scientific reports

Check for updates

OPEN

# Prognostic significance of the QuickSIN score for future hearing threshold deterioration

Jiayue Liu[1]✉, Haotian Zou[2], Eleanor M. Simonsick[3], Joshua Stohl[4] & Tobias Overath[1]

**About 10% of audiology patients who experience hearing difficulties in noise have clinically normal hearing thresholds in quiet. While it has been suggested that hearing difficulties in noise might be a precursor for the subsequent development of clinical hearing loss, there is so far no direct evidence that supports this hypothesis. This study aimed to determine whether hearing difficulties in noise, as measured by the Quick Speech in Noise (QuickSIN) test, could be used to identify people at risk of the development and the progression of clinical hearing loss, using a large dataset of 1128 individuals in the Baltimore Longitudinal Study of Aging (BLSA). A linear mixed model analysis revealed that individuals with a poorer QuickSIN score at the preceding audiological assessment were likely to reveal a more substantial deterioration in clinical audiometric thresholds at their subsequent visit than those with a better QuickSIN score at their previous assessment. These findings offer valuable insights for early interventions and monitoring strategies for individuals at risk of hearing loss.**

Hearing loss is one of the most prevalent health conditions in aging. An estimated 55% of the US population in their seventies suffer from clinical hearing loss, with this percentage rising to 80% for people older than 80 years[1]. However, hearing difficulties in noisy environments may occur prior to the development of clinical hearing loss[2,3]. Plomp analyzed data from 700,000 households and estimated that, for someone to experience hearing difficulties in noise, the average hearing loss at 500, 1000 and 2000 Hz is 24 dB HL (hearing level), while that for experiencing hearing difficulties in quiet is 35 dB HL[4]. Additionally, Parthasarathy et al. examined over 100,000 audiology patient records and found that about 10% of patients reported hearing difficulties, even though their audiometric thresholds were clinically normal[5].

Hearing difficulties result from a combination of attenuation and distortion of the input signal[4,6]. Attenuation is often caused by the loss or dysfunction of hair cells[7,8] and causes a decrease in audibility. Distortion has multiple underlying factors, which can be classified into two general categories: (1) auditory peripheral factors and (2) cognitive factors[9]. Auditory peripheral factors include degraded temporal resolution[9–11] and impaired frequency selectivity of the auditory system[5,12] both of which may contribute to hearing difficulties in noise. Cognitive factors involve cognitive demand for auditory processing, attention, working memory, multi-sensory integration, listening effort, and cortical inhibition[5,13–15]. Hearing difficulties in noise are considered an expected perceptual consequence of distortion.

Attenuation leads to worsening of the pure tone audiometry thresholds and can often be treated by sound amplification (e.g., via hearing aids[16]). Distortion, on the other hand, currently does not have a similarly straightforward measure nor clinical treatment. Some speech in noise tests, such as the Quick Speech-in-Noise (QuickSIN)[17] test, word in noise (WIN)[18] and hearing in noise test (HINT)[19] may capture the perceptual impact of distortion. However, these speech in noise tests are not commonly administered in the audiology clinic. Instead, speech perception in quiet measures such as the speech discrimination score (SDS, also referred as word recognition score, WRS)[20] are routinely administered. Recent research has proposed replacing speech perception in quiet tests with speech in noise tests as part of the audiological routine[2] for two main reasons: (1) although speech perception in quiet is commonly used to predict hearing aid outcomes, difficulties understanding speech in noise remain a primary complaint among patients, including those with hearing aids[21]; and (2) speech in noise tests have higher validity than speech in quiet, for example with respect to informing real-world hearing difficulties[22] predicting hearing aid outcomes[23] and detecting vestibular schwannoma[24].

Despite their differences, attenuation and distortion are influenced by similar risk factors, including aging and noise exposure[25–30]. Distortion appears to be more sensitive to these risk factors than attenuation, where

[1]Department of Psychology and Neuroscience, Duke University, Durham, USA. [2]Department of Biostatistics and Bioinformatics, Duke University, Durham, USA. [3]Intramural Research Program , National Institute on Aging, Baltimore, USA. [4]North American Research Laboratory, MED-EL US, Durham, USA. ✉email: liu.jiayue@outlook.com

speech perception in noise is often noticed before an official diagnosis of clinical hearing loss[5,28,29,31–35]. Given this risk profile of distortion, speech perception in noise may serve as an early predictor of the development or the progression of clinical hearing loss. However, to date there exists no direct or conclusive evidence that individuals who experience hearing difficulties in noise (a potential consequence of distortion) are at greater risk of developing clinical hearing loss later in life[36]. Given the prevalence of hearing loss[37] and its potential impact on cognition[38,39], social engagement[40,41] and quality of life[42], early identification of hearing difficulties and prevention of clinical hearing loss could significantly improve quality of life as individuals age. In the present study, we investigated the prognostic ability of the QuickSIN test to predict a subsequent deterioration of audiometric thresholds.

We did so using data from the Baltimore Longitudinal Study of Aging (BLSA), which, initiated in 1958, comprises over 3200 individuals[43]. Presently, more than 1000 individuals remain actively engaged in the study. In the BLSA, pure tone audiometry and QuickSIN[17] are acquired as part of general hearing assessments. Pure tone audiometry is the gold standard for diagnosing hearing loss, with pure tone thresholds determining the degree of clinical hearing loss[44]. QuickSIN assesses speech perception in babble noise and is considered an accurate reflection of real-world hearing abilities, since it better simulates real-life experiences[2,22].

Importantly, in the BLSA these assessments are longitudinally conducted on the same individuals. The BLSA dataset thus enables an examination of the predictive significance of QuickSIN for subsequent changes in audiometric thresholds on a large sample. We hypothesized that QuickSIN performance predicts hearing threshold degradation. If so, this would suggest that individuals experiencing perceptual difficulties in noise might be at a higher risk of developing clinical hearing loss sooner, and/or experience a faster progression of clinical hearing loss, than those who do not. In combination with an increasing adoption of QuickSIN in clinical settings, this could lead to a greater likelihood of detecting hearing difficulties at an earlier stage, potentially increasing patient awareness and enabling more timely clinical interventions.

