International Journal of Audiology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/iija20

Discrimination of degrees of auditory performance from the digits-in-noise test based on hearing status

Nicole M. Armstrong , Berthe C. Oosterloo , Pauline H. Croll , Mohammad Arfan Ikram & André Goedegebure

To cite this article: Nicole M. Armstrong , Berthe C. Oosterloo , Pauline H. Croll , Mohammad Arfan Ikram & André Goedegebure (2020): Discrimination of degrees of auditory performance from the digits-in-noise test based on hearing status, International Journal of Audiology, DOI: 10.1080/14992027.2020.1787531

To link to this article: https://doi.org/10.1080/14992027.2020.1787531

Published online: 16 Jul 2020.

INTERNATIONAL JOURNAL OF AUDIOLOGY
https://doi.org/10.1080/14992027.2020.1787531

ORIGINAL ARTICLE

OPEN ACCESS Check for updates

Discrimination of degrees of auditory performance from the digits-in-noise test based on hearing status

Nicole M. Armstrong[a], Berthe C. Oosterloo[b,c], Pauline H. Croll[b,c,d], Mohammad Arfan Ikram[c] and André Goedegebure[b]

[a]Laboratory of Behavioral Neuroscience, National Institute on Aging, Baltimore, MD, USA; [b]Department of Otorhinolaryngology, Head and Neck Surgery, Erasmus University Medical Center, Rotterdam, the Netherlands; [c]Department of Epidemiology, Erasmus University Medical Center, Rotterdam, the Netherlands; [d]Department of Radiology and Nuclear Medicine, Erasmus University Medical Center, Rotterdam, the Netherlands

ABSTRACT

Objective: To discriminate among degrees of auditory performance of the Digits-in-Noise (DIN) test.

Design: We performed Pearson's correlations and age- and sex-adjusted linear regression models to examine the correlation between pure-tone average (PTA) from pure-tone audiometric tests and speech recognition thresholds (SRT) from the DIN test. Then, optimal SRT cut-points by PTA-defined hearing status (0–25 dB HL [normal], 26–40 dB HL [mild hearing loss], 41–50 dB HL [moderate hearing loss]) were compared across three methods: Youden, Nearest, and Liu. SRT-defined categories of auditory performance were compared to PTA-defined hearing categories to examine the convergence of similar categories.

Study Sample: 3422 Rotterdam Study participants aged 51–98 years between 2011 and 2014

Results: The correlation between SRT and PTA was 0.65 (95% Confidence Interval: 0.63, 0.67) in the overall sample. The variability of SRT explained by PTA after age and sex adjustment was 54%. Optimal cut-points for the overall sample across the three methods were: ≤ -5.55 dB SNR (normal); > -5.55 to ≤ -3.80 dB SNR (insufficient performance); > -3.80 dB SNR (poor performance). When comparing the SRT- or PTA-defined categories, 59.8% had concordant hearing categories and 40.2% had discordant hearing categories.

Conclusions: Discrimination of degrees of auditory performance may add greater utility of the DIN test.

ARTICLE HISTORY
Received 24 December 2019
Revised 17 June 2020
Accepted 19 June 2020

KEYWORDS
Signal-to-noise ratio; pure-tone audiometry; hearing loss; older adults; speech reception thresholds

Introduction

Age-related hearing loss is a highly prevalent condition affecting up to two-thirds of adults aged 70 years and older (Homans et al. 2017). Older individuals have a higher prevalence of hearing loss with greater levels of severity than younger and middle-aged individuals (Agrawal, Platz, and Niparko 2008; Goman and Lin 2016; Homans et al. 2017). Among older adults with hearing loss, a major complaint is difficulty hearing speech in noisy environments (Pronk, Deeg, and Kramer 2018; Pronk et al. 2013). This difficulty requires additional attentional resources to be recruited to aid in understanding verbal communication in noisy environments, which may have cascading effects, leading to loneliness, social isolation, and depression. Aside from age, prevalence of hearing loss differs between men and women with men having a higher prevalence of hearing loss, as compared to women (Agrawal, Platz, and Niparko 2008; Goman and Lin 2016; Homans et al. 2017). Both age and sex need to be considered when examining the prevalence of hearing loss.

To test the ability to comprehend speech in noisy environments, assessments have been developed to measure the speech reception threshold (SRT). SRT is the difference between the level of presented speech and background noise at which an individual can correctly reproduce 50% of words or sentences. This measure is considered to be more representative of an individual's hearing ability in real-life situations than pure-tone audiometry or speech recognition in quiet environments (Grant and Walden 2013; Houtgast and Festen 2008; Taylor 2003). One such assessment is the Digits-in-Noise (DIN) test, which requires individuals to repeat three spoken numbers (a digit triplet) that are presented through a listening device while continuous noise is playing in the background (Smits, Goverts, and Festen 2013). These tests are strongly correlated with pure-tone averages (PTA), ranging from 0.7 to 0.8 from variations in PTA definitions and sample characteristics (Jansen et al. 2010; Koole et al. 2016; Potgieter, Swanepoel, and Smits 2018b; Potgieter et al. 2018a; Smits, Kapteyn, and Houtgast 2004; Watson et al. 2012). Thus, sensitivity and specificity of the DIN is >80% in terms of detection of hearing loss (Denys et al. 2019; Folmer et al. 2017; Jansen et al. 2013; Koole et al. 2016; Potgieter, Swanepoel, and Smits 2018b; Potgieter et al. 2018a; Smits, Kapteyn, and Houtgast 2004; Vercammen et al. 2018; Vlaming et al. 2014; Watson et al. 2012). The DIN test has been designated as the National Hearing Test in The Netherlands. Several versions of the DIN in different

CONTACT André Goedegebure a.goedegebure@erasmusmc.nl Department of Otorhinolaryngology, Head and Neck Surgery, Erasmus MC, P.O. Box 2040, Rotterdam, 3000CA, the Netherlands

languages have been developed (Dawes et al. 2014; Watson et al. 2012; Zokoll et al. 2012).

