearing). Extended high-frequency thresholds (8 kHz) contributed the largest share of the predicted hearing age.

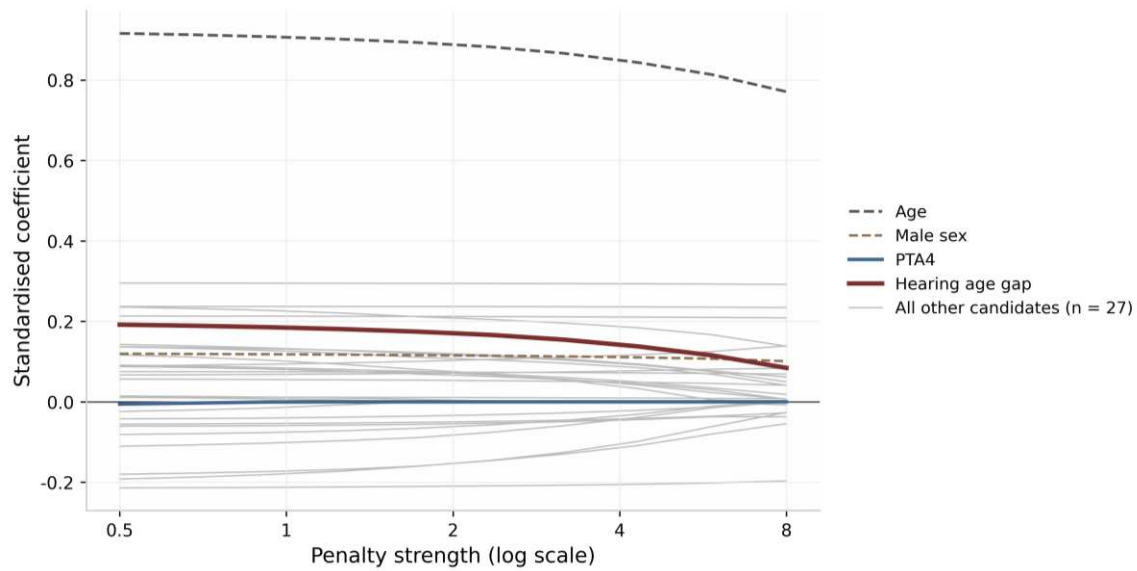


Figure S7. Elastic-net Cox feature selection from the full pool of 31 candidate variables ($n = 12,129$; 763 deaths; L1 ratio 0.9). Each line traces the standardised coefficient of one candidate variable across increasing penalty strengths. The hearing age gap (red) was retained at every penalty examined, whereas PTA4 (blue) and the individual 8 kHz thresholds were eliminated at stronger penalties. Age (dashed grey) and male sex (dashed brown) are shown for reference; all other candidates are shown in light grey.