## Materials and methods

### Study population

This study utilizes data from the BLSA from 1,128 individuals, collected between 2012 and 2023. The length of the visit interval was typically determined by the individual's age, with those aged below 60 years being assessed every four years, individuals aged between 60 and 80 years seen every two years, and those above 80 years tested annually. This study was approved by the National Institutes of Health (NIH) Intramural Institutional Review Board (IRB). Informed consent was obtained from all participants. All methods were performed in accordance with the NIH IRB (#03AG0325). Hearing assessments including audiometry, QuickSIN and SDS, as well as self-reported hearing and screening questions about noise exposure history and hearing aid usage, were conducted during each visit.

### Hearing assessment

Pure-tone audiometry was conducted to measure hearing thresholds, the softest sounds that can be heard, at frequencies of 0.5, 1, 2, 4, and 8 kHz in each ear[45]. The pure tone audiometry was conducted using the Hughson Westlake procedure[46]. The audiometric threshold is defined as the lowest intensity at which the participant was able to detect the signal 50% of the time. Lower thresholds represent better hearing. The pure tone average (PTA) was calculated by averaging the audiometric pure tone thresholds from both ears at 0.5, 1, 2, 4 and 8 kHz. In this study, we follow standard convention by defining normal hearing (NH) as all hearing thresholds below 25 dB HL, mild hearing loss (mild HL) as all hearing thresholds below or equal to 40 dB HL but not NH, and hearing loss (HL) as at least one hearing threshold above 40 dB HL.

The QuickSIN test[17] assesses speech perception in noise and involves presenting participants with diotic sentences in varying signal-to-noise ratio (SNR) levels from 0 to 25 dB SNR (70 dB HL presentation level). Two lists (List 1 and 2) from the QuickSIN test were presented to both ears of each participant diotically. Each list contained six sentences, and each sentence had five key words, where each correct identification of a key word was worth 1 point. Participants were asked to verbally repeat the sentence after they heard it. Each sentence was played at a different SNR. The SNR decreased from 25 to 0 dB, with a 5-dB step size. The SNR loss (ranging from −4.5 to 25.5) was then calculated by subtracting the total score from all six sentences of the list from 25.5; thus, a perfect performance score corresponds to a score of −4.5. The SNR loss indicates the severity of speech perception difficulties in noise: the higher the QuickSIN SNR loss, the worse the speech perception in noise. A QuickSIN SNR loss below 3 is considered normal. The average of the SNR loss scores from the two lists was used as the QuickSIN score in subsequent analyses.

### Data filtering

As of November 2023, a total of 1128 individuals in the BLSA dataset had undergone comprehensive hearing tests, with 1046 individuals completing both audiometry and QuickSIN during a single visit at least once. Figure 1 illustrates the data filtering process for identifying those individuals that fit our inclusion criteria. We first excluded visits that contained missing or non-responding frequencies in audiometry (i.e., the participant did not respond to the highest possible sound level at that frequency) or missing or invalid scores in QuickSIN. Since we were interested in markers for the early development and progression of clinical hearing loss, only those data from participants with a minimum of two visits in the filtered subset were selected for analysis, resulting in 664 individuals and 2094 visits.

The primary analysis included individuals from all three groups: NH, mild HL, and HL. Since we were especially interested in markers for the early development and progression of clinical hearing loss, we separately analyzed the data from the subgroup of individuals with mild HL and NH (having at least one visit where all audiometric thresholds were $\leq$ 40 dB HL). Additionally, we also queried the data of the subgroup of NH

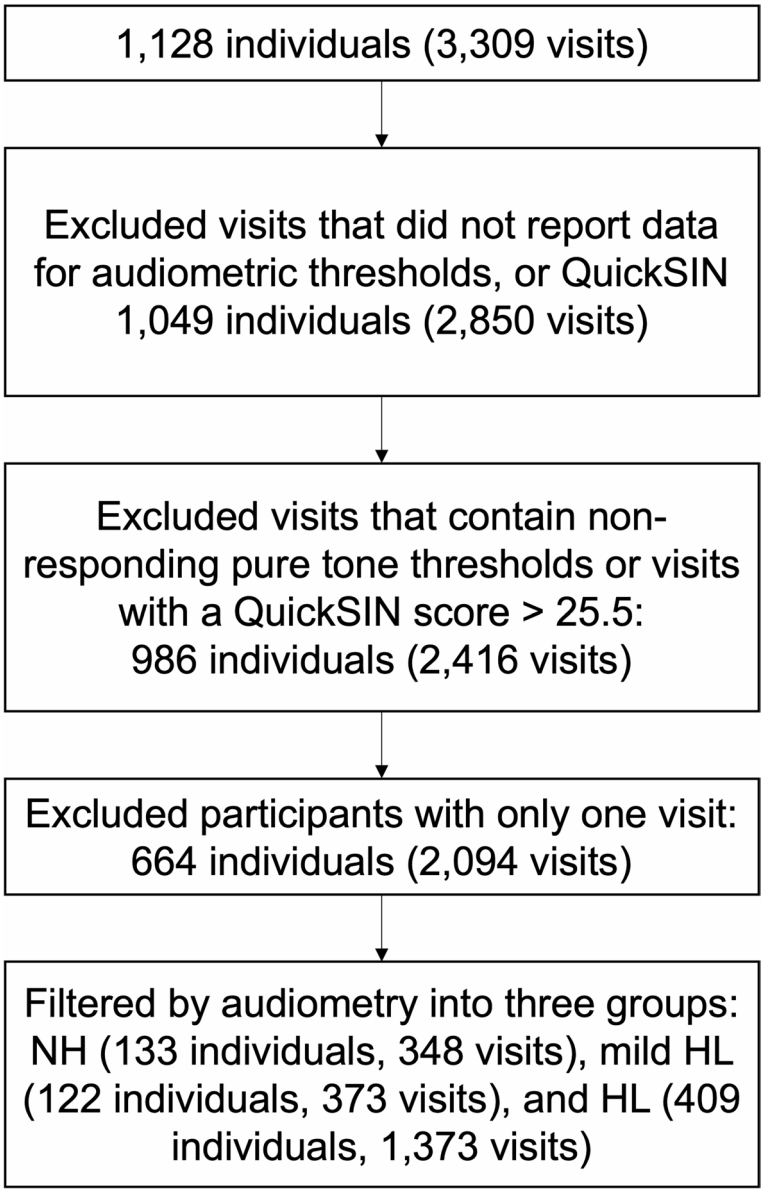


**Fig. 1**. The data filtering process used to remove individuals and visits that did not fit inclusion criteria.

individuals (having at least one visit where all thresholds were $\leq$ 25 dB HL) in order to assess the predictive power of speech measures prior to any clinically detectable hearing loss.

