Original Article

Enlarged Vestibular Aqueduct and Associated Inner Ear Malformations: Hearing Loss Prognostic Factors and Data Modeling from an International Cohort

Haroon S. Saeed[1], Martin Fergie[2], Kristianna Mey[3], Niels West[3], Michael Bille[3], Per Caye-Thomasen[3,4], Robert Nash[5], Shakeel R. Saeed[6,7], Stavros M. Stivaros[8], Graeme Black[9], Iain A. Bruce[1,10]

[1]Royal Manchester Children's Hospital, Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK
[2]Division of Informatics, Imaging and Data Sciences, University of Manchester, University of Manchester, School of Biological Sciences, Manchester, UK
[3]Department of Otorhinolaryngology, Head and Neck Surgery, and Audiology, Copenhagen Hearing and Balance Center (CHBC), Copenhagen University Hospital, Copenhagen, Denmark
[4]University of Copenhagen, Faculty of Health and Medical Sciences, Copenhagen, Denmark
[5]Department of Paediatric Otolaryngology, Great Ormond Street Hospital for Children, NHS Foundation Trust, London, UK
[6]The Royal National ENT and Eastman Dental Hospitals, University College London Hospitals NHS Foundation Trust, London, UK
[7]University College London Ear Institute, London, UK
[8]Department of Paediatric Radiology, Royal Manchester Children's Hospital, Manchester, UK
[9]Manchester Centre for Genomic Medicine, Manchester Academic Health Sciences Centre, Manchester University NHS Foundation Trust, Manchester, UK
[10]Department of Infection, Immunity and Respiratory Medicine, University of Manchester, School of Biological Sciences, Faculty of Biology, Medicine and Health, Manchester, UK

ORCID IDs of the authors: H.S.S. 0000-0002-2624-0518, M.F. 0000-0002-9531-6109, K.M. 0000-0003-3784-8331, N.W. 0000-0002-4556-1309, M.B. 0000-0002-2368-1537, P.C.-T. 0000-0002-6837-694X, R.N. 0000-0002-7861-3734, S.R.S. 0000-0002-2316-5655, S.M.S. 0000-0003-3023-541X, G.B. 0000-0001-8727-6592, I.A.B. 0000-0003-0831-4760

BACKGROUND: There is a need to operationalize existing clinical data to support precision medicine in progressive hearing loss (HL). By utilizing enlarged vestibular aqueduct (EVA) and its associated inner ear abnormalities as an exemplar, we model data from a large international cohort, confirm prognostic factors for HL, and explore the potential to generate a prediction model to optimize current management paradigms.

METHODS: An international retrospective cohort study. Regression analyses were utilized to model frequency-specific HL and identify prognostic factors for baseline average HL severity and progression. Elastic-net regression and machine learning (ML) techniques were utilized to predict future average HL progression based upon routinely measurable clinical, genetic, and radiological data.

RESULTS: Higher frequencies of hearing were lost more severely. Prognostic factors for HL were the presence of incomplete partition type 2 (coefficient 12.95 dB, $P = .011$, 95% CI 3.0-22 dB) and presence of sac signal heterogeneity ($P = .009$, 95% CI 0.062-0.429) on magnetic resonance imaging. Elastic-net regression outperformed the ML algorithms (R^2 0.32, mean absolute error 11.05 dB) with coefficients for baseline average hearing level and the presence of sac heterogeneity contributing the most to prediction outcomes.

CONCLUSION: Incomplete partition type 2 and endolymphatic sac signal heterogeneity phenotypes should be monitored closely for hearing deterioration and need for early audiological rehabilitation/cochlear implant. Preliminary prediction models have been generated using routinely collected health data in EVA. This study showcases how international collaborative research can use exemplar techniques to improve precision medicine in relatively rare disease entities.

KEYWORDS: Hearing loss, machine learning, modeling, neurotology, sensorineural hearing loss, statistics

INTRODUCTION

Hearing loss (HL) affects approximately 1-3 per 1000 newborns, and despite this, there is a relative paucity of clinical guidance-related specifically to individualized surveillance and targeted audiological rehabilitation in cases of progressive deafness in children. Most clinical research efforts have been focused on early identification by newborn screening programs. Progressive HL has been shown to affect almost 50% of children with confirmed HL after newborn screening.[1] A greater prognostic understanding of the anticipated severity and trajectory of progressive HL would ensure that clinical resources are focused on individuals with the greatest need for intensive audiological surveillance. As a result, individuals at risk of early deterioration of hearing would be identified before the debilitating consequences of inadequate language acquisition and access to mainstream education. Likewise, parents would benefit from a greater understanding of their child's hearing journey, increasing motivation and engagement with hearing services.

