

A MultiCenter Analysis of Factors Associated with Hearing Outcome for 2,735 Adults with Cochlear Implants

Trends in Hearing
Volume 25: 1–17
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/23312165211037525
journals.sagepub.com/home/tia



Benjamin Goudey^{1,2}, Kerrie Plant³, Isabell Kiral¹,
Antonio Jimeno-Yepes¹, Annalisa Swan¹, Manoj Gambhir¹,
Andreas Büchner⁴, Eugen Kludt⁴ , Robert H. Eikelboom^{5,6,7} ,
Cathy Sucher^{5,6}, Rene H. Gifford⁸ , Riaan Rottier³ ,
and Hamideh Anjomshoa¹ 

Abstract

While the majority of cochlear implant recipients benefit from the device, it remains difficult to estimate the degree of benefit for a specific patient prior to implantation. Using data from 2,735 cochlear-implant recipients from across three clinics, the largest retrospective study of cochlear-implant outcomes to date, we investigate the association between 21 preoperative factors and speech recognition approximately one year after implantation and explore the consistency of their effects across the three constituent datasets. We provide evidence of 17 statistically significant associations, in either univariate or multivariate analysis, including confirmation of associations for several predictive factors, which have only been examined in prior smaller studies. Despite the large sample size, a multivariate analysis shows that the variance explained by our models remains modest across the datasets ($R^2 = 0.12\text{--}0.21$). Finally, we report a novel statistical interaction indicating that the duration of deafness in the implanted ear has a stronger impact on hearing outcome when considered relative to a candidate's age. Our multicenter study highlights several real-world complexities that impact the clinical translation of predictive factors for cochlear implantation outcome. We suggest several directions to overcome these challenges and further improve our ability to model patient outcomes with increased accuracy.

Keywords

cochlear-implant, predictive factor

Received 7 October 2020; Revised 5 July 2021; accepted 10 July 2021

Introduction

For decades, cochlear implants (CIs) have been an effective intervention to restore some hearing in individuals who have been impacted by significant permanent hearing loss (HL) (Eshraghi et al., 2012). Implantation is considered as an option when the auditory nerve is intact but hearing aids (HAs) are no longer able to compensate adequately for the loss of hearing. In general, clinical outcomes for implantees have improved substantially since the introduction of CIs (Hoppe et al., 2019) and it is recognized that the vast majority of individuals who are eligible for implantation showed improved hearing outcomes. Despite the well-documented success of CIs in restoring some hearing for the majority of patients, individual patient outcomes, and satisfaction vary with some individuals achieving strong improvements in hearing, while a minority show little or no improvement (Rubinstein et al., 1999; Gantz et al., 1993; Pisoni et al., 2017; Boisvert et al., 2020).

An expansive body of research aims to quantify differences in hearing performance and identify factors that

¹IBM Research Australia, Southbank, Victoria, Australia

²School of Computing and Information Systems, University of Melbourne, Parkville, Victoria, Australia

³Cochlear Limited, Sydney, New South Wales, Australia

⁴Medizinische Hochschule Hannover, Hannover, Niedersachsen, Germany

⁵Ear Science Institute Australia, Subiaco, Western Australia, Australia

⁶Ear Sciences Centre, The University of Western Australia, Nedlands, Western Australia, Australia

⁷Department of Speech Language Pathology and Audiology, University of Pretoria, South Africa

⁸Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN, USA

Corresponding author:

Hamideh Anjomshoa, IBM Research Australia, Southbank, Victoria, Australia.

Email: hamideh.anjomshoa@unimelb.edu.au



may account for the variation in observed performance (Blamey et al., 1996, 2013; Dowell et al., 2004; Lazard et al., 2012; Roditi et al., 2009; Shea III et al., 1990; Summerfield & Marshall, 1995; Waltzman et al., 1995). However, identification and interpretation of such predictive factors remain challenging (Pisoni et al., 2017). Several factors are associated with CI outcomes, including etiology of HL, duration of HL, duration of HA use prior to implantation, patient age at implantation, preoperative hearing scores, and percentage of stimulating electrodes (Zhao et al., 2020; Blamey et al., 2013, 1992). While the association with the hearing outcome is clear, the strength of evidence for many of these associations varies substantially across studies (Zhao et al., 2020). Similar trends are observed for multivariate analysis whereby the reported ability of predictive factors to explain variance in performance outcomes varies from as low as 10% (Blamey et al., 2013) to up to 31% for implantees with prelingual HL (Kraaijenga et al., 2016). Several reasons proposed to explain this variability include small patient cohorts (Zhao et al., 2020), differing performance criteria (Gaylor et al., 2013), and differences in cohort demographics (Boisvert et al., 2020; Leigh et al., 2016).

Larger multicenter studies or meta-analyses of the literature may provide a more accurate depiction of true effect size for predictive factors compared to smaller, single-center cohorts but such studies come with potential complications. The largest analysis of individual-level participant data can be seen in the landmark papers by Lazard et al. (2012) and Blamey et al. (2013), which examined numerous factors in 2,251 patient records across 15 clinics, the largest studies to date. However, these studies did not analyze differences in effect sizes or the impact of cohort differences between centers. Moreover, the chosen hearing outcome, a rank-based composite of heterogeneous hearing tests from across the different clinics may reduce the predictive power of the measured factors (Goldberg et al., 2014; Harrell & Frank, 2015). More recently, Zhao et al. (2020) conducted a meta-analysis of 1,095 participants from 13 different studies. They identified duration of HL, preimplantation pure tone average (PTA), preimplantation word recognition tests, and age at implantation as significantly associated with multiple performance outcomes across the different studies, albeit with substantial interstudy differences. This study was unable to conduct a metaregression due to limited data available across studies. Given the limited number of large multicenter studies, a recent review of the CI literature explicitly recommended larger meta-analyses from multiple datasets be conducted to improve the reliability of conclusions related to predictive factors (Boisvert et al., 2020).

To address some of the challenges of prior studies of predictive factors for cochlear implantation outcomes, we evaluated a cohort of 2,735 individual CI patients, the largest joint study of CI outcome to date. We focused on a single outcome metric, postoperative monosyllabic word

recognition scores (WRSs) at 12 months after implantation, with data provided by three different clinics, Vanderbilt University Medical Center (VUMC), Ear Science Institute Australia (ESIA), and Medizinische Hochschule Hannover (MHH). We investigated predictive factors pertaining to patient demographics, hearing-related measurements, clinical history, and etiology of HL and examined their effects within and across the clinics (with the latter subject to measurement availability across the three cohorts). Examining 21 predictive factors across three large centers, we found further refinement of several known associations and more definitive evidence for several predictive factors, which have only been examined in relatively smaller studies. Our study highlights the real-world complexities in understanding the relationship of predictive factors and cochlear implantation outcomes, including differences in cohort criteria, data collection, and definitions of predictive factors, that will impact the clinical translation of findings from this area, and we suggest several possibilities to mitigate the impact of these discrepancies.

Methods

Participating Clinics

This study was based on records from three different clinics: VUMC, ESIA, and MHH. Ethics approvals and data privacy protection practices were implemented. All patient data were deidentified and meet data compliance requirements for local patient data privacy laws and international law for General Data Protection Regulation. Each clinic implemented its own standard practice and preimplant test protocol for CI candidacy and post-implant evaluations.

Exclusion Criteria

This study focuses on the hearing performance of adults with a single implantation whose HL was within the range typically considered for cochlear implantation (moderate to profound HL). To ensure that the patient records met these criteria across the clinics, we removed any individual where: age at implantation was <18 years, postoperative word score was not recorded, a second CI was received sooner than 12 months after the first implant, implantation was conducted before 2003, or where data entered were spurious (e.g., incorrect age and missing surgery date). Additionally, we excluded a small number of records with a preoperative four frequency PTA of <60 dB HL or a WRS <50% in the implanted ear as the HL of these individuals was outside of the normal range considered for cochlear implantation. Applying the exclusion criteria, 2,735 patient records out of an initial cohort of 6,500 remained. Table 1 presents a breakdown of individual records per dataset and patient demographics.

Table 1. Dataset Demographics.

	VUMC	ESIA	MHH	Combined datasets
Number of patients	491	293	1,951	2,735
Age at implantation (years)	64.0 (15.3)	62.8 (15.0)	55.5 (17.4)	57.8 (17.2)
WRS(CI)	43.8 (23.1)	42.1 (23.1)	51.2 (28.9)	48.9 (27.6)
WRS(HA)	8.3 (12.1)	7.6 (11.5)	4.2 (9.6)	6.0 (10.9)
YRS – D _i	26.5 (17.9)	29.8 (18.7)	8.9 (13.2)	14.9 (17.414)
PTA _i	98.7 (19.2)	116.8 (14.0)	98.8 (17.2)	100.7 (18.2)
PTA _c	85.2 (25.8)	87.5 (28.1)	77.8 (28.4)	80.2 (28.2)
N. female (%)	220 (44.8%)	—	1,072 (54.9%)	1,292 (52.9%)
N. prelingual HL (%)	38 (7.7%)	47 (16.0%)	161 (8.3%)	246 (9.0%)

All entries show mean (and standard deviation in brackets), except N. female and N. prelingual HL, where we report the number and percentage of participants. ^aVUMC = Vanderbilt University Medical Center; ESIA = Ear Science Institute Australia; MHH = Medizinische Hochschule Hannover; WRS(CI) = 12-month post-operative word recognition score percentage correct; WRS(HA) = preoperative word recognition score percentage correct with hearing aid; YRS – D_i = years of severe to profound hearing loss in the implanted ear; PTA_{i/c} = pure tone average in implanted/contralateral ears.

Measure of Performance

The most common method to evaluate a patient’s progress with a HA or CI is to conduct monosyllabic word recognition and sentence recognition tests. The present study used different monosyllabic WRSs at the different clinics; Consonant-Nucleus-Consonant (Peterson & Lehiste, 1962) scores for data acquired in Australia (ESIA) and the United States of America (VUMC), and the Freiburg monosyllabic test scores for data collected in Germany (MHH) (Hahlbrock, 1953, 1960). The tests are not identical, with differences in the number of words, the words themselves, and how common the words are in the language. However, they do have a similar structure and were conducted in comparable settings. If predictive factors are associated with WRS, we would expect to see comparable results across the clinics. While we recognize that differences in the word recognition tests across clinics may impact our analysis, a formal comparison of the impact was beyond the scope of this work.