### Statistical analysis

PTA and QuickSIN each contributed one value per individual and visit. We hypothesized that QuickSIN predicts future changes in PTA. Thus, we tested whether QuickSIN from the previous visit could predict the change in PTA at the subsequent visit. The primary analysis employed a linear mixed model (LMM, shown in expression 1), with change in PTA from the baseline visit ( $\mathrm{PTA}_i - \mathrm{PTA}_0$) as the outcome variable. The term 'baseline visit' refers to the first visit of a participant during which a hearing assessment that was analyzed in this dataset was conducted. The subscript $i$ denotes the $i$th visit, with $i = 0$ indicating the baseline visit. We will refer to the measures collected at the baseline visit as 'baseline' measures. The QuickSIN score from the previous visit ( $\mathrm{QuickSIN}_{i-1}$) was included to predict the PTA change at the current visit $i$. We use the term 'lagged' to refer to measures collected during the previous visit $i - 1$. Note that the previous visit may or may not correspond to the baseline visit, depending on whether more than one follow-up visits occurred. The predictor variables included sex, baseline PTA, baseline age, time elapsed from the baseline visit ( $\mathrm{Time}_i$), lagged QuickSIN (the QuickSIN score from the previous visit, $\mathrm{QuickSIN}_{i-1}$), as well as the interaction term between Time and lagged QuickSIN. A random intercept was included for each participant to account for inter-individual variability.

$$\begin{aligned}(\mathrm{PTA}_i - \ \mathrm{PTA}_0) \ \sim \ & \mathrm{Time}_i + \mathrm{sex} + \mathrm{PTA}_0 + \mathrm{age}_0 + \mathrm{QuickSIN}_{i-1} \\ & + \mathrm{Time}_i \ \times \mathrm{QuickSIN}_{i-1} + (1 \mid \text{individual ID}).\end{aligned} \tag{1}$$

To test if lagged QuickSIN and the baseline measures significantly add to the prediction of PTA change, we compared the models in expression 1 and 2. In model 2 the lagged QuickSIN and its interaction with time were left out compared to model 1.

$$(\mathrm{PTA}_i - \ \mathrm{PTA}_0) \ \sim \ \mathrm{Time}_i + \mathrm{sex} + \mathrm{PTA}_0 + \mathrm{age}_0 + (1 \mid \text{individual ID}). \tag{2}$$

To better compare the predictive performance of these LMMs, we conducted leave-one-out cross-validation (LOOCV). In this process, data from one participant were designated as the validation set, while data from all other participants were used to train the model. This procedure was repeated for each participant in turn. For the training mean squared error (MSE), predictions were calculated using two approaches: (1) predictions based on the full model, incorporating both fixed effects and random intercepts, and (2) predictions based solely on the fixed effects. For the validation MSE, only predictions based on fixed effects were calculated, as the random intercept was not available for the validation participant (whose data were excluded from the training set). The training and validation MSEs allow us to (a) compare predictions based solely on fixed effects (e.g., Time, Sex, Age and Baseline PTA) to those generated by the full model, which incorporates both fixed effects and the random intercept, and (b) assess whether the models exhibit overfitting.

Several more common measures were also calculated for comparing of these LMMs, including Akaike information criterion (AIC), Bayesian information criterion (BIC) and log likelihood. AIC and BIC are estimators of the prediction error, where a better model prediction results in a smaller AIC or BIC value[47]. Compared to AIC, BIC places a heavier penalty on models with more parameters. Log likelihood denotes how well the model fits the data, where a better model is associated with a higher log likelihood. A likelihood ratio test comparing the log likelihood values of these two models calculates the associated $p$-value of the comparison[48].

Data filtering and data cleaning were conducted with Python Pandas[49]. The statistical analyses were conducted in R version 4.1.2[50], while the linear mixed model analyses were conducted using the nlme package[51].

## Results

### Baseline characteristics and change of PTA as a function of age

Table 1 provides an overview of the demographics and test measurements of individuals at the time of their baseline visit, i.e., the first visit of each individual that was included in the analyses. Note that age at the baseline visit varied substantially across the entire participant cohort. Additionally, note that the maximum number of follow-up years were relatively short (12 years), because the current hearing test battery was only adopted in October 2012. The initial analysis encompassed the entire group, considering all individuals. Subsequently, we narrowed our focus to separate subgroups to examine the data more closely for early signs of developing clinical hearing loss, first examining individuals with NH and mild HL, and then only NH individuals. This sequential approach enabled a detailed exploration of the effects within the subgroup experiencing the onset of clinical hearing loss.

Figure 2 displays the individual trajectories of logarithmically spaced PTA values as a function of age, showcasing an exponential increase in PTA over time. This observed trend is similar across both male and female individuals.

### Predictive significance of QuickSIN for a subsequent change in PTA

Past research has shown that speech perception in noise, such as QuickSIN, might capture additional information (beyond PTA, age and sex) that could better reflect individuals' real-world hearing experiences[2,9,52,53]. We were particularly interested in whether QuickSIN predicted the development and progression of clinical hearing loss. In this section, we investigate whether QuickSIN can be used to predict a subsequent change in PTA, since speech in noise difficulties might precede audiometric threshold elevation. Specifically, we employed a LMM (Expression 1) to examine the effect of lagged QuickSIN on a change in PTA since the baseline visit. In this way, we investigate the predictive significance of QuickSIN for a subsequent change in PTA.

| | All individuals (HL, mild HL, NH) (100%) | Mild HL & NH individuals (38.4%) | NH individuals (20.7%) |
|---|---|---|---|
| # Individuals | 664 | 255 | 133 |
| # Visits | 2094 | 721 | 348 |
| Age (years) | 68.63 (13.15) | 58.61 (12.53) | 53.18 (12.60) |
| Female (%) | 56.46 | 66.27 | 65.41 |
| Baseline PTA (dB) | 29.96 (14.63) | 15.99 (6.26) | 11.52 (3.99) |
| Baseline QuickSIN (SNR Loss) | 2.36 (3.25) | 0.61 (1.74) | 0.27 (1.23) |
| Baseline SDS | 90.81 (15.11) | 97.41 (2.89) | 97.98 (2.39) |
| Max follow-up (years) | 5.53 (2.47) | 6.05 (2.25) | 6.16 (2.15) |

**Table 1**. Characteristics of the individuals at the time of their baseline visit. . Means are reported, with standard deviation in parentheses.