While there are established PTA cut-points to define peripheral hearing categories, there are no standardised cut-points for the DIN test. These cut-points tend to vary by study. In a previous study using Rotterdam Study data, Koole et al. (2016) reported that an appropriate cut-point for the DIN test to identify abnormal hearing would be generally in the range between 0 and −5 decibels signal-to-noise ratio (dB SNR). Smits, Merkus, and Houtgast (2006) reported using the following categories: good when SRT >−5.5 dB SNR, insufficient when −5.5 ≤ SRT ≤ −2.8 dB SNR, and poor hearing ability when SRT >−2.8 dB SNR. Dawes et al. (2014) used the following cut-points: < −5.5 dB SNR for normal auditory performance, −5.5 dB SNR to −3.5 dB SNR for insufficient auditory performance, and > −3.5 dB SNR for poor auditory performance. The investigators of these studies did not examine whether this cut-point would be appropriate by age and sex, especially since there are differences in the prevalence estimates of hearing loss by these characteristics (Goman and Lin 2016). Taken together, age and sex may result in different sets of cut-points, since the overall cut-points reported by other studies may not reflect differences in both age and sex. These cut-points can be used to identify older adults with a relatively poor auditory functional performance by age and sex. Additionally, these cut-points can be used in prospective cohort studies as well as studies that only use DIN, not pure-tone audiometry.

The aims of the study are threefold. First, we created optimal cut-points of SRT based on PTA-defined categories in the overall sample to examine agreement across different calculation methods for the determination of optimal cut-points and to investigate whether these cut-points remained consistent with other studies. Second, as an exploratory aim, we evaluated whether these optimal cut-points should be corrected for age and/or sex, due to differences in prevalence of hearing loss, as defined by PTA, by age and sex reported by other population-based studies. Third, we compared SRT- and PTA-defined hearing categories to determine whether there were incongruent categories between SRT and PTA. We hypothesised that individuals with incongruent hearing categories could have differences in hearing ability not found when administered when only pure-tone audiometry is evaluated. Pure-tone audiometry does not test speech recognition abilities in background noise (Kramer et al. 1996; Smits, Kapteyn, and Houtgast 2004), while the DIN can (Smits, Goverts, and Festen 2013; Taylor 2003).

Materials and methods

Participants

This cross-sectional study was embedded in the Rotterdam Study, a prospective cohort study of determinants and consequences of aging (Ikram et al. 2017). Enrolment criteria are described in more detail elsewhere (Ikram et al. 2017). Briefly, from the population registry of Ommoord, a suburb of Rotterdam, The Netherlands, adults aged ≥45 years were invited to participate in the Rotterdam Study. Enrolment took place on a voluntary basis. In 2011, hearing assessments were introduced into the study protocol. This study included participants with both completed pure-tone audiometry and DIN ($N = 3667$). Participants were excluded if their SRT was greater than two standard deviations (SD) above the mean ($n = 169$) or their PTA was >50 dB HL ($n = 76$) to avoid that the noise would be presented below threshold levels. Exclusion criteria were defined to ensure the examination of the comprehension of digits through the background noise and lack of reliability with large intra-test SD. The DIN would be a test of hearing ability if the participant did not comprehend the digits through the background noise. The analytic sample consisted of 3422 participants. The institutional review board (Medical Ethics Committee) and the review board of The Netherlands Ministry of Health, Welfare, and Sports approved this study, and participants provided written informed consent.

Pure-tone audiometry

Pure-tone audiometry, a measure of peripheral auditory system sensitivity, was performed in a soundproof booth by one trained health care professional (Ikram et al. 2017). A computer-based audiometry system (Decos Technology Group, version 210.2.6 with AudioNigma interface) and TDH-39 headphones were used. dB hearing thresholds were measured according to the ISO-standard 8253 (International Organisation for Standardisation [ISO], 2010). Air conduction was tested for both ears at the following frequencies: 250, 500, 1000, 2000, 4000, and 8000 hertz (Hz). Masking was done according to the method of Hood (Hood 1960). A speech-frequency PTA of air conduction thresholds at 250, 500, 1000, 2000, 4000 Hz was calculated for each ear. The best hearing ear for every participant was determined by taking the average threshold over all hearing frequencies, as identified by the lowest hearing thresholds of one of the ears. When hearing thresholds were equal in both ears, alternately the right or the left ear was chosen. We also examined PTA in low (250, 500, 1000 Hz), middle (500, 1000, 2000 Hz), and high (2000, 4000, 8000 Hz) frequencies. We used WHO-defined hearing categories, defined by PTA: normal hearing (0−25 dB HL), mild hearing loss (26−40 dB HL), moderate hearing loss (41−60 dB HL), and severe hearing loss (>60 dB HL) (Stevens et al., 2013; Ikram et al. 2017). The last category does not play a role in the current study, as we excluded all participants with PTA >50 dB HL.

Digits-in-noise test

SRT from DIN, a 3-minute speech-in-noise test, was used in separate analyses. After pure-tone audiometry was performed, the DIN was administered. The DIN was measured in the best ear (average of 500 and 4000 Hz in the pure tone audiogram) using TDH-39P headphones with MX-41/AR cushions. When the average loss in both ears was equal, left and right ears were measured alternatively.

Detailed testing procedures regarding the DIN have been previously reported (Smits, Goverts, and Festen 2013). Briefly, the DIN test consists of lists with 24 digit triplets. A pre-recorded male-spoken speech-signal consisted of three consecutive digits (from 0 to 9). The digits are presented against a background of 65 dB of sound pressure level (SPL). The starting level of the speech is determined by repeatedly presenting the first triplet, increased more intensely by 4 dB, until it is first heard correctly. The initial starting level was at −8 dB SNR. Participants' echoed triplet was rated correct or incorrect by an experienced audiometrist. The measurement then follows an adaptive up-down procedure, using 2-dB steps. If the participant repeated the triplet incorrectly, the next triplet was presented 2 dB more intensely. A stable signal-to-noise ratio (SNR) is reached after the first four presentations. The overall SNR was calculated by

Table 1. Sample characteristics from the Rotterdam Study ($N = 3422$).