Across all clinics, WRS tests for pre- and postoperative evaluation of the implanted ear were conducted in free-field at conversational level with a HA, denoted WRS(HA), or with a CI, denoted WRS(CI), while the contralateral ear was masked appropriately. The presentation level of the monosyllabic word test ranged from 60 to 65 dB sound pressure level (SPL) root mean squared (RMS), differing across a site with respect to the calibration procedure applied. This level was consistent across pre- and postimplant visits in each clinic. Additionally, the MHH dataset included records of a preoperative test conducted with headphones at a range of loudness levels. The maximum score under these conditions was recorded as PB_{max} (Hoppe et al., 2019).

In our study, we used the postoperative WRS (evaluated on whole words, not on the phoneme level) that was acquired

closest to 12 months after implantation, with the time after surgery varying from 6 to 24 months.

Predictive Factors Studied

In addition to demographic and hearing-related influences, we investigated factors relating to the impact of a patient’s clinical history and their etiology. These clinical factors were collected across the three clinics independently and, as such, some differ in their exact definition across the three sites, reflecting the current reality of clinical data collection in this space. While many factors were collected by all clinics, we highlight when a factor was only available in a subset of datasets, as this limits the ability to evaluate performance across all individuals.

Patient Demographics

Age at implantation: The age of the patient at the time of receiving a CI.

Native speaker: Hearing tests were conducted in English (VUMC and ESIA) and German (MHH). This binary field indicates whether an implantee self-identifies as a native (or bilingual) speaker in the test language. Since this field was only available in the MHH dataset, a value of 1 means that German is the implantee’s native language. This has previously been shown to impact speech perception of cochlear implantees (Van Wijngaarden et al., 2002; Kilman et al., 2015).

Gender: Gender recorded as either female or male was available in the VUMC and MHH datasets, with female coded as 1 in any regression analysis. While this field was recorded at ESIA, it was unavailable for this study.

Hearing-Related Measurements

Preoperative WRS: The WRS recorded during the most recent visit prior to implantation, as defined in the “Measure of Performance” section and measured with a HA. This score was recorded for the to-be-implanted ear ($WRS(HA)_i$), the contralateral ear ($WRS(HA)_c$), and both ears ($WRS(HA)_b$). $WRS(HA)_c$ was recorded in VUMC only.

Preoperative PTA: The PTA is the mean hearing threshold across 0.5, 1, 2, and 4 kHz frequencies. During audiogram tests, the patient was unaided with no HA and the assessment was under air conduction. This audiometric score was recorded for the to-be-implanted ear and the contralateral ear, denoted PTA_i and PTA_c , respectively. In addition, we introduced the variable $PTA_{sat,i}$, which is 1 if any of the frequencies used to calculate PTA_i reached saturation, i.e., the individual’s “true” PTA was >110 dB and hence exceeded the capacity of the audiometer.

PBmax: The maximum WRS measured in the to-be-implanted ear via headphones while the presentation level of the stimulus was varied. This was only available in the MHH dataset.

Clinical History

Prelingual HL: A binary variable where 1 indicates an individual with diagnosed HL at the age of two years or younger (VUMC and ESIA), or at the age of four years or younger (MHH).

Course of HL: This describes whether the patient had sudden (acute) or progressive HL, as defined by a clinician. Available in MHH only, this binary variable is coded with progressive loss as 1. Note: this measurement is different from the etiology of “sudden hearing loss as many individuals without this etiology may describe their hearing loss as sudden and vice versa.”

Years of HL: The time span between an individual being identified with HL of any degree and implantation date. This information is based on the patient’s response to several questions. This value was recorded for the to-be-implanted ear ($YRS - HL_i$) and the contralateral ear ($YRS - HL_c$) in the MHH dataset.

Side of the implant: This binary variable refers to the side of the ear that received a CI, with left coded as 1.

Years of severe to profound deafness: The time span between onset of severe to profound HL and implantation is recorded in all three datasets and is denoted for the to-be-implanted and contralateral ears as $YRS - D_i$ and $YRS - D_c$, respectively. This information is based on a series of questions (e.g. when did the patient stopped using the phone?), all of which attempt to identify a point in time when the functional hearing reached a level that would indicate a severe/profound HL.

However, the exact questions and criteria used to translate the patient’s responses into a single estimate of duration differ across clinics. This reflects the way in which this measurement is collected within current clinical practice. As such, we expected to see differences in this measure across the different clinics and a key outcome was the extent to which we observed variation.

Years of HA use: Available in ESIA and VUMC, the duration of HA use prior to CI implantation is recorded for the to-be-implanted and contralateral ears, denoted $YRS - HA_i$ and $YRS - HA_c$, respectively. As HA usage is not broken down by year (which would indicate if usage was inconsistent), we assume that HA use was consistent for the entire reported duration.

Etiology. The underlying reason for the patient’s HL, if known, was commonly recorded as free text by clinics. Although there is no standardized methodology for categorizing etiological data, an approach was followed to combine the data into groups after consulting with subject matter experts.

The grouping of conditions led to a 13-category variable, with the following classes: noise-induced, otosclerosis, Meniere’s disease, congenital syndrome, childhood or congenital illness, genetics, (chronic) otitis media & infections, trauma, sudden HL, ototoxicity & streptomycin, meningitis, others (containing all recorded etiologies that did not fit into a category with sufficient values to be meaningful or were recorded as “other” in the original datasets), and unknown (if etiology was recorded as unknown in the original datasets or was missing). Since otosclerosis and meningitis have been described as predictive factors in the past literature (Blamey et al., 2013), they are not grouped with other etiologies.

In the MHH dataset, the course of HL was captured separately to etiology. For example, patients with meningitis can exhibit acute or progressive HL, as do patients with an etiology that falls in the “sudden HL” group. The “sudden HL” etiology was used to describe patients whose cause of HL is unknown but occurs in one or several acute episodes.

Imputation of Missing Hearing Assessments

Audiograms were collected at the last visit before implantation to investigate the degree of HL and were used to compute PTA. In some circumstances (i.e., patients reaching the limit of their hearing at low frequencies, or patients not being able to hear at maximum loudness and providing no response), the resulting audiogram data may not include all frequencies. If the data indicate that the patient reached the limit or had no response at that particular frequency, no response was recorded at max limits, and the missing values were imputed to 125 dB HL.

Similar to audiogram frequencies, all datasets contained missing values for the preoperative $WRS(HA)$ of the

to-be-implanted ear. Since missing WRS values again may indicate that the assessed patient may have had more severe HL than the average tested participant, removing these patients would bias our analyses. Hence, we imputed any missing preoperative WRS(HA) values with 0 if all measured PTA values were equal to or above 110 dB HL, mimicking the event in which a patient does not provide any correct answers during a word recognition task. Our rule-based imputation for audiograms and WRS(HA) affected nearly 25% of the records across the three datasets.

Statistical Analysis

Univariate distributions of predictive factors provide insight into the make-up of the different cohorts. Age at implantation, years of HL, WRS(HA), and WRS(CI) were estimated using kernel density estimation as provided in the Seaborn Python package (Waskom, 2021). Differences in the distributions across the clinics were quantified using a two-sided Kolmogorov–Smirnov test.

Within each clinic, associations between predictive factors and WRS(CI) were evaluated using linear regression models, either with the predictive factor alone (univariate analysis) or with a combination of predictive factors (multivariate). The resulting coefficient (β) for each term in the model indicates how one unit change in the predictive factor of interest will affect WRS(CI), assuming all other variables in the model (if present) are fixed. To conduct a meta-analysis across the three cohorts, we used a mixed linear modeling approach with a random intercept for each clinic.

Interaction analysis was evaluated by fitting models with and without an interaction term and determining whether the interaction significantly improved the model's fit via a likelihood ratio test.

Significance for each predictive factor was determined using a two-sided Wald test. Significant predictive factors were those that had a p -value < 0.05 after applying Benjamini–Hochberg correction (Benjamini & Hochberg, 1995).

Forest Plots

Given a large number of variables being considered across three separate datasets and the combined dataset, we visualized the effect of predictive factors using forest plots, which are commonly reported in the epidemiology literature. For a given set of predictive factors, whether uni- or multivariate, we created a forest plot showing effect size, confidence interval, and associated p -value. Each plot shows a different predictive factor along the y -axis, with colored bars indicating the four different datasets (three clinics and a meta-analysis). Each point corresponds to the coefficient for a predictive factor in a given dataset

with bars indicating the 95% confidence interval of this coefficient. Asterisks above each bar indicate significance after Benjamini–Hochberg correction. We also report the number of available observations as a part of the predictive factor name on the x -axis, with the following order: VUMC, ESIA, MHH, and the combination dataset. If no observations were available for a given predictive factor in a dataset, we replace the number of observations with a dash. The y -axis, showing effect size, uses a pseudo-log scale, which allows us to show a range of effect sizes on the same plot.

Results

Distribution of Data Across Clinics

As summarized in Table 1, data between clinics differed in several ways, including demographic factors and the clinical history of patients. To explore the similarity of these cohorts, we analyzed the distribution of patients' age and hearing performance before and after implantation. Figure 1 displays the density of age, years of deafness in implanted ear, PTA in both ears, WRS(HA) scores, and WRS(CI) scores for each dataset, visualizing the relative spread of data for each of the clinics.

The distributions of all fields shown in Figure 1 differ significantly in MHH compared to VUMC and ESIA ($p < 1 \times 0.001$ for MHH vs VUMC or ESIA for all six features). In particular, there is a strong difference in years of severe to profound deafness. There were many individuals in the MHH dataset who were implanted within a few years of qualifying as severe or profoundly deaf, while the other clinics tended to have the majority of patients meeting these criteria for over 15 years. Years of severe to profound deafness and PTA, both in the implanted ear, were also significantly different between ESIA and VUMC ($p = 5 \times 10^{-4}$ and 1×10^{-16} , respectively). PTA in the implanted ear in ESIA shows a strong peak at 120 dB HL, indicative of a large number of patients with no hearing at the maximum limits of the audiometer. Remaining comparisons were not significant (Online Supplemental Table 1).