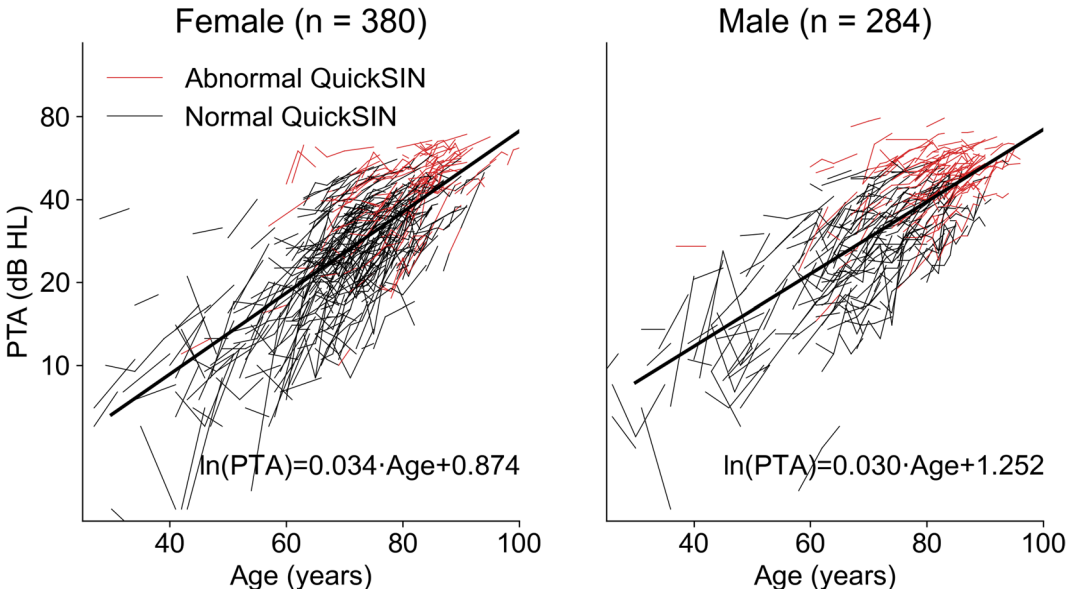


**Fig. 2**. PTA change (in dB HL) as a function of age for all included individuals (see Fig. 1 for inclusion criteria). Each trace corresponds to an individual. Thin traces in red denote instances where the QuickSIN score of the preceding visit exceeded 3, while thin traces in black indicate the QuickSIN score of the previous visit was below 3. The bold black lines are the predicted mean trajectories estimated by a linear model with ln(PTA) as the outcome variable and age as the independent variable, with equations representing the linear regression models for each sex.

| | **Estimate (SE, *p*-value)** | | |
|---|---|---|---|
| Predictors | All individuals | Mild HL & NH individuals | NH individuals |
| # of individuals | 664 | 255 | 133 |
| # of observations | 1430 | 466 | 215 |
| PTA (0.5–8 kHz) at the baseline visit | −0.043(0.016, **0.004**) | −0.098 (0.048, **0.041**) | −0.357 (0.085, <**0.001**) |
| Time (# years since the baseline visit) | 1.166 (0.045, <**0.001**) | 1.067 (0.074, <**0.001**) | 0.641 (0.100, <**0.001**)) |
| Age at the baseline visit | 0.102 (0.016, <**0.001**) | 0.135 (0.026, <**0.001**) | 0.125 (0.029, <**0.001**) |
| Sex | 0.081 (0.289, 0.779) | −0.035 (0.494, 0.943) | −0.573 (0.592, 0.335) |
| Lagged QuickSIN | 0.097 (0.063, 0.122) | −0.33 (0.192, 0.086) | −0.964 (0.408, **0.021**) |
| Lagged QuickSIN x Time | 0.011 (0.010, 0.256) | 0.064 (0.030, **0.033**) | 0.216 (0.067, **0.002**) |

**Table 2**. LMM results for PTA (0.5–8 kHz). PTA change as a function of lagged QuickSIN (expression (1)). *P*-values less than 0.05 are shown in bold.

The model fitting statistics of the LMM revealed that both PTA and Age at the baseline visit were predictive of a change in PTA thresholds, and this was the case for all groups (Table 2). The LMM estimate value denotes the predicted change at the next visit, controlling for all other factors. For example, for every year increase in age at the baseline visit, the predicted PTA at the subsequent visit deteriorates by 0.1 dB (All individuals; Age), controlling for all other factors in the LMM. Concretely, for two people with the same sex, baseline PTA, QuickSIN from the previous visit and number of years past the baseline visit, if person A was one year older at the baseline visit than person B, person A's predicted PTA would be 0.1 dB worse than that of person B at the subsequent visit. In the Mild HL & NH group, a higher lagged QuickSIN score, indicative of poorer speech perception in noise, significantly predicted a larger change in PTA with aging (Mild HL & NH individuals; Lagged QuickSIN x Time), and this effect was also present in the other NH subgroup (NH individuals; Lagged QuickSIN x Time).

The interaction between Lagged QuickSIN and Time with respect to change in PTA for the NH subgroup is visualized in Fig. 3, which plots the LMM prediction of PTA change for different representative Lagged QuickSIN scores and Time. As outlined in Methods, a QuickSIN score of −1.5 denotes near perfect speech perception in noise (the best score for QuickSIN is −4.5, where all 30 key words in a list were identified), while QuickSIN scores of 1.5 and 5 indicate normal and mild hearing difficulties in noise, respectively. The LMM revealed that the higher (worse) the Lagged QuickSIN score (from the previous visit), the steeper the increase in PTA thresholds since the baseline visit. Concretely, while a QuickSIN score of −1.5 predicted a PTA deterioration of