Sample Characteristics	Overall ($N = 3422$)	Normal Hearing ($n = 1583$)	Mild Hearing Loss ($n = 1422$)	Moderate Hearing Loss ($n = 417$)	*p* Value for the difference among hearing categories
Age Category, *n* (%)					<0.001
<60 years	929 (27.2)	686 (43.3)	221 (15.5)	22 (5.3)	
60–70 years	1783 (52.1)	802 (50.7)	821 (57.7)	160 (38.4)	
>70 years	710 (20.8)	95 (6.0)	380 (26.7)	235 (56.4)	
Male, *n* (%)	1499 (43.8)	597 (37.7)	679 (47.8)	223 (53.5)	<0.001
PTA of all frequencies, dB HL, mean (SD)	22.0 (10.0)	13.4 (4.1)	26.1 (4.2)	40.5 (4.1)	<0.001
PTA of high frequencies, dB HL, mean (SD)	29.7 (15.3)	17.0 (6.7)	36.3 (8.5)	55.8 (7.9)	<0.001
PTA of middle frequencies, dB HL, mean (SD)	19.2 (9.5)	11.7 (4.3)	22.3 (5.0)	36.7 (5.7)	<0.001
PTA of low frequencies, dB HL, mean (SD)	13.1 (7.4)	8.7 (4.4)	14.8 (5.9)	24.0 (7.3)	<0.001
Speech recognition threshold, in dB SNR, mean (SD)	−4.4 (3.3)	−6.2 (1.5)	−4.2 (2.3)	1.5 (4.3)	<0.001

SD: standard deviation; PTA: pure-tone average in the better ear; dB HL: decibels of hearing level; dB SNR: decibels of speech-to-noise ratio. Note: normal hearing was defined by a PTA ranging from 0 to 25 dB HL in the better ear. Mild hearing loss was defined by a PTA ranging from 26 to 40 dB HL in the better ear. Moderate hearing loss was defined by a PTA ranging from 41 to 50 dB HL in the better ear. *p*-values for the differences among hearing categories are derived from either analysis of variance (ANOVA) for continuous variables or chi-squared tests for categorical variables.

taking the average SNR of the digit triplets from the 5th presentation to the 24th presentation. The 24th triplet is not presented, but its level is calculated from the response to the 23rd triplet. All examinations took place in a sound-treated booth with a clinical audiometer (Decos audiology workstation, version 210.2.6, with AudioNigma interface; Decos Audiology, Inc., Peachtree City, GA). Lower SRTs indicate better performance.

Statistical analysis

First, we used chi-squared tests for categorical variables and analysis of variance for continuous variable to examine differences in sample characteristics by hearing category. All analyses were performed using Stata 15.1 (StataCorp 2017). Type I error level was set to 0.05 for analyses.

Second, we examined the Pearsons correlation coefficients between the SRT and PTA in the overall sample and by age group, sex, and age group and sex. We also used multivariable linear regression models to examine the amount of variability of SRT explained by PTA (R^2) with adjustment of age group, sex, and two-way interaction between age group and sex.

Third, we used three methods of cut-point estimation to define appropriate cut-points for the SRT from the DIN test to compare across the methods. The first was the Youden Index, which maximises the sum of sensitivity and specificity. Youden Index is the vertical distance between the 45 degree line and the point of the receiver-operating characteristic (ROC) curve, where higher values are better than lower values (Youden 1950). The second was the Nearest method, which finds the cut-point on the receiver-operating curve closest to (0,1), the point with perfect sensitivity and specificity (Coffin and Sukhatme 1997). The third was the Liu method, which maximises the product of sensitivity and specificity (Liu 2012). Standard errors for each method were bootstrapped 1000 times to estimate 95% confidence intervals for each of the cut-points. We also used ROCs to examine the area under the curve values for the optimal cut-points. The reference standard was based on PTA-defined degrees of hearing loss. The targeted disorder was the degrees of hearing loss (mild and moderate vs. normal). The DIN was the index test.

Lastly, after defining the categories based on the cut-point estimation, we compared the SRT-defined categories to the PTA-defined categories. We used the kappa statistic to determine the agreement in the categorisation of hearing status between SRT and PTA, since SRT and PTA were moderately correlated. Then, we examined the incongruent pairs to determine if the incongruency of the categorisation may have some clinical meaning.

Data availability

Data can be obtained on request. Requests should be directed towards the management team of the Rotterdam Study (secretariaat.epi@erasmusmc.nl), which has a protocol for approving data requests. Because of restrictions based on privacy regulations and informed consent of the participants, data cannot be made freely available in a public repository.

Results

Characteristics of study sample

Table 1 contains the characteristics for the overall sample and by hearing status defined by PTA. The mean age of participants was 64.9 (Standard Deviation, [SD] = 7.0) years, ranging from 51.5 to 98.6 years. Overall, there were 929 (27.2%) participants <60 years, 1783 (52.1%) participants aged 60–70 years, and 710 (20.8%) participants >70 years. Approximately, 43.8% ($n = 1499$) of the sample was male. The mean PTA of all frequencies was 22.0 (SD = 10.0) decibels of hearing level (dB HL). The mean SRT was −4.4 (SD = 3.3) dB SNR. All hearing categories differed in age category, sex, PTA, and SRT (Table 1).

Table 2. The estimated optimal cut-points of the speech reception threshold in decibels of signal-to-noise ratio (dB SNR) for the comparison of hearing statuses by three methods.