While WRS are scaled to be between 0 and 100, the tests included 50, 25, and 20 monosyllabic word items for VUMC, ESIA, and MHH, respectively. The difference in the number of words or, in particular, word lists being used may account for some of the variability in testing outcome, in particular, with differences between VUMC and the remaining clinics, with the smaller word lists likely to result in noisier WRS due to increased sampling error (Thornton & Raffin, 1978). These differences between clinics highlight some of the underlying properties that may affect the interpretation of predictive factor associations across the clinics.

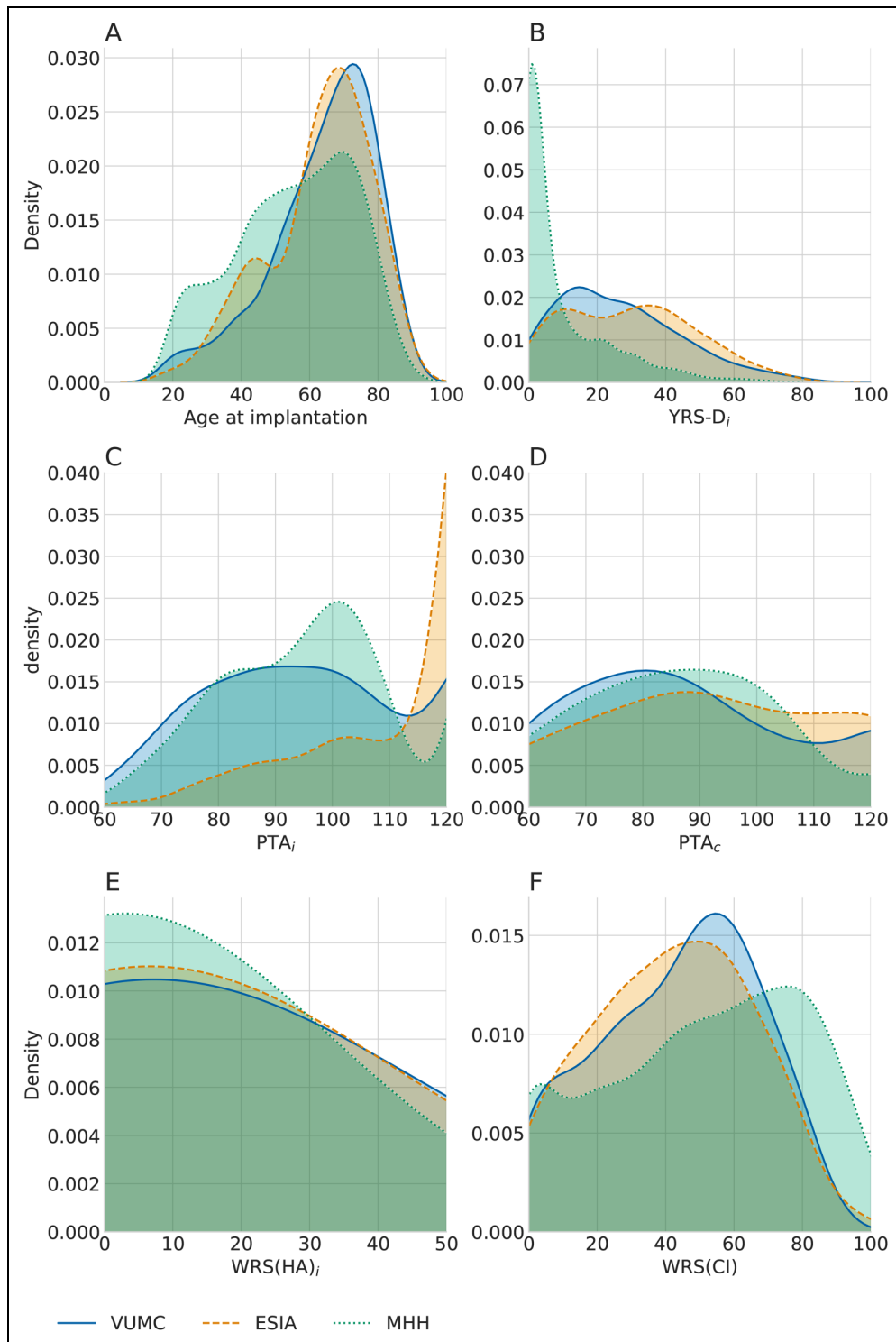


Figure I. Differences in key predictive factors across clinics. (A) Distribution of age, (B) distribution of years of severe to profound hearing loss, (C) distribution of pure tone average (PTA) in the implanted ear, (D) distribution of PTA in contralateral ear, (E) distribution of preoperative WRS(HA)_i, and (F) distribution of postoperative WRS(CI) of the implanted ear. Density plots indicate the estimated probability density function of a measurement, highlighting the spread of values across patients in each of the clinics. The area under any section of the curve indicating the probability of an individual having a value within that range.

Analysis of Univariate Predictive Factors

We initially conducted a univariate analysis, considering predictive factors of patient demographics, hearing-related measurements, and clinical history. The resulting effect sizes, confidence intervals, and significance levels are shown in Figure 2, with a detailed description of forest plots in Figure and further details given in the Online Supplemental Table 2.

Overall, we found significant associations for 16 of the 21 common predictive factors across the different datasets. The most significant associations were prelingual HL ($p = 7.1 \times 10^{-32}$) in MHH, years of HA use in the implanted ear in VUMC ($p = 1.0 \times 10^{-6}$) and PTA in the contralateral ear for ESIA ($p = 9.6 \times 10^{-3}$). The latter is the only significant predictive factor observed in the ESIA dataset, likely due to the limited sample size. Many predictive factors that have been reported in the literature were found to be significant in our analysis in either an individual clinic or in the combined meta-analysis. The following sections report findings for each factor broken down by category.

Patient Demographics

Age at implantation (available in all datasets): Results were mixed for the relationship between the age of the patient at implantation and their post-operative WRS. Lower age is associated with better WRS(CI) in VUMC and ESIA ($\beta = -0.18$, $p = 0.008$ and $\beta = -0.12$, $p = 0.17$, respectively), but higher age is associated with better WRS(CI) at MHH and the combined analysis ($\beta = 0.17$, $p = 6 \times 10^{-6}$ and $\beta = 0.9$, $p = 0.003$, respectively). Except for ESIA, all associations are significant. Further interpretation of these results using multivariate analysis in the “Multivariate Analysis” section indicates this may be due to differences related to varying etiologies in different age groups across clinics.

Native speaker (available only in MHH): We observe that native German speakers have a significant advantage, with an expected 19.7 points increase in WRS(CI) compared to nonnative German speakers ($p = 6.1 \times 10^{-14}$). This is in line with previous studies of hearing perception by individuals with and without CIs (Ji et al., 2014).

Gender (available in VUMC and MHH): In the VUMC and MHH datasets, gender shows no significant difference ($p = 0.78$ and $p = 0.35$, respectively), concordant with previous studies (Lazard et al., 2012).

Hearing-Related Measurements

Preoperative WRS(HA) (available in all datasets): WRS(HA) in the implanted ear (WRS(HA)_i) is significantly associated with better WRS(CI) in the combined, MHH and VUMC datasets ($p = 6 \times 10^{-6}$, 4×10^{-4} , 5×10^{-4} ,

respectively) with similar effect sizes observed across these three analyses ($\beta = -0.28$, -0.37 and -0.33 , respectively). In ESIA, the effect size is weak and the result is not significant ($\beta = 0.06$, $p = 0.61$). Similar results can be observed for the bilateral case, WRS(HA)_b , with a similar or weaker effect size. WRS(HA)_c was only available in VUMC and shows a positive effect that is not significant after multiple testing correction ($\beta = 0.11$, $p = 0.032$).

PTA (available in all datasets): Having a poorer (i.e. larger)

PTA was significantly associated with worse WRS(CI) for both ears, except in ESIA. Stronger and more significant effects are observed for PTA in the implanted ear compared to the contralateral ear ($\beta = -0.23$, $p = 3 \times 10^{-14}$ vs $\beta = -0.05$, $p = 0.005$), respectively, in the combined analysis. This result is consistent with the expectation that poorer starting conditions in the ear that is to be implanted will negatively affect the postoperative performance. The dichotomized version of PTA saturated also shows a strong association (e.g. $\beta = -5.76$, $p = 6 \times 10^{-7}$) for the implanted ear in the combined analysis. While the effect size here appears higher than that of the continuous-valued PTA_i , these cannot be compared because that PTA_i has a high range of values, while $\text{PTA}_{\text{sat},i}$ can only be 0 or 1. However, we note that $\text{PTA}_{\text{sat},i}$ is less significant than PTA_i .

PB_{max} (available only in MHH): PB_{max} exhibits strong association with WRS(CI) ($\beta = 0.19$, $p = 4.3 \times 10^{-8}$) in the MHH dataset. Given recent comparisons between PB_{max} and WRS(HA) (Hoppe et al., 2019), we further examine 356 patients that had both values measured to determine whether differences between the two tests could be observed with respect to their influence on postoperative WRS. Across these individuals, Online Supplemental Table 3 shows that WRS(HA) has a marginally higher effect size ($\beta = 0.16$ vs $.14$ for WRS(HA) and PB_{max}, respectively) but with a higher standard deviation and hence a less significant association with WRS(CI) than PB_{max} ($p = 0.19$ vs $.01$ for WRS(HA) and PB_{max}, respectively).

Clinical History

Prelingual HL (available in all datasets): Out of all factors relating to the clinical history, patients with prelingualHL have the largest association with poorer outcome. This association is significant except for in ESIA ($p = 7.4 \times 10^{-5}$, 8×10^{-32} , 4×10^{-33} for VUMC, MHH, and combined analysis, respectively), with prelingual HL predicting a 15 and 28 point lower postoperative word score for VUMC and MHH, respectively.