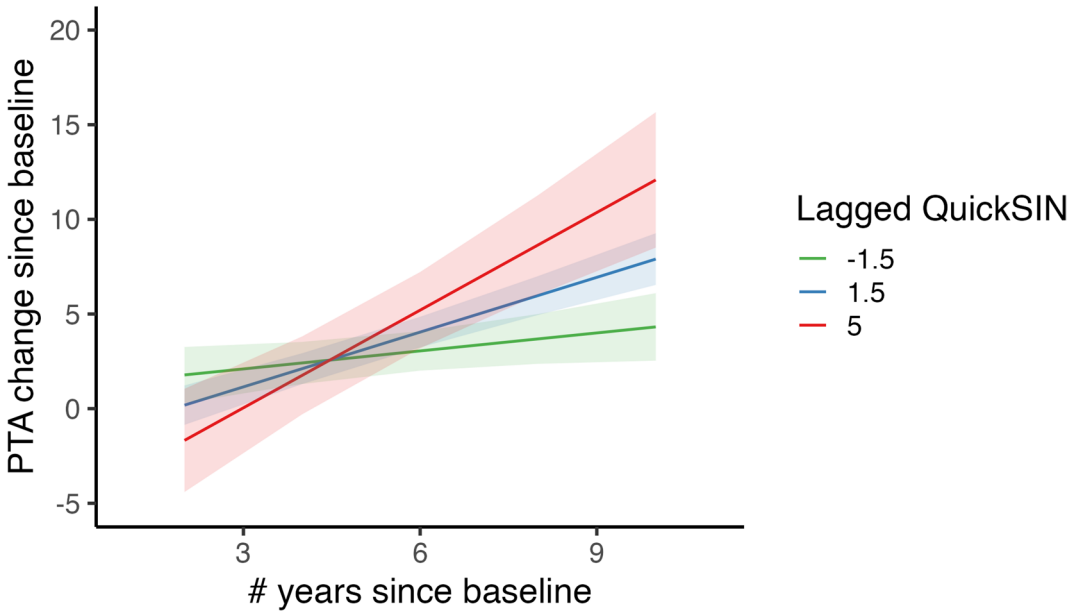


**Fig. 3**. Interaction between Time and Lagged QuickSIN on PTA change since the baseline visit. The LMM-predicted PTA change since the baseline visit is shown for the NH subgroup. The x-axis represents the number of years since the baseline visit. The y-axis represents the LMM predicted outcome, i.e., the PTA change since the baseline visit. The predicted values correspond to different representative QuickSIN scores of −1.5 (near perfect speech perception in noise), 1.5 (normal speech perception in noise), and 5 (mild speech perception difficulties in noise), and are shown in different colors. The shaded area represents the 95% confidence interval.

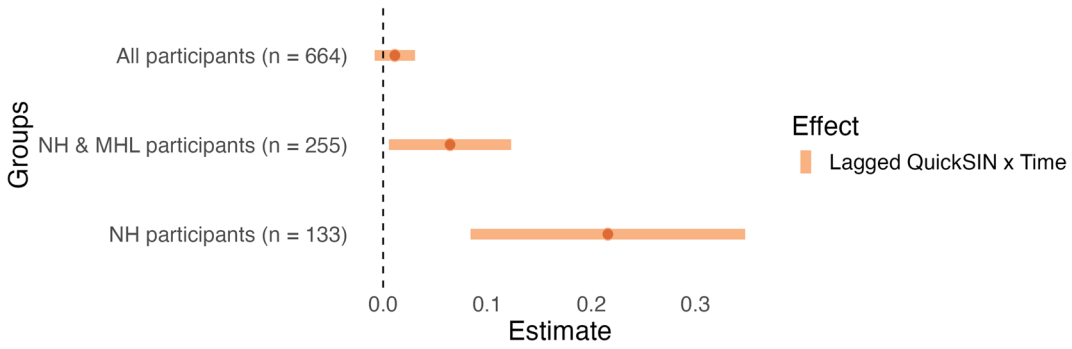


**Fig. 4**. Forest plot of the LMM estimates of the interactions. The bars represent the 95% confidence interval for the coefficients. The estimate of the interaction between the Lagged QuickSIN and Time was positive and significantly different from zero for the two subgroups.

~ 3 dB over the next 10 years (presumably reflecting normal PTA decline), a QuickSIN score of 5 predicted a PTA deterioration of about four times that much (~ 12 dB) over the same time period.

For these longitudinal analyses, we used the raw values of variables for the LMM to maximize interpretability of the findings. However, longitudinal analyses sometimes also introduce latent variables to reduce the noise for repeated measures (e.g. test-retest variability). In the Supplemental Material, we show that the pattern of the above results persists with the use of latent variables for QuickSIN.

Figure 4 visualizes these findings by depicting the estimates of the interaction of Lagged QuickSIN x Time, split up for the three groups. In all three groups, the worse the QuickSIN score from the previous visit, the stronger the deterioration in PTA in the subsequent visit.

To further investigate the predictive significance of QuickSIN for developing clinical hearing loss, we conducted a likelihood ratio test among two LMMs on the subgroup with NH individuals. LMM (1) was the full model (expression 1), LMM (2) was the full model with QuickSIN and its interaction with time removed (expression 2). The model that included the Lagged QuickSIN scores (expression 1) was significantly better at predicting a change in PTA than the one without (expression 2) ($p = 0.0042$) (Table 3).

| | $MSE_{tr\_full}$ | $MSE_{tr\_fixed}$ | $MSE_{val}$ | AIC | BIC | logLike | Test | likeRatio | *p*-value |
|---|---|---|---|---|---|---|---|---|---|
| LMM (1) | 5.62 | 11.70 | 13.12 | 1147.22 | 1177.55 | −564.61 | | | |
| LMM (2) | 5.98 | 12.28 | 13.51 | 1154.14 | 1177.74 | −570.07 | 1 vs. 2 | 10.92 | **0.0042** |

**Table 3**. Results for the comparison between models (1) and (2). $MSE_{tr\_full}$: mean squared error of the full model on the training set with leave-one-out cross validation. $MSE_{tr\_fixed}$: mean squared error of the model using only fixed effects (excluding the random intercept) on the training set with leave-one-out cross validation. $MSE_{val}$: mean squared error of the model using only fixed effects (excluding the random intercept) on the validation set with leave-one-out cross validation. AIC: Akaike information criterion; BIC: bayesian information criterion; loglike: log likelihood; likeratio: likelihood ratio. *P*-values less than 0.05 are shown in bold.

## Discussion

Despite a substantial amount of research into the causes and consequences of hearing difficulties in noise, to date there exists no conclusive evidence that individuals experiencing hearing difficulties in noise are at risk of developing clinical hearing loss sooner, or experiencing a faster deterioration of hearing. In this study, we examined whether hearing difficulties in noise could predict a subsequent deterioration of PTA. We found that among individuals with NH or mild HL, those with a worse QuickSIN score at the previous visit were more likely to have a deterioration in PTA at the subsequent visit than individuals with a better QuickSIN score.

The association between QuickSIN and PTA as individuals age could stem from the fact that speech perception in noise requires a highly integrated auditory system, which is arguably not as important for traditional pure tone audiometry. Aging and noise exposure might initially cause deficits in speech perception in noise before becoming evident in audiometric assessments[5,11,12,28,54,55].