	Cut-point	Bootstrapped SE	95% CI	Sensitivity	Specificity	AUC	Youden Index (J)	SE (J)
Overall Sample								
Mild Hearing Loss vs. Normal Hearing (REF) (N = 3005)								
Empirical Cutpoint Estimation								
Liu	−5.60	0.18	(−5.96, −5.24)	0.69	0.75	0.72		
Youden	−5.55	0.27	(−6.07, −5.03)	0.69	0.75	0.72	0.44	0.02
Nearest	−5.55	0.12	(−5.79, −5.31)	0.69	0.75	0.72		
Moderate Hearing Loss vs. Normal Hearing (REF) (N = 2000)								
Empirical Cutpoint Estimation								
Liu	−3.80	0.16	(−4.12, −3.48)	0.94	0.95	0.95		
Youden	−3.80	0.18	(−4.15, −3.45)	0.94	0.95	0.95	0.90	0.01
Nearest	−3.80	0.13	(−4.05, −3.55)	0.94	0.95	0.95		
Males								
Mild Hearing Loss vs. Normal Hearing (REF) (N = 1276)								
Empirical Cutpoint Estimation								
Liu	−5.55	0.22	(−5.98, −5.12)	0.70	0.76	0.73		
Youden	−4.95	0.31	(−5.56, −4.34)	0.59	0.87	0.73	0.46	0.02
Nearest	−5.55	0.11	(−5.76, −5.34)	0.70	0.76	0.73		
Moderate Hearing Loss vs. Normal Hearing (REF) (N = 820)								
Empirical Cutpoint Estimation								
Liu	−4.05	0.19	(−4.42, −3.68)	0.95	0.95	0.95		
Youden	−4.05	0.23	(−4.49, −3.61)	0.95	0.95	0.95	0.90	0.02
Nearest	−4.05	0.13	(−4.31, −3.79)	0.95	0.95	0.95		
Females								
Mild Hearing Loss vs. Normal Hearing (REF) (N = 1729)								
Empirical Cutpoint Estimation								
Liu	−5.30	0.21	(−5.71, −4.89)	0.65	0.79	0.72		
Youden	−5.30	0.21	(−5.70, −4.90)	0.65	0.79	0.72	0.44	0.02
Nearest	−5.75	0.21	(−6.15, −5.35)	0.74	0.69	0.72		
Moderate Hearing Loss vs. Normal Hearing (REF) (N = 1180)								
Empirical Cutpoint Estimation								
Liu	−3.80	0.29	(−4.38, −3.22)	0.96	0.95	0.95		
Youden	−3.80	0.30	(−4.38, −3.22)	0.96	0.95	0.95	0.89	0.03
Nearest	−3.80	0.28	(−4.35, −3.25)	0.94	0.95	0.94		

SE: standard error; CI: confidence interval; REF: reference group; AUC: area under the receiver-operating curve.

Correlations between signal-to-noise ratio and pure-tone average and variability of signal-to-noise ratio explained by pure-tone average

The correlation between SRT and PTA in the better ear was 0.65 (95% Confidence Interval, CI: 0.63, 0.67) in the overall sample. The correlations between SRT and PTA of low, middle, and high frequencies in the better ear were 0.42 (95% CI: 0.39, 0.45), 0.67 (95% CI: 0.65, 0.69), and 0.63 (95% CI: 0.61, 0.65) respectively. The correlation was 0.53 (95% CI: 0.49, 0.58) among those < age 60, 0.66 (95% CI: 0.63, 0.69) among those aged 60–70 years, and 0.61 (95% CI: 0.56, 0.66) among those aged >70 years. The similar correlations could be attributed to the inclusion criteria. Males (r = 0.69, 95% CI: 0.67, 0.72) had a higher correlation between SNR and PTA of all frequencies in the better ear than females (r = 0.62, 95% CI: 0.59, 0.65) in the overall sample and by age group. Additionally, using a linear regression model, the variability of SRT explained by PTA was 54%, even after adjusting by sex and age.

Optimal cut-points for signal-to-noise ratio

Table 2 shows the optimal cut-points for the DIN test for the overall sample and by sex. Since results were similar across the three methods, we only report results from the Liu method in this section. Results from all three methods are written in Table 2. The optimal cut-point for DIN test was −5.55 (95% CI: −6.07, −5.03) dB SNR to discriminate between mild hearing loss and normal hearing. The sensitivity was 0.69, while the specificity was 0.75. The area under the receiver-operative curve (AUC) was 0.72 (95% CI: 0.78, 0.81). When discriminating moderate hearing loss from normal hearing, the optimal cut-point for DIN test was −3.80 dB SNR (Table 2). The sensitivity was 0.94, while the specificity was 0.95. The AUC was 0.95 (95% CI: 0.94, 0.96). Males and females differed slightly in terms of optimal cut-points (Table 2).

Optimal cut-points differed by age group (Table 3). Among the three age groups, the optimal cut-point ranged from −5.55 dB SNR (among adults younger than age 60) to −5.05 dB SNR (among adults older than age 70) when discriminating between mild hearing loss and normal hearing. There was even a wider range in optimal cut-points when discriminating between moderate hearing loss and normal hearing. The range was from −4.55 dB SNR (among adults younger than age 60) to −3.20 dB SNR (among adults older than age 70). Supplemental Figure 1 features the receiver-operating curves (ROC) by age group and hearing status to illustrate where the optimal cut-point was located on the ROC when using the Liu method. Males and females had similar cut-points by age and hearing status (Table 3).

Comparison of categories defined by both pure-tone average and signal-to-noise ratio

Using the optimal cut-points for SRT, we defined the categories of auditory performance. Normal auditory performance was defined as ≤ −5.55 dB SNR. Insufficient auditory performance was defined as > −5.55 to ≤ −3.80 dB SNR. Poor auditory performance was defined as > −3.80 dB SNR. We compared the mean hearing thresholds by frequency for the categories defined by both PTA and SRT separately. Figure 1 shows the distribution

Table 3. Comparison of the optimal cut-points among the methods by hearing status, age category, and sex.