Years of HL (available in MHH): This factor measures the duration for which a patient experienced any HL before they received an implant. It is negatively associated with

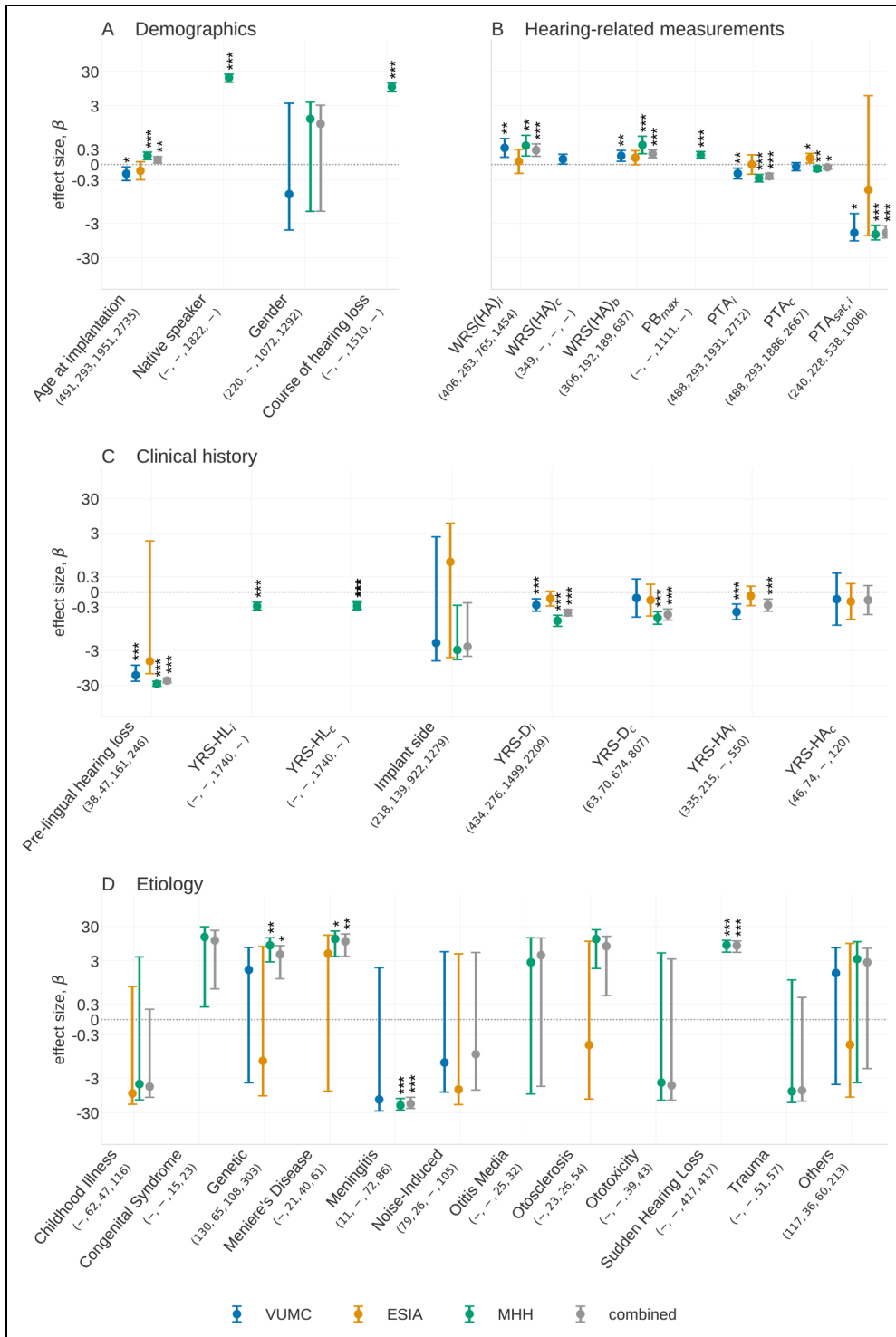


Figure 2. Univariate predictive factors showing (A) demographic factors, (B) hearing-related measurements, (C) clinical history, and (D) etiology groups. The number of individuals in each category in each dataset is shown beneath the x-axis label of each predictive factor, with numbers corresponding to VUMC, ESIA, MHH, and the combined cohort, respectively. A dash indicates no individuals carry a predictive factor in a given dataset. In (d), the forest plot, described in the “Forest Plots” section is showing the associations of the etiology categories compared to the effect when the etiology of hearing loss is unknown. All etiology classes with <10 observations have been removed from this plot but may still contribute to the combined analysis. Significance levels are indicated above bars as * $p < .05$, ** $p < .01$, and *** $p < .001$. VUMC = Vanderbilt University Medical Center; ESIA = Ear Science Institute Australia; MHH = Medizinische Hochschule Hannover.

postoperative hearing performance for the implanted ear (β : -28 , $p = 3.9 \times 10^{-12}$). Similar though marginally weaker effects are observed for the contralateral ear with (β : -26 , $p = 1.3 \times 10^{-9}$).

Side of the implant (available in all datasets): We found no significant relationship between implant side and hearing outcomes. However, a negative association in the MHH and combined analyses ($\beta = -2.8$ and -2.3 , respectively) is close to significance after multiple testing ($p = 0.03$ in both analyses).

Years of severe to profound deafness (available in all datasets): Greater years of deafness (i.e., more time between the onset of severe to profound deafness and implantation) in the implanted ear leads to lower WRS(CI) in VUMC, MHH, and the combined datasets (β : -0.26 , -0.56 and -0.41 , respectively), with these effects being significant ($p = 4.3 \times 10^{-5}$, 1.6×10^{-23} , and 5.9×10^{-34} , respectively). Years of deafness in the contralateral ear is significant in MHH and combined datasets ($p = 2.6 \times 10^{-15}$ and 9.4×10^{-15} , respectively), but not in ESIA and VUMC datasets ($p = 0.32$ and 0.54). It is worth reiterating that these results will be influenced by differences in the definition of what constitutes severe HL among the clinics despite the relatively consistent effects. Course of HL (available in MHH): Individuals who lost their hearing progressively show significantly higher WRS(CI) than patients with acute HL ($\beta = 10.67$, $p = 2.3 \times 10^{-11}$).

Years of HA use (available in ESIA and VUMC): Years of HA use in the implanted ear shows a significant negative association ($\beta = -0.26$, $p = 2.8 \times 10^{-5}$) in the combined meta-analysis of ESIA and VUMC. Years of HA use in the contralateral ear showed no association ($\beta = -0.16$, $p = 0.28$) but was limited by the smaller sample size ($n = 120$).

Etiology. Etiology data are available in all datasets. Etiology coefficients indicate the expected change in WRS(CI) for that etiology compared to the change in WRS(CI) when the cause of HL is unknown. Overall, we see a lot of variability across the datasets in terms of the strength and spread of the associations, likely because the distribution of etiology classes differs dramatically across the datasets and is likely to be different within the “unknown” category (Online Supplemental Table 4). Nevertheless, Figure 2 reveals that some etiology classes show a significant and consistent effect:

Genetic: Individuals with genetic HL have mixed results across the clinics, with the MHH and the combined analysis showing significant better postoperative outcome in WRS(CI) ($\beta = 8.4$, $p = 0.004$ and $\beta = 4.6$, $p = 0.015$, respectively), whereas VUMC and ESIA show no significant association ($p = 0.57$ and $.83$, respectively). A number of conditions are grouped under this category and it is likely that the distribution of cases differs

across the datasets.

Meningitis: Meningitis shows a significant negative association (poorer outcome) in MHH and the combined analysis ($\beta = -18.3$, $p = 1 \times 10^{-7}$ and $\beta = -16.6$, $p = 4 \times 10^{-8}$, respectively) and a nonsignificant but similar effect size in VUMC ($\beta = -12.4$, $p = 0.09$). The analysis of VUMC was hampered by low number of observations ($n = 11$). The strong negative effect observed is consistent with the prior literature (Kraaijenga et al., 2016).

Meniere’s disease: An etiology of Meniere’s disease is positively associated with outcomes in the combined analysis ($\beta = 11.0$, $p = 0.002$) compared to an unknown etiology. This association is primarily driven by the significant effects in MHH with similar though substantially weaker trends seen over a smaller population in the ESIA dataset.

Sudden HL: Individuals with sudden HL show a significant positive association in the combined analysis ($\beta = 8.27$, $p = 9 \times 10^{-8}$) largely driven by observations in the MHH dataset.

Otosclerosis: Otosclerosis is almost significant in the MHH dataset ($\beta = 12.8$, $p = 0.023$ which does not meet threshold for significance for multiple testing correction). While ESIA has a similar number of observations, the wide confidence intervals indicate no evidence of an association.

Multivariate Analysis

While the previous section examined the relationship of individual predictive factors with WRS(CI), it did not capture which measurements provide independent information, and how the relationship with WRS(CI) changes in the presence of other variables. To explore this further, we repeated the analysis by analyzing all variables simultaneously. As predictive factors in this analysis must be measured in all observations, the number of individuals and predictive factors in this multivariate analysis was reduced compared to the univariate analysis.

Results of the multivariate analysis are shown in Figure 3, with further details in Online Supplemental Table 5. The combined analysis shows having a higher PTA_i ($p = 8 \times 10^{-3}$) or greater years of deafness prior to implantation ($p = 1.5 \times 10^{-17}$) or having an implantation on the left ear ($p = 0.03$) all lead to worse outcomes. Specific etiologies of hearing remain important for explaining hearing performance. While each etiology is significant in at least one dataset, we found genetics, Meniere’s disease, meningitis, otosclerosis, sudden HL and the “others” category are significantly associated with outcome, all with better WRS(CI) except meningitis in the combined analysis.