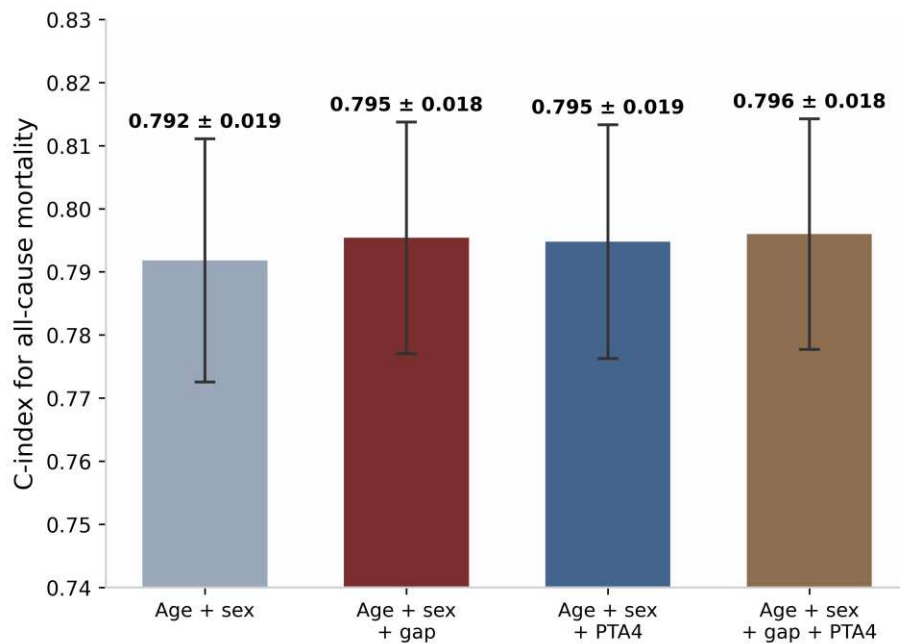


Figure S8. C-index of survey-weighted Cox models for all-cause mortality. Adding the hearing age gap to age and sex improved discrimination comparably to adding PTA4; the model containing both achieved the highest C-index (0.797).

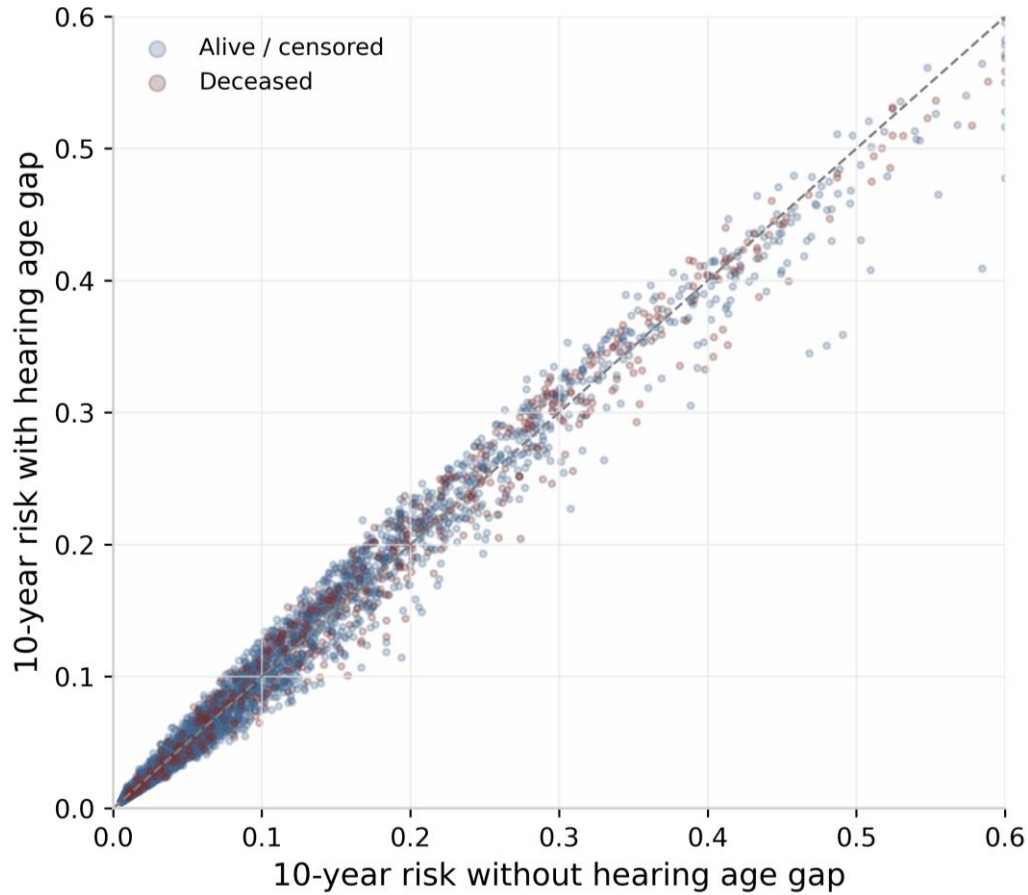


Figure S9. Reclassification of predicted 10-year mortality risk when the hearing age gap is added to a model containing age, sex, and PTA4. Each point is one of 5,000 randomly sampled participants; deceased participants (red) lie predominantly above the identity line, indicating appropriate upward reclassification (continuous NRI = 0.247).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTablesandFigures.docx](#)