	Overall ($N = 3422$)		Males ($N = 1499$)		Females ($N = 1923$)	
	Cut-point	95% CI	Cut-point	95% CI	Cut-point	95% CI
Younger than Age 60						
Mild Hearing Loss vs. Normal Hearing (REF) (N = 907)						
Empirical Cutpoint Estimation						
Liu	−5.60	(−5.96, −5.24)	−5.60	(−5.93, −5.27)	−5.80	(−6.34, −5.26)
Youden	−5.55	(−5.93, −5.17)	−5.55	(−5.90, −5.20)	−5.75	(−6.38, −5.12)
Nearest	−5.55	(−5.86, −5.24)	−5.55	(−5.87, −5.23)	−5.75	(−6.19, −5.31)
Moderate Hearing Loss vs. Normal Hearing (REF) (N = 708)						
Empirical Cutpoint Estimation						
Liu	−4.60	(−5.54, −3.66)	−3.90	(−4.53, −3.27)	−4.60	(−7.21, −1.99)
Youden	−4.55	(−5.52, −3.58)	−3.80	(−4.48, −3.12)	−4.55	(−7.66, −1.44)
Nearest	−3.80	(−4.72, −2.88)	−3.80	(−4.53, −3.07)	−4.55	(−7.47, −1.63)
Ages 60 to 70						
Mild Hearing Loss vs. Normal Hearing (REF) (N = 1623)						
Empirical Cutpoint Estimation						
Liu	−5.20	(−5.75, −4.65)	−5.20	(−5.77, −4.63)	−5.40	(−5.83, −4.97)
Youden	−5.15	(−5.63, −4.74)	−4.95	(−5.39, −4.51)	−5.30	(−5.92, −4.68)
Nearest	−5.75	(−6.17, −5.33)	−5.75	(−6.24, −5.26)	−5.75	(−6.18, −5.32)
Moderate Hearing Loss vs. Normal Hearing (REF) (N = 962)						
Empirical Cutpoint Estimation						
Liu	−4.10	(−4.75, −3.45)	−4.10	(−5.02, −3.18)	−3.90	(−4.61, −3.19)
Youden	−4.05	(−4.74, −3.36)	−4.05	(−4.98, −3.12)	−3.80	(−4.48, −3.12)
Nearest	−4.05	(−4.48, −3.62)	−4.05	(−4.65, −3.45)	−3.80	(−4.47, −3.13)
Age >70						
Mild Hearing Loss vs. Normal Hearing (REF) (N = 475)						
Empirical Cutpoint Estimation						
Liu	−5.10	(−5.89, −4.31)	−5.00	(−6.07, −3.93)	−6.00	(−6.98, −5.02)
Youden	−5.75	(−7.25, −4.25)	−4.05	(−5.60, −2.50)	−5.90	(−7.13, −4.67)
Nearest	−5.05	(−5.71, −4.39)	−4.95	(−5.67, −4.23)	−5.55	(−6.35, −4.75)
Moderate Hearing Loss vs. Normal Hearing (REF) (N = 330)						
Empirical Cutpoint Estimation						
Liu	−3.30	(−4.04, −2.56)	−3.30	(−4.07, −2.53)	−3.30	(−4.21, −2.39)
Youden	−3.20	(−3.90, −2.50)	−3.10	(−3.96, −2.24)	−3.20	(−4.11, −2.29)
Nearest	−3.20	(−3.77, −2.63)	−3.10	(−3.95, −2.25)	−3.20	(−3.91, −2.49)

CI: confidence interval.

of hearing thresholds from pure-tone audiometry by frequency for each of the hearing categories either defined by PTA or SRT. While there is more variability across all frequencies among the PTA-defined hearing categories, there is not as much variability among the SRT-defined hearing categories (Figure 1).

Then, we compared the categories of hearing status defined by SRT to those defined by PTA (Supplemental Table 1). There were 2046 (59.8%) participants with congruent categories, and 1376 (40.2%) participants had incongruent categories. The kappa between the categories was low (kappa = 0.38, SE of kappa = 0.01). Supplemental Figure 2 shows the mean hearing thresholds by frequency of the pure-tone audiometric test by incongruent categories based on SRT- and PTA-defined categories. Incongruent groups had similar mean hearing thresholds at lower frequencies (250, 500, and 1000 Hz). Starting at 2000 Hz, incongruent groups have higher mean threshold values, suggesting that the differences between congruent and incongruent groups lie in higher frequencies.

Discussion

In a sample from a population-based cohort study of older adults, we created optimal cut-points for the DIN test, based on PTA-defined hearing categories. First, we examined the relationship between the SRT and PTA. PTA did not fully explain SRT. The SRT-defined categories were: ≤ -5.55 dB SNR (normal), > -5.55 to ≤ -3.80 dB SNR (insufficient), and > -3.80 dB SNR (poor) in the overall sample. Second, we determined expected SRT-defined categories based on demographics and audiometric thresholds to identify groups that may under- or over-perform on auditory tasks, i.e. DIN test. We found differences in optimal cut-points by age group, sex, and combined age group and sex. Comparison of the categories of hearing status based on cut-points and PTA showed that 59.8% had congruent categories with a kappa of 0.38.

A previous study by Koole et al. (2016) in Rotterdam Study data studied the potential of the DIN test for hearing loss screening in an elderly population. Our current study is an addition to this previous paper as the analyses examining the use of the DIN test to estimate age-related hearing loss differed in a few ways. First, we restricted the sample to those with a PTA <50 dB HL. We did this to ensure that we were not testing hearing, but comprehension of digits through the background noise. Second, we examined the optimal cut-points of SRT by hearing status, age group, sex, and combined age group and sex. The optimal cut-points from the overall sample differed slightly between males and females as well as across age groups. Third, we used different methods to compare and determine the optimal cut-points, instead of using several SRT values, to examine the sensitivity and specificity of SRT based on PTA-defined categories. Our cut-points were different from the ones that are reported in Koole et al. (2016) for the overall sample, as we also provided more detailed categorisation of hearing level, distinguishing hearing level across three categories of hearing level instead of two. Finally, we examined the incongruent categories to determine if their hearing thresholds across frequencies differed from one another.