In addition, we found associations for factors for which the data only allowed for the inclusion of a subset of datasets. Years of deafness is significant in VUMC, MHH, and the

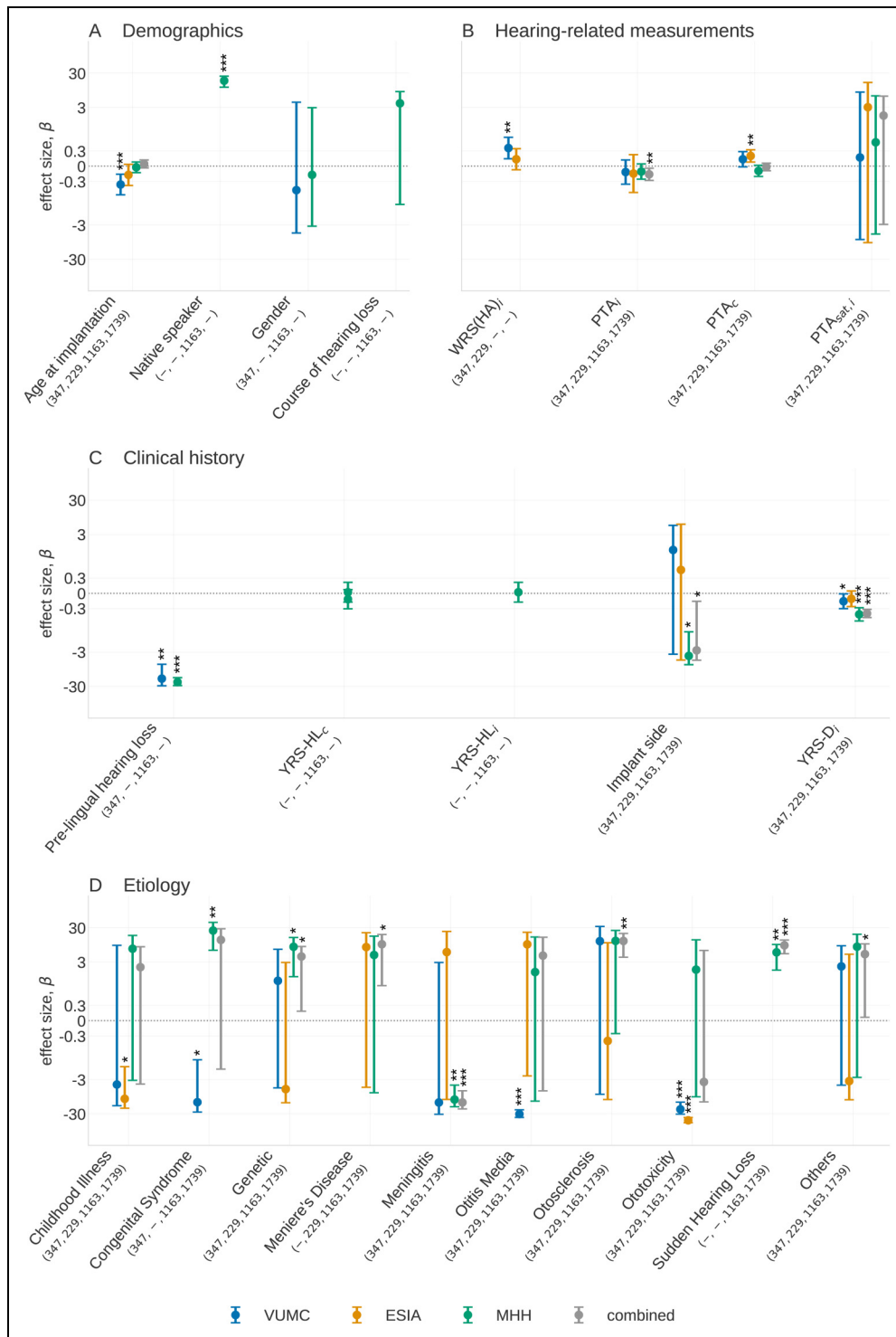


Figure 3. Multivariate predictive factors for three datasets and the combined dataset. The number of individuals in each category in each dataset is shown beneath the x-axis label of each predictive factor, with numbers corresponding to VUMC, ESIA, MHH, and the combined cohort, respectively. A dash indicates no individuals carry a predictive factor in a given dataset. In (D), forest plot, description in the “Forest Plots” section, is showing the associations of the etiology categories compared to the effect when the etiology of hearing loss is unknown. All etiology classes with <10 observations have been removed from this plot but may still contribute to the combined analysis. Significance levels are indicated above bars as * $p < .05$, ** $p < .01$, and *** $p < .001$. (A) Demographics, (B) hearing-related measurements, (C) clinical history, and (D) etiology. VUMC = Vanderbilt University Medical Center; ESIA = Ear Science Institute Australia; MHH = Medizinische Hochschule Hannover.

Table 2. R^2 and R^2_{adj} Values for Each Dataset for Regression Analysis using Multiple Features in Each Cohort. We Report Number of Individuals Included in Each Column.

	VUMC $n = 347$	ESIA $n = 229$	MHH $n = 1, 163$	Combined $n = 1, 739$
R^2	0.18	0.13	0.21	0.12
R^2_{adj}	0.12	0.06	0.20	0.12

VUMC = Vanderbilt University Medical Center; ESIA = Ear Science Institute Australia; MHH = Medizinische Hochschule Hannover.

combined analysis ($p = 0.03$, 7.2×10^{-10} and 1.5×10^{-17} , respectively). Whether the implant recipient was native in the test language ($p = 3.6 \times 10^{-8}$) is significant in MHH, and whether HL occurred prelingually is found to be a significant predictive factor in VUMC ($p = 1.7 \times 10^{-3}$), MHH ($p = 6.9 \times 10^{-14}$), and combined (1.5×10^{-17}). WRS(HA) is significant in VUMC ($p = 1 \times 10^{-3}$) and was not present in MHH or the combined analysis. In terms of etiology, ototoxicity is significant in both ESIA and VUMC datasets and the congenital syndrome is significant in VUMC and ESIA but has an opposite effect. Sudden HL, meningitis, and genetics are significant in both MHH and combined datasets.

Although most of these associations are in line with those found in the univariate analysis and provide further evidence of their relevance, we notice some interesting differences. Patient age is no longer a significant factor in MHH or the combined analysis. Across the VUMC and ESIA datasets, we observe a moderate negative impact of age. The different results between clinics are likely due to the fact that both VUMC and ESIA have a relatively higher representation of older versus younger recipients while MHH has a wider spread of ages. Similarly, $PTA_{\text{sat},i}$ and the course of HL lose significance when analyzed alongside other factors. This is likely due to the strong correlation between $PTA_{\text{sat},i}$ and PTA_i and between the course of HL and certain etiologies.

A multivariate analysis also reveals the significant association of the implant side in the MHH dataset ($\beta = -3.85$, $p = 0.01$) and the combined analysis ($\beta = -2.65$, $p = 0.03$), where improved outcomes are associated with the implantation of the right ear. Given this factor became significant in the multivariate analysis, it may indicate that the limited association, when considered univariately, is due to confounding by other factors. Similar associations have been shown previously in studies of pre- and postlingual children, but sample sizes in these studies were small and effect sizes were more modest than the one shown in this analysis (Kraaijenga et al., 2017; Mohammed & Sarwat, 2014).

Variance Explained

The multivariate analysis in the previous section highlights the relationship between predictive risk factors with the postoperative outcome, conditioned on all other available factors.

We used these same models to examine how much of the variability of WRS(CI) can be explained using the combination of risk factors.

Table 2 shows that the proportion of variance explained across all clinics and the combined analysis, with some variability expected due to the differences in both sample sizes and available features as per the multivariate analysis in the previous section. The reported R^2 is moderate in each dataset. The highest R^2 of 0.21 was observed in the MHH dataset. The smaller R^2 for the other clinics is likely related to the reduced number of available clinical factors and the more limited sample size. The combined analysis is further restricted in terms of the number of features, with critical features such as prelingual HL and WRS(HA) unavailable across the three cohorts, resulting in a low R^2 of 0.12.

To account for potential overfitting in smaller datasets with larger feature sizes, we also reported R^2_{adj} (Ostertagová, 2012). We observed a substantial drop in performance for VUMC and ESIA (to 0.12 and 0.06, respectively), highlighting the differences in the number of factors in the multivariate model relative to the sample size. In contrast, we see little difference for MHH given the significantly larger sample size. While R^2_{adj} is not commonly found in the literature related to cochlear implantation predictive factors, we believe it provides an important indication of whether the variance explained may be due to the overfitting of the data, even if the values themselves are less interpretable.

These reported values are comparable with or higher than those presented in previous studies in Blamey et al. (1996) ($R^2 = 0.21$), Blamey et al. (2013) ($R^2 = 0.10$), and Lazard et al. (2012) ($R^2 = 0.22$). Interestingly, many of the features in the previous studies were not used in this multivariate analysis as they were unavailable in the datasets used. Thus, we expect the explained variance would increase if the measurements that appear to be significant in one or more datasets were more complete.

Interaction Analysis

There has been little investigation to date whether there are significant nonadditive effects between established predictive factors. To explore this, we conducted an interaction analysis

among all major predictive factors in the dataset, focusing on those predictive factors that were shown to be significant in the univariate analysis. This results in a total of 72 interactions, with several more being excluded if the feature combinations were redundant (i.e., interactions between PTA on the contralateral and implanted ear, or PTA and PTA_{sat}).

The most significant finding is a strong synergistic effect between years of deafness in the contralateral ear and patient age at implantation ($p = 5.73 \times 10^{-5}$ in the combined dataset), shown in Figure 4. The plot indicates that WRS(CI) is substantially worse for young individuals with a long history of deafness, seen by the strong slope in the line for those with higher (68th percentile) years of deafness. In contrast, while the duration of deafness has only a moderate impact on individuals who are older when they receive an implant, exemplified in Figure 4 by the limited change for those with fewer (34th percentile) years of deafness as age is varied. As such, this finding indicates that years of deafness matter more when considered relative to a candidate's age.

Similar findings were seen for years of deafness in the implanted ear ($p = 1.03 \times 10^{-3}$ in the combined dataset). Some of these interactions were replicated in ESIA and MHH alone, with little effect shown in VUMC (Online

Supplemental Table 6), which may be influenced by the strong differences in the distribution of years of severe to profound deafness across the clinics. Moreover, a multivariate regression analysis using all available observations in the MHH dataset demonstrated that the interaction term provides significant contributions independent of the other main effects ($p = 0.006$). We also found that the interaction in MHH remains significant even if individuals with prelingual HL were removed from the analysis ($p = 1.3 \times 10^{-6}$).