One potential mechanism for hearing difficulties in noise preceding audiometric threshold shift that has attracted intensive research in the past decade is cochlear synaptopathy, which describes the loss of synapses between inner hair cells (IHCs) and auditory nerve fibers (ANFs)[28]. An influential study in mice showed that a 70% reduction of ANFs led to a 17-dB shift in behavioral thresholds for tone detection in noise, while no threshold shift was observed in quiet listening conditions[33]. This suggests that cochlear synaptopathy can induce hearing difficulties in noise without resulting in a clinically detectable threshold shift. Post-mortem research in both animals and humans has shown that the rate of ANF loss precedes and exceeds that of outer hair cells (OHCs) or IHCs in the context of aging or noise exposure[7,28,29,34,35]. Consequently, perceptual difficulties in noise may be a precursor to developing clinically elevated hearing thresholds (which are thought to be primarily due to the loss of OHCs and IHCs)[56].

The prediction estimates for a subsequent deterioration of PTA became stronger when we limited the analysis from the group of individuals with either mild HL or NH to the group with only NH individuals (put differently, when we excluded participants who already had clinically detectable HL). This suggests that the predictive significance was higher in individuals at early, possibly subclinical stages of hearing loss, or even without any signs of hearing loss. The comparatively larger confidence interval of this effect, however, is likely a result of the reduction in sample size, from over 255 individuals and over 721 observations (mild HL and NH subgroup) to 133 individuals and 348 observations (NH subgroup). Follow up analyses, once the sample sizes of the NH and the Mild HL groups in the BLSA have increased further, are therefore necessary to confirm these findings. The prognostic significance of QuickSIN, however, is diminished in the group with all individuals. One potential explanation is that as clinical hearing loss worsens, hearing difficulties in noise become increasingly dominated by audibility[4], making QuickSIN less reflective of the underlying risk factors for future PTA deterioration.

The underlying etiology behind the prognostic significance of QuickSIN scores for PTA changes remains unclear. As mentioned in the Introduction, speech perception deficits in noise are complex and are influenced by various factors including both attenuation and distortion. The latter includes aspects such as cognitive factors[5,15,57], auditory temporal resolution[9], and frequency selectivity[5]. While the current study focused on audiological measures (QuickSIN and PTA), the BLSA dataset al.so includes measures of cognitive function[45,58]. We did not include these cognitive factors here as our aim was to specifically focus on whether QuickSIN can predict changes in clinical pure tone thresholds. One future direction would be to investigate whether and how cognitive function – such as working memory, attention, and executive processes - are related to changes in QuickSIN and PTA, and how, in turn, cognitive functions may affect speech in noise performance.

While the BLSA offers a tremendous opportunity to the hearing research community, it has some limitations. First, the maximum follow-up duration for standardized hearing assessments that include QuickSIN is currently only 12 years (as of November 2023), which is short compared to the average lifespan of humans. Second, the measures collected in the BLSA may not fully capture the nuances of the participants' hearing profiles. For instance, audiometric thresholds were only assessed at 0.5, 1, 2, 4, and 8 kHz. Including additional frequencies such as 0.75, 1.5, 3, and 6 kHz, along with lower frequencies (e.g., 0.125, 0.25 kHz) and extended high frequencies (>8 kHz), could offer a more detailed understanding of hearing function. Furthermore, QuickSIN was administered using diotic presentation, preventing per-ear analyses. Third, the dataset currently only has 133 individuals with normal hearing at their baseline visit, which is suboptimal for the linear mixed model we conducted. A larger dataset, or more data from the BLSA in the future, will likely lead to more definitive findings.

In our analyses using linear predictive models, we assumed that the change in PTA was a linear function of time and its interaction with QuickSIN. However, such a linear relationship might not hold true, especially at older ages and/or over longer individual tracking times[59,60]. As depicted in Fig. 2, the relationship between PTA

shift and age was exponential. While the 12-year tracking period (and an average tracking period of 5–6 years) in this analysis may mitigate this to some extent, future research may need to explore nonlinear transformations of time or PTA when more data during longer periods of time become available. Furthermore, pure tone thresholds are typically not uniform across frequency, but are more degraded towards higher frequencies[27,60]. In the LMM, we treated PTA in a simplified manner by averaging thresholds across all tested frequencies (0.5 to 8 kHz, at octave intervals) and both ears. A more granular analysis of frequency-specific thresholds could offer deeper insights into the nuances of hearing changes across different frequencies.

We used a LMM approach to investigate the prognostic significance of QuickSIN on future PTA change in an imbalanced longitudinal dataset (different number of visits across participants, at variable time intervals). The LMM incorporated a random intercept for each individual to effectively capture between-individual variability across repeated measures. However, since the random intercepts are specific to the individuals included in the model, this method has limited generalizability. Consequently, the LMM predicted outcomes for new participants poorly using only the fixed effects ($MSE_{val}$). The MSEs based on fixed effects alone were comparable between the training and validation sets ($MSE_{tr\_fixed}$ vs. $MSE_{val}$), indicating that the model was not overfitted to the data. Furthermore, predictions using the full model, which included the random intercept, resulted in a significantly lower MSE ($MSE_{tr\_full}$). This finding suggests the existence of substantial individual variability that cannot be explained by fixed effects such as Age, Sex, Time, or any other fixed factors included in our model. To improve clinical applicability, future models should explore more generalizable approaches, such as machine learning or time-to-event analysis, so that individuals at higher risk of hearing loss can be identified and in order to provide actionable insights for novel cases[61–63]. It is also worth noting that the distribution of participants within the current dataset (or other/future datasets) should be carefully considered when conducting such analyses (i.e., machine learning methods or time-to-event analyses). Older participants have more frequent visits and are thus overrepresented compared to younger ones, leading to a dataset that is skewed toward older individuals.

In summary, the results of our analyses suggest that speech perception in noise does indeed offer valuable insights into the functional status of the auditory system. Therefore, clinicians should consider including tests that assess speech perception in noise (e.g., QuickSIN) as part of their standard audiological assessments. Individuals with a poorer QuickSIN score should take particular care to protect their hearing from noise exposure and ototoxic drugs. They may also leverage hearing technologies such as mild gain hearing aids to improve their experiences and quality of life in noisy environments[64].

## Data availability

The data analyzed in this study are part of the large BLSA dataset; applications to use BLSA data should be made online (https://www.blsa.nih.gov/blsa-data-use).