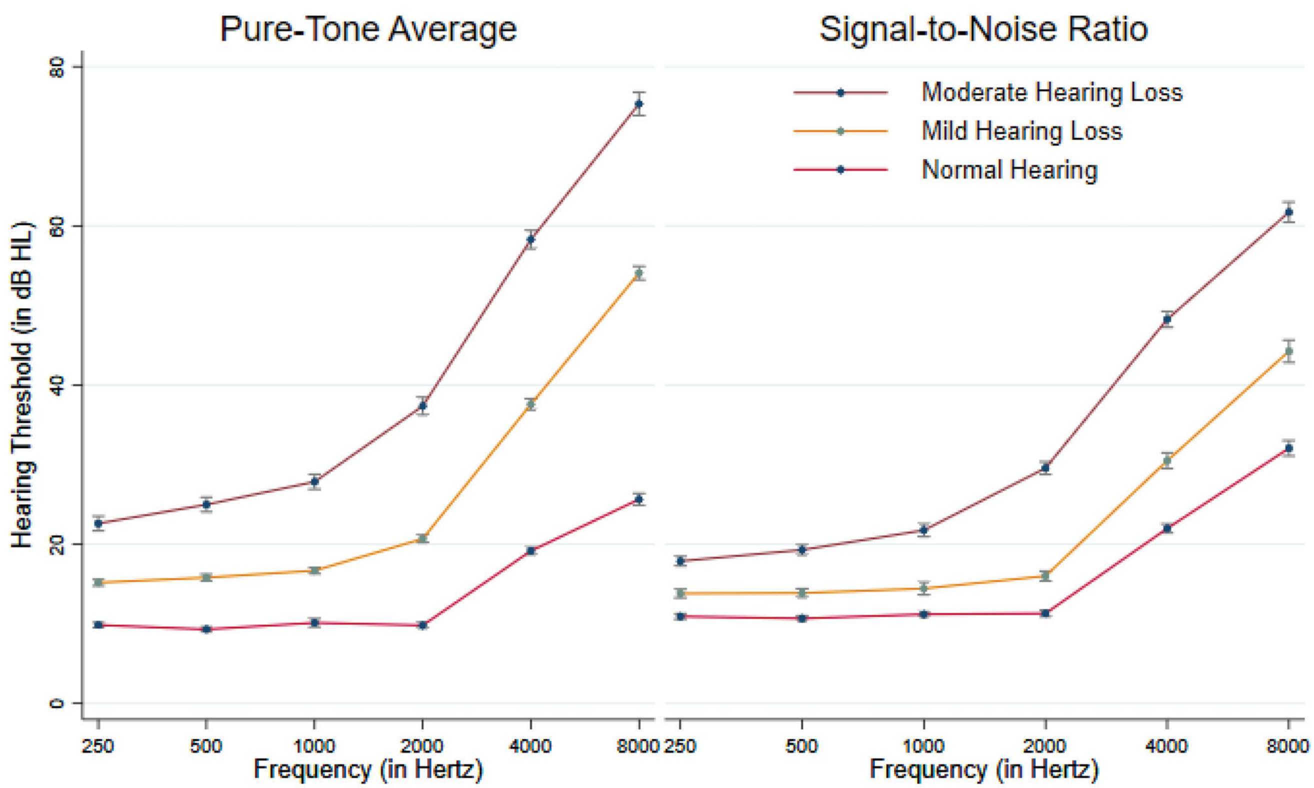

Figure 1. Hearing thresholds (air conduction) averaged per frequency by hearing status defined by either pure-tone average or signal-to-noise ratio. This figure shows the differences in hearing thresholds averaged per frequency among moderate hearing loss, mild hearing loss, and normal hearing groups. WHO-defined hearing categories are defined by pure-tone average: normal hearing (0–25 decibels of hearing level [dB HL]), mild hearing loss (26–40 dB HL), moderate hearing loss (41–50 dB HL). The hearing categories defined by speech reception thresholds are: ≤ -5.55 decibels of signal-to-noise ratio (decibel (dB) SNR) (normal auditory performance); > -5.55 to ≤ -3.80 dB SNR (insufficient auditory performance); > -3.80 dB SNR (poor auditory performance).

Similar to Koole et al. (2016), we found that explained variability of SRT by PTA was 54%. Additionally, the sensitivity, specificity, and AUC were much lower when discriminating between mild hearing loss and normal hearing, as compared to discriminating between moderate hearing loss and normal hearing in all investigated (sub)groups. The cut-points we established for meaningful degrees of hearing loss were similar to the ones reported by Dawes et al. (2014) from UK Biobank data. Their cut-points were < -5.5 dB SNR for normal auditory performance, -5.5 dB SNR to -3.5 dB SNR for insufficient auditory performance, and > -3.5 dB SNR for poor auditory performance. However, they evaluated a sample of middle−aged adults ranging from 40 to 69 years of age, so we show that the cut-points do not differ as much in a population-based sample with a wider age range. Additionally, their cut-points were from statistically defined definitions not based on audiometric or functional levels.

There were differences in the cut-points among the age groups. The younger group with normal hearing had better speech reception thresholds than the oldest group with normal hearing did, which could have been attributed to differences in proportions of individuals with moderate hearing loss. The older age groups have a greater proportion of moderate hearing loss, while the younger age group has a greater asymmetric distribution of hearing loss towards the lower bound of the moderate hearing loss category. All auditory tasks consist of both peripheral and central aspects. Speech-in-noise tests involve more central processing than audiometry, but a simple speech-in-noise test, e.g. DIN, has more peripheral contributions than one, such as the coordinate response measure (Brungart et al. 2001).

Men and women had similar optimal cut-points for SRT-defined categories in the overall sample and by age group. This is in contrast to other cross-sectional studies that found that older men have significantly poorer SRTs than older women (Dawes et al. 2014; Pronk et al. 2013; Smits, Merkus, and Houtgast 2006). While the patterns of the estimated cut-points mirror findings from other studies, the corresponding 95% confidence intervals overlap with one another, suggesting that there are no significant differences between men and women overall and by age group. This could be attributed to our exclusion criteria, as men may have an earlier onset of peripheral hearing loss than women, thus having a higher probability of a PTA >50 dB HL.

When we examined the incongruent hearing categories defined by either SRT or PTA, we found that these groups did not differ much in terms of hearing thresholds at lower frequencies. These groups began to differentiate around 2000 Hz with the moderate hearing loss (SRT)/mild hearing loss (PTA) having the highest mean hearing threshold at that frequency, as compared to the other groups. This is in line with an older study by DePaolis, Janota, and Frank (1996) in which they showed that speech intelligibility has a functionality peak at 2000 Hz. High-frequency thresholds may better predict DIN performance than middle/low frequencies. High-frequency hearing loss typically shows hearing within the normal range from 250 Hz to 1000 Hz, but outside of the normal range starting at 2000 Hz. Decreased hearing acuity in high frequency ranges, from 2000 to 8000 Hz, is a primary sign of hearing loss (Wingfield, Tun, and McCoy 2005). Those with high-frequency hearing loss have difficulty understanding speech in noise and high-frequency consonant sounds. Our results confirm that adults with more pronounced high-frequency losses are at risk for relatively poor understanding in noisy conditions. However, we did use a standard DIN test with long-term average speech spectrum noise.