Informed by the interaction analysis, we repeated the multivariate analysis in the previous sections to determine whether the observed interaction increases the variance explained by WRS(CI). As both years of deafness in the implanted or contralateral ear are not present in all datasets, we could not evaluate the impact on the combined cohort. Focusing instead on the MHH dataset, we found that the R^2 of the model was increased from 0.21 to 0.22 when the years of deafness in the implanted ear and age at implantation was present (years of deafness in the contralateral ear being absent from the multivariate regression in the "Multivariate Analysis" section). While this improvement is not significant ($p = 0.11$) it does provide yet further evidence that the interaction may be useful for prediction, and the effect may be

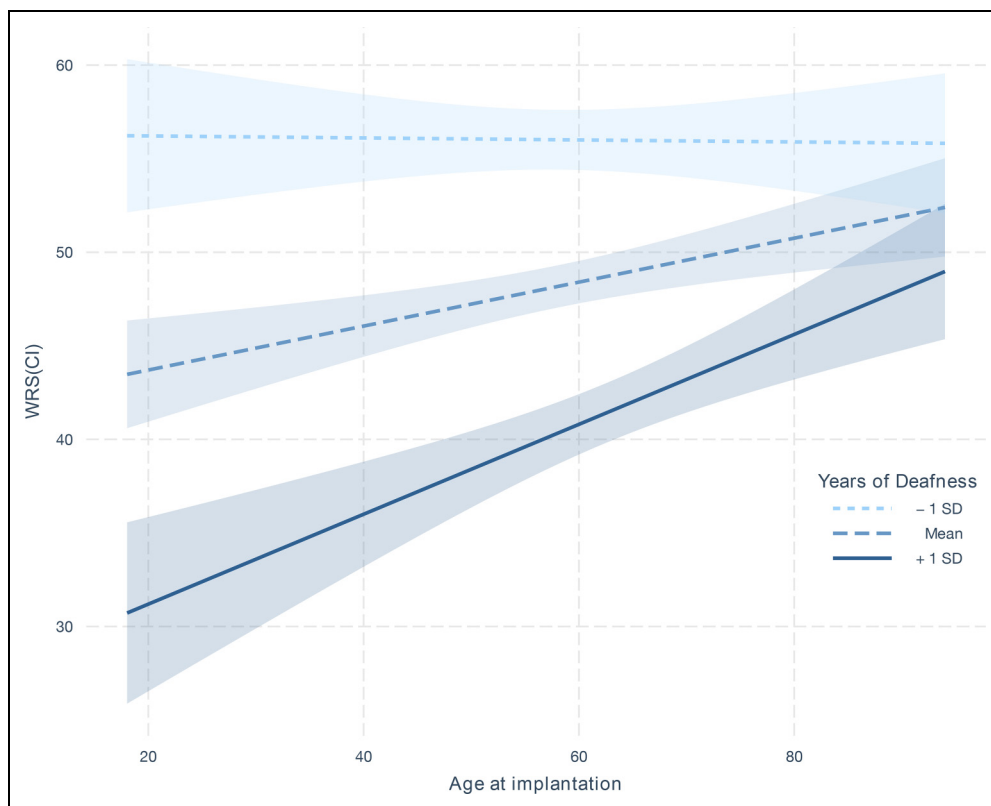


Figure 4. Interaction effect of age of implantation on WRS(CI) as a function of years of deafness. The three lines show the estimated linear relationship between age of implantation on WRS(CI) when years of deafness is set to the 34th, 50th, and 68th percentile (i.e., one standard deviation (SD) below the mean, the mean, and one SD above the mean). The shaded intervals around the lines indicate the 95% confidence interval derived from the coefficient estimates in the linear model.

potentially stronger if years of deafness in the contralateral ear were used.

Conclusion and Discussion

Although most adult cochlear implantees have improved hearing outcomes, the degree of improvement in hearing and speech perception varies widely (Pisoni et al., 2017; Boisvert et al., 2020). Despite having been studied extensively over the last 25 years, the exact relationship between preoperative predictive factors and hearing outcome after cochlear implantation is not fully understood (Boisvert et al., 2020; Zhao et al., 2020). In this study, we analyzed a large cohort of adult cochlear implantees ($n = 2,735$) from three different clinics from three countries, including the largest cohort of adult cochlear implantees from an individual clinic to date (MHH, $n = 1,951$). Our analysis provides further refinement of effect size for 21 predictive factors, replicating or confirming 17 significant associations. Exploring associations across multiple datasets highlights a number of complexities that may help to explain the variability across the literature and suggests several avenues to improve the study of factors that influence cochlear implantation outcomes.

We found significant univariate associations for 16 predictive factors, many consistent with the previous literature, with one additional association found through a multivariate analysis. In line with previous studies (Blamey et al., 2013; Lazard et al., 2012; Rubinstein et al., 1999; del Mar Medina et al., 2017; Francis et al., 2005; Zhao et al., 2020), the negative effect of a longer duration of HL prior to implantation across all three clinics was shown to have a strong and consistent effect size for the implanted ear. Years of deafness in the contralateral ear were also significant in the combined analysis, but was only significant in MHH individually. Although many of these associations are not significant in ESIA, this is likely to be a function of sample size, given ESIA has a far smaller sample size ($n = 293$), compared to VUMC ($n = 491$) or MHH ($n = 1,951$).

Several factors are understudied for their explicit role as a predictive factor for cochlear implantation but nevertheless have expected outcomes. As per Kraaijenga et al. (2016), we found that individuals with prelingual HL had worse postoperative WRSs. The impact of being a native language speaker was found to have a highly significant effect, similar to findings by Kilman et al. (2015) and Van Wijngaarden et al. (2002). The course of HL (sudden vs progressive) also had a highly significant effect in the univariate analysis, whereby progressive HL lead to improved outcomes, in line with previous studies (Clark, 2006; Battmer et al., 1995). However, we found that the course of HL was not significant in the multivariate analysis, indicating the information it carries may be reflected in other measurements.

Etiology of HL emerged as another strongly associated predictive factor, adding to the existing evidence investigating the relationship between certain causes for HL and CI outcomes (Blamey et al., 2013; Lazard et al., 2012; Boisvert et al., 2020; Janeschik et al., 2013). As with many studies of etiology in the literature to date, inconsistency in data collection limits our ability to interpret some of these findings. Given the clinics reported a large number of different etiology classes (including uncontrolled free text fields), we grouped etiology values from each clinic into 11 standardized groupings. Although standardizing the data ensure each category had a larger number of observations, it also discards valuable information, for example, grouping different genetic conditions into only one category (e.g., deafness due to mutations in the *connexin* gene family vs deafness due to autosomal recessive syndromes become a single etiology). Moreover, etiology information was only known and available in 55% of the data making subsequent analysis difficult. Despite these challenges, patient etiology remains highly informative of the outcome. As such, refining the collection of etiology-related information may substantially help explain the remaining variance in hearing outcome.

Hearing performance before implantation is known as an indicator of postimplantation hearing capabilities. Our results for the implanted ear are consistent with the previous literature (Blamey et al., 2013; Lazard et al., 2012; Boisvert et al., 2020; Zhao et al., 2020), showing that individuals with a better hearing before implantation are likely to have stronger outcomes postimplantation. Our analysis also paid particular attention to the influence of the preimplantation performance of the contralateral ear, given mixed conclusions in the previous literature. Lazard et al. (2012) showed a positive association between PTA_c and CI outcomes in the implanted ear while Plant et al. (2016) suggested a negative association. Our study found significant univariate associations for PTA_c but varying effect sizes in the univariate analyses. Only ESIA showed a significant association in the multivariate analysis, consistent with Plant et al. (2016). The inconsistency of these findings, along with those in the literature, may indicate that the hearing performance of the contralateral ear is influenced by cohort demographics. Further investigations are required to understand these associations.

Hoppe et al. (2019) hypothesized that PB_{max} , the maximum WRS of a patient using headphones that are achieved through assessment at multiple signal levels, may prove to be a more informative performance indicator than WRS(HA). In the univariate analysis performed on 1,111 implantees from MHH, we found that the effect size of PB_{max} , while significant, is not substantially different from WRS(HA). Moreover, a direct comparison of effect sizes on 333 individuals for whom both measures were taken found the effect sizes of WRS(HA) and PB_{max} were very similar. The differences in outcomes with those of Hoppe

et al. (2019) may reflect differences in cohorts between the two studies, especially the inclusion of participants with prelingual HL, and further investigation of the relative utility of WRS(HA) and PB_{\max} is warranted.

We shed further light on the relevance of implantation age in adults, which has also been a point of contention in the literature (Kraaijenga et al., 2016; Schwab et al., 2015; Zhao et al., 2020; Blamey et al., 2013; Holden et al., 2013). Our univariate analyses showed mixed results, revealing both positive (ESIA and VUMC) and negative (MHH) associations in the individual datasets. Similar results are seen when only analyzing patients with postlingual HL, indicating that the effect is not driven by differences in pre-/postlingual HL. When accounting for other factors through multivariate analysis, a significant but mild negative association was observed for two datasets. This is in line with results from Blamey et al. (1996), Blamey et al. (2013), Holden et al. (2013) and the meta-analysis of Zhao et al. (2020).

To further investigate the varying outcomes across different datasets with respect to age, we examined interactions between predictive factors and found significant nonadditive effects for the age of implantation and years of severe to profound deafness prior to implantation. We showed that this interaction was statistically significant, replicated across two of the three clinics, and improved the variance explained. We also demonstrated that this improvement was not caused by the presence of patients with prelingual HL. However, the underlying drivers of this association are unclear. It is possible that the observed interaction is a statistical artifact due to the uncaptured information about the underlying etiology of HL. An alternative explanation, discussed in a recent review (Simon et al., 2020), is that auditory deprivation has different effects on the brain structure, specifically the primary or secondary auditory cortex, between younger and older individuals, and this may be reflected in auditory and speech perception. This supports a feasible mechanism underpinning the observed statistical association. Future work will be required to further clarify the true driver of this interaction.

We further found a strong association of the implant side and postoperative hearing performance, with right-ear implantation leading to better results. Similar results have previously been reported on adults (Liang et al., 2020) and children with pre- and postlingual HL, as summarized in a meta-analysis by Kraaijenga et al. (2017). Significant interactions with other factors, including prelingual HL and gender, were not found. Possible explanations include hemispherical asymmetries in auditory processing (Mills & Rollman, 1980; Schönwiesner et al., 2007; Brown & Nicholls, 1997) or the well-established right-ear-advantage that has been linked to the language dominance of the left hemisphere (Hugdahl, 2009).

Despite the large cohort, the multivariate analysis only explains a modest 13%–21% of the variance in postoperative WRSs. The meta-analysis on the combined dataset explains 12% of variance, with the lower performance potentially

attributable to the smaller number of included features that overlapped across all datasets. Our results are in line with previous studies, which were able to explain up to 31% (Kraaijenga et al., 2016) of variance for a cohort of patients with prelingual HL and up to 22% (Lazard et al., 2012) for a cohort including a wider range of patients. Any reported R^2 value needs to be discussed in context data samples and a number of input features since models using small datasets and a large number of features may overfit the available data. We therefore also report the adjusted variance explained, which lies between 6% and 20% in single clinics and at 12% for the combined analysis.