Received: 30 June 2024; Accepted: 16 June 2025
Published online: 01 July 2025

## References

1. Lin, F. R., Niparko, J. K. & Ferrucci, L. Hearing loss prevalence in the united States. *Arch. Intern. Med.* **171**, 1851 (2011).
2. Fitzgerald, M. B., Gianakas, S. P., Qian, Z. J., Losorelli, S. & Swanson, A. C. Preliminary guidelines for replacing word-recognition in quiet with speech in noise assessment in the routine audiologic test battery. (2023).
3. Smith, M. L., Winn, M. B. & Fitzgerald, M. B. A Large-Scale study of the relationship between degree and type of hearing loss and recognition of speech in quiet and noise. *Ear Hear.* **45**, 915–928 (2024).
4. Plomp, R. Auditory handicap of hearing impairment and the limited benefit of hearing aids. *J Acoust. Soc. Am* **63**, (1978).
5. Parthasarathy, A., Hancock, K. E., Bennett, K., DeGruttola, V. & Polley, D. B. Bottom-up and top-down neural signatures of disordered multi-talker speech perception in adults with normal hearing. *eLife* **9**, e51419 (2020).
6. Plomp, R. & Mimpen, A. M. Speech-reception threshold for sentences as a function of age and noise level. *J. Acoust. Soc. Am.* **66**, 1333–1342 (1979).
7. Wu, P., O'Malley, J. T., de Gruttola, V. & Liberman, M. C. Primary neural degeneration in Noise-Exposed human cochleas: correlations with outer hair cell loss and Word-Discrimination scores. *J. Neurosci.* **41**, 4439–4447 (2021).
8. Stebbins, W. C., Hawkins, J. E., Johnsson, L. G. & Moody, D. B. Hearing thresholds with outer and inner hair cell loss. *Am. J. Otolaryngol.* **1**, 15–27 (1979).
9. Humes, L. E. Factors underlying individual differences in speech-recognition threshold (SRT) in noise among older adults. *Front. Aging Neurosci.* **13**, 702739 (2021).
10. Henry, K. S. Animal models of hidden hearing loss: does auditory-nerve-fiber loss cause real-world listening difficulties? *Mol. Cell. Neurosci.* **118**, 103692 (2022).
11. Schneider, B. A. & Pichora-Fuller, M. K. Age-related changes in temporal processing: implications for speech perception. *Semin Hear.* **22**, 227–240 (2001).
12. Lorenzi, C., Gilbert, G., Carn, H., Garnier, S. & Moore, B. C. J. Speech perception problems of the hearing impaired reflect inability to use Temporal fine structure. *Proc. Natl. Acad. Sci.* **103**, 18866–18869 (2006).
13. Gómez-Álvarez, M., Johannesen, P. T., Coelho-de-Sousa, S. L., Klump, G. M. & Lopez-Poveda, E. A. The relative contribution of cochlear synaptopathy and reduced inhibition to age-related hearing impairment for people with normal audiograms. *Trends Hear.* **27**, 23312165231213191 (2023).
14. Karunathilake, I. M. D. et al. Effects of aging on cortical representations of continuous speech. *J. Neurophysiol.* **129**, 1359–1377 (2023).
15. Wong, P. C. M., Uppunda, A. K., Parrish, T. B. & Dhar, S. Cortical mechanisms of speech perception in noise. *J. Speech Lang. Hear. Res.* **51**, 1026–1041 (2008).
16. Dillon, H. *Hearing Aids* (Thieme Medical, 2012).
17. Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J. & Banerjee, S. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.* **116**, 2395–2405 (2004).
18. Wilson, R. H. Development of a speech-in-multitalker-babble paradigm to assess word-Recognition performance. *J. Am. Acad. Audiol.* **14**, 453–470 (2003).
19. Nilsson, M., Soli, S. D. & Sullivan, J. A. Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. *J. Acoust. Soc. Am.* **95**, 1085–1099 (1994).