The strengths of the study are its large population-based sample and use of both pure-tone audiometry and DIN tests. There are some limitations of this study. The accuracy of the DIN test depends on attention. Fatigability may have influenced the accuracy of the test, due to full-day visit of assessments for the Rotterdam Study. Also, the majority of the sample is white, so racial/ethnic differences in hearing status were not evaluated. However, the Ommoord neighbourhood is representative of the Dutch population, thus tending to be generalisable to populations of other Western European countries. Furthermore, these optimal cut-points hold true for the standard DIN, but not for newer versions of the DIN using different stimulus approaches. While it is probable that the functional impact of a certain SRT from the DIN test will be different between younger and older people, our SRT-defined categories are not meant to address this difference. Rather, our SRT-defined categories are meant to identify those who under- or over-perform on the DIN test. Lastly, there could have been overestimation of the congruence of SRT- and PTA-defined categories, since the SRT-defined categories were defined using PTA cut-points. However, the kappa was low between SRT-defined and PTA-defined categories.

Conclusions

This is a large-scale study in which optimal cut-points to discriminate degrees of hearing loss were defined for the DIN test in the overall sample. The SRT-defined categories were: ≤ -5.55 dB SNR (normal), > -5.55 to ≤ -3.80 dB SNR (insufficient), and > -3.80 dB SNR (poor). We did not find many differences in optimal cut-points by age group, sex, and combined age group and sex, although the variability of SRT explained by PTA, age, and sex was 54% and the correlation between PTA and SRT was 0.65. When examining the incongruent groups, we found differences in hearing thresholds at high frequency ranges, from 2000 to 8000 Hz. Measuring hearing thresholds at higher frequency ranges could be more sensitive to the variability of cut-points, independent from PTA. This suggests the added clinical utility of the DIN test at higher stimulus levels in adults with less severe degrees of hearing loss.

Acknowledgements

The authors thank all the participants of the study for their time and Tekla Enser for performing the pure-tone audiometry and the speech-in-noise tests. This study is based on data collected in the context of the Rotterdam Study.

Disclosure statement

André Goedegebure and Pauline H. Croll report funding by Cochlear Ltd. The other authors have no conflicts of interest to disclose.

Funding

The Rotterdam Study is supported by the Erasmus Medical Centre and Erasmus University, Rotterdam, The Netherlands; the Organisation for Scientific Research; the Netherlands Organisation for Health Research and Development; the Research Institute for Diseases in the Elderly; the Netherlands Genomics Initiative; the Ministry of Education, Culture, and Science; the Ministry of Health, Welfare, and Sports; the European Commission (DG XII); and the Municipality of Rotterdam. Hearing-related research within the Rotterdam Study is partly funded by The Heinsius Houbolt Foundation. Separately, NMA is supported by the Intramural Research Programme at the National Institute on Aging, Baltimore, MD, USA.

Author contributions

NMA wrote the manuscript and analysed the data. All authors discussed the results and implications and commented on the manuscript at all stages, and they were responsible for the study design.

ORCID

Nicole M. Armstrong http://orcid.org/0000-0002-2479-1076

References

Agrawal, Y., E. A. Platz, and J. K. Niparko. 2008. "Prevalence of Hearing Loss and Differences by Demographic Characteristics among US Adults: Data from the National Health and Nutrition Examination Survey, 1999-2004." *Archives of Internal Medicine* 168 (14): 1522–1530. doi:10.1001/archinte.168.14.1522.

Brungart, D. S., B. D. Simpson, M. A. Ericson, and K. R. Scott. 2001. "Informational and Energetic Masking Effects in the Perception of Multiple Simultaneous Talkers." *The Journal of the Acoustical Society of America* 110 (5 Pt 1): 2527–2538. doi:10.1121/1.1408946.

Coffin, M., and S. Sukhatme. 1997. "Receiver Operating Characteristic Studies and Measurement Errors." *Biometrics* 53 (3): 823–837. doi:10.2307/2533545

Dawes, P., H. Fortnum, D. R. Moore, R. Emsley, P. Norman, K. Cruickshanks, A. Davis, et al. 2014. "Hearing in Middle Age: A Population Snapshot of 40- to 69-Year Olds in the United Kingdom." *Ear and Hearing* 35 (3): e44–e51. doi:10.1097/AUD.0000000000000010.

Denys, S., M. Hofmann, A. van Wieringen, and J. Wouters. 2019. "Improving the Efficiency of the Digit Triplet Test Using Digit Scoring with Variable Adaptive Step Sizes." *International Journal of Audiology* 58 (10): 670–677. doi:10.1080/14992027.2019.1622042.

DePaolis, R. A., C. P. Janota, and T. Frank. 1996. "Frequency Importance Functions for Words, Sentences, and Continuous Discourse." *Journal of Speech, Language, and Hearing Research* 39 (4): 714–723. doi:10.1044/jshr.3904.714.

Folmer, R. L., J. Vachhani, G. P. McMillan, C. Watson, G. R. Kidd, and M. P. Feeney. 2017. "Validation of a Computer-Administered Version of the Digits-in-Noise Test for Hearing Screening in the United States." *Journal of the American Academy of Audiology* 28 (2): 161–169. doi:10.3766/jaaa.16038.

Goman, A. M., and F. R. Lin. 2016. "Prevalence of Hearing Loss by Severity in the United States." *American Journal of Public Health* 106 (10): 1820–1822. doi:10.2105/AJPH.2016.303299.

Grant, K. W., and T. C. Walden. 2013. "Understanding Excessive SNR Loss in Hearing-Impaired Listeners." *Journal of the American Academy of Audiology* 24 (4): 258–273. doi:10.3766/jaaa.24.4.3.

Homans, N. C., R. M. Metselaar, J. G. Dingemanse, M. P. van der Schroeff, M. P. Brocaar, M. H. Wieringa, R. J. Baatenburg de Jong, et al. 2017. "Prevalence of Age-Related Hearing Loss, Including Sex Differences, in Older Adults in a Large Cohort Study." *The Laryngoscope* 127 (3): 725–730. doi:10.1002/lary.26150.

Hood, J. D. 1960. "The Principles and Practice of Bone Conduction Audiometry: A Review of the Present Position." *The Laryngoscope* 70: 1211–1228. doi:10.1288/00005537-196009000-00001.

Houtgast, T., and J. M. Festen. 2008. "On the Auditory and Cognitive Functions That May Explain an Individual's Elevation of the Speech Reception Threshold in Noise." *International Journal of Audiology* 47 (6): 287–295. doi:10.1080/14992020802127109.