The analyses conducted in this work highlight several practical considerations for the standardization of data collected in this space. One such issue is around inconsistency in the definitions of the preoperative factors that are commonly collected. For example, the “years of severe to profound deafness” feature showed a clear difference in distribution between the clinics that is unlikely to be due to cohort make-up alone. Although this predictive factor has been shown to be important in numerous studies (Blamey et al., 2013; Lazard et al., 2012; Zhao et al., 2020), we reexamined the studies that were used to form a recent meta-analysis of the duration of HL (Zhao et al., 2020) and found definitions of HL across the included studies varied. This included asking a patient whether they could use the phone (Rubinstein et al., 1999), asking whether HAs were useful (del Mar Medina et al., 2017) or were based on PTA (Francis et al., 2005). In our study, we found similar variability across the three centers and while this may reduce the precision of estimated effect sizes within our analysis, it also reflects the current nature of data collected and analyzed in the field.

Our analysis was also complicated by the inconsistency of predictive factor availability across the three centers. This is an issue that has been observed in previous meta-analyses of cochlear implantation outcomes (Zhao et al., 2020; Lazard et al., 2012) due to the inconsistent data collection across centers. Such data missingness, while common in the medical domain, limits the ability to examine combined effects of each factor or conduct comparable multivariate analyses across datasets. Although the future exploration of imputation techniques may mitigate this issue and increase the variance explained, such approaches are not a substitute for improved data collection. Moreover, the use of artificial intelligence and machine learning, which are helping advance many health care-related fields but remain nascent in predicting CI outcomes (Chen & Asch, 2017; Crowson et al., 2020a, 2020b), will be greatly improved with larger, more cohesive datasets, which enable more accurate predictive performance.

Given this, a key outcome of this study is to further highlight how standardization of data collection is critical to further improve our understanding of outcomes related to cochlear implantation, supporting many other efforts to

advance this cause. According to a 2018 analysis (Adunka et al., 2018), this process will be a global effort that requires collaboration between manufacturers, HA distributors, cochlear implantation clinics, hospitals, and governments to integrate audiological data with other health data, following the example set by the UK NHS Hearing Health Informatics Collective initiative. Although such standardization is difficult to implement, such systems have been implemented successfully in cancer pathology reporting (Srigley et al., 2009), where adherence to a controlled vocabulary has enabled greater rigor in epidemiological studies across multiple institutions, leading to improved patient outcomes (Williams et al., 2015).

There are several limitations of our current analysis that may impact the interpretation of results. For some predictive factors, our analysis led to different results for the three datasets despite applying the same inclusion criteria for patients. Interpreting such results is difficult given the datasets differ in a number of factors, including cohort differences, testing protocol and/or setup differences, and patient selection criteria, due to differing regulatory rules across countries.

Moreover, despite analyzing the largest number of preoperative factors in a predictive factor analysis, our dataset lacked certain features that have previously been shown to be associated with outcome performance. These include previously studied factors about the implantation itself such as electrode placement (Holden et al., 2013), insertion depth (James et al., 2019), implant brand, and the number of active electrodes during stimulation (Lazard et al., 2012). Information about rehabilitation, a patient's domestic or work environment, medication, comorbidities, cognitive or education level, and social interaction are also likely to increase the explained variance.

A further limitation stems from the selected hearing outcome variable. The use of a single test at a single time point may have limited the power of our analysis, given that hearing outcome is known to change across tests and across time. Instead, combining multiple measures of hearing performance and evaluating these longitudinally may provide a cleaner signal that improves both our ability to detect significant predictive factors and to predict an individual's likely outcome. In addition, the assessment of hearing performance tested with a low number of monosyllabic words introduces limitations on the granularity of our data, as well as contributing to the variance of the measurement results. Finally, tests may not be conducted if they are perceived by the patient as too hard, particularly on the lower end of word recognition performance. In that case, alternative test protocols were used by clinicians that were not included in our analysis.

The results presented in this paper largely align with previous studies, with several findings advancing our understanding of the relationships between predictive factors. Although the findings explain only a modest amount of the variability of WRS outcomes, they shed

light on the nature of several interactions and highlight the ongoing need for data standardization. These results will be complemented by an exploration of nonlinear machine learning-based approaches to explain variation on CI outcome. Furthermore, it is increasingly apparent that predictive factors that are currently being collected are insufficient for strong predictive performance. Instead, factors that more directly capture aspects of the implantation itself, the individual's environment (before and after implantation), and the nature of their HL are promising candidates to consider. These directions combined should lead to significant improvements in our ability to explain implantation outcomes beyond improvements in sample size alone.

Acknowledgments

This research has been conducted primarily while all authors employed by their institutional affiliations. The authors would like to thank Dwarikanath Mahapatra, Benjamin Scott Mashford, Christine Schieber, Kerry Halupka, Suman Sedai, Andrei Pavlov, Andrew Rawlinson, Jianbin Tang, and Gregory Cameron. Without their contributions, insightful discussion, and assistance in preprocessing the three datasets, this study would not have been possible.


Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The collection of the VUMC dataset was supported by a research project grant no. NIH NIDCD R01 DC13117 (principal investigator: Gifford).

ORCID iDs

Eugen Kludt  <https://orcid.org/0000-0001-7030-7604>
 Robert H. Eikelboom  <https://orcid.org/0000-0003-2911-5381>
 Rene H. Gifford  <https://orcid.org/0000-0001-6662-3436>
 Riaan Rottier  <https://orcid.org/0000-0001-7299-5412>
 Hamideh Anjomshoa  <https://orcid.org/0000-0002-0074-2405>

Supplemental Material

Supplemental material for this article is available online.

References

- Adunka O. F., Gantz B. J., Dunn C., Gurgel R. K., & Buchman C. A. (2018). Minimum reporting standards for adult cochlear implantation. *Otolaryngology–Head and Neck Surgery*, *159*, 215–219. <https://doi.org/10.1177/0194599818764329>
- Battmer R., Gupta S., Allum-Mecklenburg D., & Lenarz T. (1995). Factors influencing cochlear implant perceptual performance in 132 adults. *The Annals of Otolaryngology, Rhinology and Laryngology*, *104*, 185–187.
- Benjamini Y., & Hochberg Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing.

- Journal of the Royal Statistical Society: Series B (Methodological)*, 57, 289–300.
- Blamey P., Arndt P., Bergeron F., Bredberg G., Brimacombe J., Facer G., Larky J., Lindström B., Nedzelski J., Peterson A., & Shipp D. (1996). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants. *Audiology and Neurotology*, 1, 293–306. <https://doi.org/10.1159/000259212>
- Blamey P., Artieres F., Başkent D., Bergeron F., Beynon A., Burke E., Dillier N., Dowell R., Fraysse B., Gallégo S., Govaerts P. J., Green K., Huber A. M., Kleine-Punte A., Maat B., Marx M., Mawman D., Mosnier I., O'Connor A. F., & ... Lazard D. S. (2013). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: An update with 2251 patients. *Audiology and Neurotology*, 18, 36–47. <https://doi.org/10.1159/000343189>
- Blamey P. J., Pyman B. C., Clark G. M., Dowell R. C., Gordon M., Brown A. M., & Hollow R. D. (1992). Factors predicting postoperative sentence scores in postlinguistically deaf adult cochlear implant patients. *Annals of Otolaryngology, Rhinology & Laryngology*, 101, 342–348. <https://doi.org/10.1177/000348949210100410>
- Boisvert I., Reis M., Au A., Cowan R., & Dowell R. C. (2020). Cochlear implantation outcomes in adults: A scoping review. *PLoS One*, 15, e0232421. <https://doi.org/10.1371/journal.pone.0232421>
- Brown S., & Nicholls M. E. (1997). Hemispheric asymmetries for the temporal resolution of brief auditory stimuli. *Perception and Psychophysics*, 59, 442–447. <https://doi.org/10.3758/bf03211910>
- Chen J. H., & Asch S. M. (2017). Machine learning and prediction in medicine—beyond the peak of inflated expectations. *The New England Journal of Medicine*, 376, 2507. <https://doi.org/10.1056/nejmp1702071>
- Clark G. (2006). *Cochlear Implants: Fundamentals and Applications. Modern acoustics and signal processing*. Springer.
- Crowson M. G., Dixon P., Mahmood R., Lee J. W., Shipp D., Le T., Lin V., Chen J., & Chan T. C. (2020a). Predicting postoperative cochlear implant performance using supervised machine learning. *Otology and Neurotology*, 41, e1013–e1023. <https://doi.org/10.1097/MAO.0000000000002710>
- Crowson M. G., Ranisau J., Eskander A., Babier A., Xu B., Kahmke R. R., Chen J. M., & Chan T. C. (2020b). A contemporary review of machine learning in otolaryngology–head and neck surgery. *The Laryngoscope*, 130, 45–51. <https://doi.org/10.1002/lary.27850>
- del Mar Medina M., Polo R., Gutierrez A., Muriel A., Vaca M., Perez C., Cordero A., & Cobeta I. (2017). Cochlear implantation in postlingual adult patients with long-term auditory deprivation. *Otology and Neurotology*, 38, e248–e252. <https://doi.org/10.1097/mao.0000000000001257>
- Dowell R. C., Hollow R., & Winton E. (2004). Outcomes for cochlear implant users with significant residual hearing: Implications for selection criteria in children. *Archives of Otolaryngology—Head & Neck Surgery*, 130, 575–581. <https://doi.org/10.1001/archotol.130.5.575>
- Eshraghi A. A., Nazarian R., Telischi F. F., Rajguru S. M., Truy E., & Gupta C. (2012). The cochlear implant: Historical aspects and future prospects. *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology*, 295, 1967–1980. <https://doi.org/10.1002/ar.22580>
- Francis H. W., Yeagle J. D., Bowditch S., & Niparko J. K. (2005). Cochlear implant outcome is not influenced by the choice of ear. *Ear and Hearing*, 26, 7S–16S. <https://doi.org/10.1097/00003446-200508001-00003>
- Gantz B. J., Woodworth G. G., Knutson J. F., Abbas P. J., & Tyler R. S. (1993). Multivariate predictors of audiological success with multichannel cochlear implants. *Annals of Otolaryngology, Rhinology & Laryngology*, 102, 909–916. <https://doi.org/10.1177/000348949310201201>
- Gaylor J. M., Raman G., Chung M., Lee J., Rao M., Lau J., & Poe D. S. (2013). Cochlear implantation in adults: A systematic review and meta-analysis. *JAMA Otolaryngology—Head & Neck Surgery*, 139, 265–272. <https://doi.org/10.1001/jamaoto.2013.1744>
- Goldberg R., Gore J. M., Barton B., & Gurwitz J. (2014). Individual and composite study endpoints: Separating the wheat from the chaff. *The American Journal of Medicine*, 127, 379–384. <https://doi.org/10.1016/j.amjmed.2014.01.011>
- Hahlbrock K. H. (1953). Über sprachaudiometrie und neue wörterteste. *Archiv Für Ohren-, Nasen- Und Kehlkopfheilkunde*, 162, 394–431.
- Hahlbrock K. H. (1960). Kritische betrachtungen und vergleichende untersuchungen der schubertschen und freiburger sprachteste. *Zeitschrift für Laryngologie, Rhinologie, Otologie Und Ihre Grenzgebiete*, 39, 100.
- Harrell J., & Frank E. (2015). *Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis*. Springer.
- Holden L. K., Finley C. C., Firszt J. B., Holden T. A., Brenner C., Potts L. G., Gotter B. D., Vanderhoof S. S., Mispagel K., Heydebrand G., & Skinner M. W. (2013). Factors affecting open-set word recognition in adults with cochlear implants. *Ear and Hearing*, 34, 342. <https://doi.org/10.1097/AUD.0b013e3182741aa7>
- Hoppe U., Hocke T., Hast A., & Iro H. (2019). Maximum preimplantation monosyllabic score as predictor of cochlear implant outcome. *HNO*, 67, 62–68. <https://doi.org/10.1007/s00106-019-0648-0>
- Hugdahl K. (2009). Dichotic listening studies of brain asymmetry. *Encyclopedia of Neuroscience*, 3, 517–522. <https://dx.doi.org/10.1016/B978-008045046-9.00295-3>
- James C. J., Karoui C., Laborde M. L., Lepage B., Molinier C.É., Tartayre M., Escudé B., Deguine O., Marx M., & Fraysse B. (2019). Early sentence recognition in adult cochlear implant users. *Ear and Hearing*, 40, 905–917. <https://doi.org/10.1097/aud.0000000000000670>
- Janeschik S., Teschendorf M., Bagus H., & Arweiler-Harbeck D. (2013). Influence of etiologic factors on speech perception of cochlear-implanted children. *Cochlear Implants International*, 14, 190–199. <https://doi.org/10.1179/1754762812y.0000000017>
- Ji C., Galvin J. J., Chang Y. p., Xu A., & Fu Q. J. (2014). Perception of speech produced by native and nonnative talkers by listeners with normal hearing and listeners with cochlear implants. *Journal of Speech, Language, and Hearing Research*, 57, 532–554. https://doi.org/10.1044/2014_jslhr-h-12-0404
- Kilman L., Zekveld A., Höllgren M., & Rönnberg J. (2015). Native and non-native speech perception by hearing-impaired listeners in noise-and speech maskers. *Trends in Hearing*, 19, 1–12. <https://doi.org/10.1177/2331216515579127>
- Kraaijenga V., Derksen T., Stegeman I., & Smit A. (2017). The effect of side of implantation on unilateral cochlear implant