20. Tillman, T. W. & Carhart, R. An expanded test for speech discrimination utilizing CNC monosyllabic words: northwestern university auditory test no. 6: https://doi.org/10.21236/AD0639638 (19996).
21. Picou, E. M. MarkeTrak 10 (MT10) survey results demonstrate high satisfaction with and benefits from hearing aids. *Semin Hear.* **41**, 021–036 (2020).
22. Fitzgerald, M. B. et al. Speech-in-Noise assessment in the routine audiologic test battery: relationship to perceived auditory disability. *Ear Hear.* **45**, 816–826 (2024).
23. Davidson, A., Marrone, N., Wong, B. & Musiek, F. Predicting hearing aid satisfaction in adults: a systematic review of speech-in-noise tests and other behavioral measures. *Ear Hear.* **42**, 1485–1498 (2021).
24. Jason Qian, Z. et al. Evaluation of asymmetries in Speech-in noise abilities in audiologic screening for vestibular Schwannoma. *Ear Hear.* **44**, 1540–1547 (2023).
25. Bramhall, N., Konrad-Martin, D., McMillan, G. P. & Griest, S. E. Auditory brainstem response altered in humans with noise exposure despite normal outer hair cell function. *Ear Hear.* **38**, e1–e12 (2017).
26. Dubno, J. R., Dirks, D. D. & Morgan, D. E. Effects of mild hearing loss and age on speech recognition in noise. *J. Acoust. Soc. Am.* **72**, S34–S35 (1982).
27. Dubno, J. R., Lee, F. S., Matthews, L. J. & Mills, J. H. Age-Related and Gender-Related changes in monaural speech recognition. *J. Speech Lang. Hear. Res.* **40**, 444–452 (1997).
28. Kujawa, S. G. & Liberman, M. C. Adding insult to injury: cochlear nerve degeneration after 'temporary' noise-induced hearing loss. *J. Neurosci.* **29**, 14077–14085 (2009).
29. Sergeyenko, Y., Lall, K., Liberman, M. C. & Kujawa, S. G. Age-related cochlear synaptopathy: an early-onset contributor to auditory functional decline. *J. Neurosci.* **33**, 13686–13694 (2013).
30. Zhao, F. & Stephens, D. A critical review of king-kopetzky syndrome: hearing difficulties, but normal hearing? *Audiol. Med.* **5**, 119–124 (2007).
31. Liu, J., Stohl, J., Lopez-Poveda, E. A. & Overath, T. Quantifying the impact of auditory deafferentation on speech perception. *Trends Hear* **28**, (2024).
32. Lopez-Poveda, E. A. Why do I hear but not understand? stochastic undersampling as a model of degraded neural encoding of speech. *Front Neurosci* **8**, (2014).
33. Resnik, J. & Polley, D. B. Cochlear neural degeneration disrupts hearing in background noise by increasing auditory cortex internal noise. *Neuron* **109**, 984–996e4 (2021).
34. Valero, M. D. et al. Noise-induced cochlear synaptopathy in rhesus monkeys (Macaca mulatta). *Hear. Res.* **353**, 213–223 (2017).
35. Wu, P. et al. Primary neural degeneration in the human cochlea: evidence for hidden hearing loss in the aging ear. *Neuroscience* **407**, 8–20 (2019).
36. Trevino, M. & Lobarinas, E. Current topics in hearing research: deafferentation and threshold independent hearing loss. *Hear. Res.* https://doi.org/10.1016/j.heares.2021.108408 (2021).
37. Wilson, B. S., Tucci, D. L., O'Donoghue, G. M., Merson, M. H. & Frankish, H. A lancet commission to address the global burden of hearing loss. *Lancet* **393**, 2106–2108 (2019).
38. Lin, F. R. Hearing loss and cognition among older adults in the united States. *J. Gerontol. Biol. Sci. Med. Sci.* **66A**, 1131–1136 (2011).
39. Dawes, P. et al. Hearing loss and cognition: the role of hearing aids, social isolation and depression. *Plos One*. **10**, e0119616 (2015).
40. Shukla, A., Cudjoe, T. K. M., Lin, F. R. & Reed, N. S. Functional hearing loss and social engagement among medicare beneficiaries. *J. Gerontol. Ser. B*. **76**, 195–200 (2021).
41. Gopinath, B. et al. Hearing-impaired adults are at increased risk of experiencing emotional distress and social engagement restrictions five years later. *Age Ageing*. **41**, 618–623 (2012).
42. Ciorba, A., Bianchini, C., Pelucchi, S. & Pastore, A. The impact of hearing loss on the quality of life of elderly adults. *Clin. Interv Aging*. **159** https://doi.org/10.2147/CIA.S26059 (2012).
43. Shock, N. W. *Normal Human Aging: The Baltimore Longitudinal Study of Aging*. (US Department of Health and Human Services, Public Health Service, National Institutes of Health, 1984)
44. Clark, J. G. Uses and abuses of hearing loss classification. *ASHA* **23**, 493–500 (1981).
45. Jiang, K. et al. Associations of audiometric hearing and speech-in-noise performance with cognitive decline among older adults: the Baltimore longitudinal study of aging (BLSA). *Front. Neurol.* **13**, 1029851 (2022).
46. Carhart, R. & Jerger, J. F. Preferred method for clinical determination of pure-tone thresholds. *J. Speech Hear. Disord.* **24**, 330–345 (1959).
47. Vrieze, S. I. Model selection and psychological theory: A discussion of the differences between the akaike information criterion (AIC) and the bayesian information criterion (BIC). *Psychol. Methods*. **17**, 228–243 (2012).
48. Pinheiro, J. & Bates, D. *Mixed-Effects Models in S and S-PLUS* (Springer Science & Business Media, 2006).
49. McKinney, W. Data structures for statistical computing in python https://doi.org/10.25080/Majora-92bf1922-00a (2010)
50. Core Team, R. R. R: A language and environment for statistical computing. (2013).
51. Pinheiro, J. et al. nlme: Linear and nonlinear mixed effects models. (2023).
52. Vermiglio, A. J., Soli, S. D., Freed, D. J. & Fang, X. The effect of stimulus audibility on the relationship between pure-tone average and speech recognition in noise ability. *J. Am. Acad. Audiol.* **31**, 224–232 (2020).
53. Wilson, R. H., McArdle, R. A. & Smith, S. L. An evaluation of the BKB-SIN, HINT, quicksin, and WIN materials on listeners with normal hearing and listeners with hearing loss. *J. Speech Lang. Hear. Res.* **50**, 844–856 (2007).
54. Humes, L. E. et al. Central presbycusis: A review and evaluation of the evidence. *J. Am. Acad. Audiol.* **23**, 635–666 (2012).
55. Liberman, M. C., Epstein, M. J., Cleveland, S. S., Wang, H. & Maison, S. F. Toward a differential diagnosis of hidden hearing loss in humans. *Plos One*. **11**, e0162726 (2016).
56. Wu, P., O'Malley, J. T., de Gruttola, V. & Liberman, M. C. Age-Related hearing loss is dominated by damage to inner ear sensory cells, not the cellular battery that powers them. *J. Neurosci.* **40**, 6357–6366 (2020).
57. Porto, L., Wouters, J. & Van Wieringen, A. Speech perception in noise, working memory, and attention in children: a scoping review. *Hear. Res.* **439**, 108883 (2023).
58. Armstrong, N. M. et al. Temporal sequence of hearing impairment and cognition in the baltimore longitudinal study of aging. *J. Gerontol. Ser. A*. **75**, 574–580 (2020).
59. Moore, D. R. et al. Relation between Speech-in-Noise threshold, hearing loss and cognition from 40–69 years of age. *Plos One*. **9**, e107720 (2014).
60. Pearson, J. D. et al. Gender differences in a longitudinal study of age-associated hearing loss. *J. Acoust. Soc. Am.* **97**, 1196–1205 (1995).
61. Clark, T. G., Bradburn, M. J., Love, S. B. & Altman, D. G. Survival analysis part I: basic concepts and first analyses. *Br. J. Cancer*. **89**, 232–238 (2003).
62. Goodfellow, I., Bengio, Y. & Courville, A. *Deep Learning* (MIT Press, 2016).
63. Guo, S. *Survival Analysis* (Oxford University Press, 2010).
64. Roup, C. M., Post, E. & Lewis, J. Mild-Gain hearing aids as a treatment for adults with self-reported hearing difficulties. *J. Am. Acad. Audiol.* **29**, 477–494 (2018).

## Author contributions

J.L. obtained data access, framed the idea of this manuscript, conducted analysis and wrote the manuscript. H.Z. provided the statistical analysis framework and revised the manuscript. E.M.S reviewed the manuscript. J.S framed the idea of this manuscript and revised the manuscript. T.O. obtained data access, framed the idea of this manuscript and revised the manuscript.

## Declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-07454-z.

**Correspondence** and requests for materials should be addressed to J.L.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.