Ikram, M. A., G. G. O. Brusselle, S. D. Murad, C. M. van Duijn, O. H. Franco, A. Goedegebure, C. C. W. Klaver, et al. 2017. "The Rotterdam Study: 2018 Update on Objectives, Design and Main Results." *European Journal of Epidemiology* 32 (9): 807–850. doi:10.1007/s10654-017-0321-4.

Jansen, S., H. Luts, P. Dejonckere, A. van Wieringen, and J. Wouters. 2013. "Efficient Hearing Screening in Noise-Exposed Listeners Using the Digit Triplet Test." *Ear and Hearing* 34 (6): 773–778. doi:10.1097/AUD.0b013e318297920b.

Jansen, S., H. Luts, K. C. Wagener, B. Frachet, and J. Wouters. 2010. "The French Digit Triplet Test: A Hearing Screening Tool for Speech Intelligibility in Noise." *International Journal of Audiology* 49 (5): 378–387. doi:10.3109/14992020903431272.

Koole, A., A. P. Nagtegaal, N. C. Homans, A. Hofman, R. J. B. de Jong, et al. 2016. "Using the Digits-in-Noise Test to Estimate Age-Related Hearing Loss." *Ear and Hearing* 37: 508–513. doi:10.1097/AUD.0000000000000282

Kramer, S. E., T. S. Kapteyn, J. M. Festen, and H. Tobi. 1996. "The Relationships Between Self-Reported Hearing Disability and Measures of Auditory Disability." *Audiology : official Organ of the International Society of Audiology* 35 (5): 277–287. doi:10.3109/00206099609071948.

Liu, X. 2012. "Classification Accuracy and Cut Point Selection." *Statistics in Medicine* 31 (23): 2676–2686. doi:10.1002/sim.4509.

Potgieter, J.-M., D. W. Swanepoel, H. C. Myburgh, and C. Smits. 2018a. "The South African English Smartphone Digits-in-Noise Hearing Test: Effect of Age, Hearing Loss, and Speaking Competence." *Ear and Hearing* 39 (4): 656–663. doi:10.1097/AUD.0000000000000522.

Potgieter, J.-M., D. W. Swanepoel, and C. Smits. 2018b. Evaluating a Smartphone Digits-in-Noise Test as Part of the Audiometric Test Battery. *The South African journal of communication disorders = Die Suid-Afrikaanse tydskrif vir Kommunikasieafwykings*. 65: a574.

Pronk, M., D. J. H. Deeg, J. M. Festen, J. W. Twisk, C. Smits, H. C. Comijs, S. E. Kramer, et al. 2013. "Decline in Older Persons' Ability to Recognize Speech in Noise: The Influence of Demographic, Health-Related, Environmental, and Cognitive Factors." *Ear and Hearing* 34 (6): 722–732. doi:10.1097/AUD.0b013e3182994eee.

Pronk, M., D. J. H. Deeg, and S. E. Kramer. 2018. "Explaining Discrepancies between the Digit Triplet Speech-in-Noise Test Score and Self-Reported Hearing Problems in Older Adults." *Journal of Speech, Language, and Hearing Research : JSLHR* 61 (4): 986–999. doi:10.1044/2018_JSLHR-H-17-0124.

Smits, C., S. T. Goverts, and J. M. Festen. 2013. "The Digits-in-Noise Test: Assessing Auditory Speech Recognition Abilities in Noise." *The Journal of the Acoustical Society of America* 133 (3): 1693–1706. doi:10.1121/1.4789933.

Smits, C., T. S. Kapteyn, and T. Houtgast. 2004. "Development and Validation of an Automatic Speech-in-Noise Screening Test by Telephone." *International Journal of Audiology* 43 (1): 15–28. doi:10.1080/14992020400050004.

Smits, C., P. Merkus, and T. Houtgast. 2006. "How We Do It: The Dutch Functional Hearing–Screening Tests by Telephone and Internet." *Clinical Otolaryngology* 31 (5): 436–440. doi:10.1111/j.1749-4486.2006.01195.x.

StataCorp. 2017. *Stata Statistical Software: Release 15*. College Station, TX: StataCorp LLC.

Stevens, G., S. Flaxman, E. Brunskill, M. Mascarenhas, C. D. Mathers, and M. Finucane, Global Burden of Disease Hearing Loss Expert Group. 2013. "Global and Regional Hearing Impairment Prevalence: An Analysis of 42 Studies in 29 Countries." *European Journal of Public Health* 23 (1): 146–152. doi:10.1093/eurpub/ckr176.

Taylor, B. 2003. "Speech-in-Noise Tests: How and Why to Include Them in Your Basic Test Battery." *The Hearing Journal* 56 (1): 40–42. doi:10.1097/01.HJ.0000293000.76300.ff.

Vercammen, C., T. Goossens, J. Wouters, and A. van Wieringen. 2018. "Digit Triplet Test Hearing Screening with Broadband and Low-Pass Filtered Noise in a Middle-Aged Population." *Ear and Hearing* 39: 825–828. doi:10.1097/AUD.0000000000000524

Vlaming, M. S. M. G., R. C. MacKinnon, M. Jansen, and D. R. Moore. 2014. "Automated Screening for High-Frequency Hearing Loss." *Ear and Hearing* 35 (6): 667–679. doi:10.1097/AUD.0000000000000073.

Watson, C. S., G. R. Kidd, J. D. Miller, C. Smits, and L. E. Humes. 2012. "Telephone Screening Tests for Functionally Impaired Hearing: Current Use in Seven Countries and Development of a US Version." *Journal of the American Academy of Audiology* 23 (10): 757–767. doi:10.3766/jaaa.23.10.2.

Wingfield, A., P. A. Tun, and S. L. McCoy. 2005. "Hearing Loss in Older Adulthood: What It is and How It Interacts with Cognitive Performance." *Current Directions in Psychological Science* 14 (3): 144–148. doi:10.1111/j.0963-7214.2005.00356.x.

Youden, W. J. 1950. "Index for Rating Diagnostic Tests." *Cancer* 3 (1): 32–35. doi:10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3.

Zokoll, M. A., K. C. Wagener, T. Brand, M. Buschermöhle, and B. Kollmeier. 2012. "Internationally Comparable Screening Tests for Listening in Noise in Several European Languages: The German Digit Triplet Test as an Optimization Prototype." *International Journal of Audiology* 51 (9): 697–707. doi:10.3109/14992027.2012.690078.