- performance in patients with prelingual and postlingual sensorineural hearing loss: A systematic review. *Clinical Otolaryngology*, *43*, 440–449. <https://doi.org/10.1111/coa.12988>
- Kraaijenga V. J., Smit A. L., Stegeman I., Smilde J. J., Van Zanten G., & Grolman W. (2016). Factors that influence outcomes in cochlear implantation in adults, based on patient-related characteristics—a retrospective study. *Clinical Otolaryngology*, *41*, 585–592. <https://doi.org/10.1111/coa.12571>
- Lazard D. S., Vincent C., Venail F., Van de Heyning P., Truy E., Sterkers O., Skarzynski P. H., Skarzynski H., Schauwers K., O’Leary S., & Mawman D. (2012). Pre-, per-and postoperative factors affecting performance of postlinguistically deaf adults using cochlear implants: A new conceptual model over time. *PLoS One*, *7*, 1–11. <https://doi.org/10.1371/journal.pone.0048739>
- Leigh J. R., Dettman S. J., & Dowell R. C. (2016). Evidence-based guidelines for recommending cochlear implantation for young children: Audiological criteria and optimizing age at implantation. *International Journal of Audiology*, *55*, S9–S18. <https://doi.org/10.3109/14992027.2016.1157268>
- Liang C., Houston L. M., Samy R. N., Xiang J., & Zhang F. (2020). The effect of side of implantation on the cortical processing of frequency changes in adult cochlear implant users. *Frontiers in Neuroscience*, *14*, 368. <https://doi.org/10.3389/fnins.2020.00368>
- Mills L., & Rollman G. B. (1980). Hemispheric asymmetry for auditory perception of temporal order. *Neuropsychologia*, *18*, 41–47. [https://doi.org/10.1016/0028-3932\(80\)90082-2](https://doi.org/10.1016/0028-3932(80)90082-2)
- Mohammed A. A., & Sarwat S. A. (2014). The side of cochlear implantation and speech intelligibility in pediatric and adult cochlear implantees. *The Egyptian Journal of Otolaryngology*, *30*, 362. <https://doi.org/10.4103/1012-5574.144977>
- Ostertagová E. (2012). Modelling using polynomial regression. *Procedia Engineering*, *48*, 500–506. <https://doi.org/10.1016/j.proeng.2012.09.545>
- Peterson G. E., & Lehiste I. (1962). Revised CNC lists for auditory tests. *Journal of Speech and Hearing Disorders*, *27*, 62–70. <https://doi.org/10.1044/jshd.2701.62>
- Pisoni D. B., Kronenberger W. G., Harris M. S., & Moberly A. C. (2017). Three challenges for future research on cochlear implants. *World Journal of Otorhinolaryngology-Head and Neck Surgery*, *3*, 240–254. <https://doi.org/10.1016/j.wjorl.2017.12.010>
- Plant K., McDermott H., Van Hoesel R., Dawson P., & Cowan R. (2016). Factors predicting postoperative unilateral and bilateral speech recognition in adult cochlear implant recipients with acoustic hearing. *Ear and Hearing*, *37*, 153–163. <https://doi.org/10.1097/aud.0000000000000233>
- Roditi R. E., Poissant S. F., Bero E. M., & Lee D. J. (2009). A predictive model of cochlear implant performance in postlingually deafened adults. *Otology & Neurotology*, *30*, 449–454. <https://doi.org/10.1097/mao.0b013e31819d3480>
- Rubinstein J., Parkinson W., Tyler R., & Gantz B. (1999). Residual speech recognition and cochlear implant performance: Effects of implantation criteria. *The American Journal of Otology*, *20*, 445–452.
- Schönwiesner M., Krumbholz K., Rübsem R., Fink G. R., & von Cramon D. Y. (2007). Hemispheric asymmetry for auditory processing in the human auditory brain stem, thalamus, and cortex. *Cerebral Cortex*, *17*, 492–499. <https://doi.org/10.1093/cercor/bhj165>
- Schwab B., Gandolfi M., Lai E., Reilly E., Singer L., & Kim A. H. (2015). The impact of age on cochlear implant performance. *International Journal of Otolaryngology and Head & Neck Surgery*, *4*, 329. <https://doi.org/10.4236/ijohns.2015.45056>
- Shea III J. J., Domico E. H., & Orchik D. J. (1990). Speech recognition ability as a function of duration of deafness in multichannel cochlear implant patients. *The Laryngoscope*, *100*, 223–226. <https://doi.org/10.1288/00005537-199003000-00002>
- Simon M., Campbell E., Genest F., MacLean M. W., Champoux F., & Lepore F. (2020). The impact of early deafness on brain plasticity: A systematic review of the white and gray matter changes. *Frontiers in Neuroscience*, *14*, 206. <https://doi.org/10.3389/fnins.2020.00206>
- Srigley J. R., McGowan T., MacLean A., Raby M., Ross J., Kramer S., & Sawka C. (2009). Standardized synoptic cancer pathology reporting: A population-based approach. *Journal of Surgical Oncology*, *99*, 517–524. <https://doi.org/10.1002/jso.21282>
- Summerfield A., & Marshall D. (1995). Preoperative predictors of outcomes from cochlear implantation in adults: Performance and quality of life. *The Annals of Otology, Rhinology and Laryngology*, *166*, 105–108.
- Thornton A. R., & Raffin M. J. (1978). Speech-discrimination scores modeled as a binomial variable. *Journal of Speech and Hearing Research*, *21*, 507–518. <https://doi.org/10.1044/jshr.2103.507>
- Van Wijngaarden S., Steeneken H., & Houtgast T. (2002). Quantifying the intelligibility of speech in noise for non-native talkers. *The Journal of the Acoustical Society of America*, *112*, 3004–3013. <https://doi.org/10.1121/1.1512289>
- Waltzman S., Fisher S., Niparko J., & Cohen N. (1995). Predictors of postoperative performance with cochlear implants. *The Annals of Otology, Rhinology and Laryngology*, *165*, 15–18.
- Waskom M. L. (2021). Seaborn: Statistical data visualization. *Journal of Open Source Software*, *6*, 3021. <https://doi.org/10.21105/joss.03021>
- Williams C. L., Bjugn R., & Hassell L. A. (2015). Current status of discrete data capture in synoptic surgical pathology and cancer reporting. *Pathology and Laboratory Medicine International*, *7*, 11. <https://doi.org/10.2147/PLMI.S64378>
- Zhao E. E., Dornhoffer J. R., Loftus C., Nguyen S. A., Meyer T. A., Dubno J. R., & McRackan T. R. (2020). Association of patient-related factors with adult cochlear implant speech recognition outcomes: A meta-analysis. *JAMA Otolaryngology–Head and Neck Surgery*, *146*, 613–620. <https://doi.org/10.1001/jamaoto.2020.0662>