To improve prognostic understanding of progressive HL, robust clinical prognostic factors (PFs) should be identified. This is because PFs form the building blocks of clinical prediction models which would ultimately position clinicians to anticipate HL and optimize pathways for repeat audiological testing, hearing aid provision, and timeframes to cochlear implant surgery (CI).[2-4]

In this study, we aim to advance prognosis research in an exemplary condition associated with progressive HL, namely, enlarged vestibular aqueduct (EVA).[4] This is the most common radiological abnormality associated with childhood sensorineural HL (SNHL), and can be concurrently associated with other inner ear abnormalities such as incomplete partition (IP) of the cochlea. Despite this, clinicians are unable to advise on the likely HL trajectory per ear, nor do they actively seek out patients with a high risk of HL progression/severity for early intervention and management. This may exacerbate parental and clinician uncertainty following early identification of HL in the neonatal period by newborn screening and subsequent confirmation of etiology in infancy using MR imaging in natural sleep. There has been growing interest in identifying exploratory PFs related to HL severity and trajectory in EVA and its associated inner ear abnormalities.[5-7] In recent years, a PF systematic review and other low risk of bias studies have identified a cluster of important HL PFs to explore further, namely, gender, zygosity (number of mutations in the Pendred gene), baseline hearing thresholds, and radiological morphology of the vestibular aqueduct.[3,8-10]

We aim to interrogate a large international database of patients with EVA to discover novel PFs for EVA HL and confirm the role of previously identified PFs. In tandem, we will explore the role of regression analysis and machine learning (ML) (a subbranch of artificial intelligence (AI)) to produce academic models for HL progression. We

wish to demonstrate that lessons learnt from the techniques used to model our clinical database can be applied widely to other conditions causing progressive SNHL—an area with a significant unmet clinical need for precision medicine.

MATERIAL AND METHODS

An international multicenter retrospective review of patients identified with EVA (UK and Denmark). Patient data were recruited from 4 tertiary CI centers. This study had ethical approval from the Ethics Committee of Copenhagen University Hospital and the National Health Service (NHS) Research Ethics Committee, the UK Health Research Authority (HRA) (IRAS 271326) and separate Danish ethical approvals. No informed consent was required.

Patients

The inclusion criterion was patients with a confirmed diagnosis of EVA on radiological imaging assessed at any of the sites up to January 2020 (DOB range 1946-2018). Radiological diagnosis was confirmed by at least 1 suitably trained radiologist. Radiological inclusion criteria for EVA were applied to all scan data according to Cincinnati criteria—the midpoint width of the vestibular aqueduct should be more than 0.9 mm and/or the width of the operculum greater than 1.9 mm.[11] Any associated inner ear abnormalities were recorded. The exclusion criterion was patients in whom longitudinal audiological data were missing (less than 2 consecutive reliable audiology results). Twenty-one candidate prognostic factors (covariates) were recorded per patient and are outlined in Table 1.

Genetic Measurements

For UK subjects, genetic data were acquired from official National Health Service regional molecular genetics laboratory reports. For Danish cohorts, preexisting EVA genetic records were acquired. All *SLC26A4* variants were classified in accordance with the American College of Medical Genetics and Genomics.[12] Variants for each patient were recorded and patients were categorized by variant type using the ClinVar online database[13] (no variant, missense variant, splice donor variant, other variant) to explore a predictive relationship with HL. "Other variant" represents a group of variants with low representation to simplify modeling. We did not categorize variants by a number of mutated alleles as our previous work failed to show an association with HL and conflicting associations are reported in the literature.[3,14]

Radiological Measurements

Specific morphological features of the enlarged endolymphatic duct (ED) and sac (ES) were measured and recorded from a subgroup of 170 patients in which MR imaging was available. Imaging features contained measurements derived from previously reported low risk of bias studies and exploratory measurements developed by the study team based upon overall morphology of the ED and ES.[5] All feature measurements were from T2-weighted axial images. Imaging features recorded are shown in Table 1. Visual schematics of the novel imaging features are provided in the online supplementary content.

Handling of Audiological Data and Statistical Analyses

Where available and complete, we collated the results of auditory brainstem responses (ABRs) and age-appropriate hearing test results at 250, 500, 1000, 2000, and 4000 Hz (pure tone audiometry, visual reinforcement audiometry, and play audiometry). Patient's age was

Table 1. Candidate Prognostic Factors Used in Modeling Analyses

Candidate Clinical/Radiological/Genetic/Audiological Prognostic Factors	Description
Incomplete partition	
Type 1	Type 1 = Absence of the modiolus and interscalar septum
Type 2	Type 2 = Cochlea has 1.5 turns and there is coalescence of the apical and middle turns (cystic apex)
Endolymphatic duct types (1-4)	
Type 1 (club shaped)	Type 1 = the duct connecting to the common crus/vestibule is widest causing a club appearance
Type 2 (hockey stick shaped)	Type 2 = the duct is concave with no obvious widening at entry to the vestibule
Type 3 (rod shaped)	Type 3 = the duct is straight, relatively thin with no curvature
Type 4 (short and/or narrow)	Type 4 = the duct is relatively short/narrow not fitting with other shape types
Endolymphatic sac mild/moderate/massive enlargement	The relative appearance of the sac is subjectively classified into mild/moderate and massive
Endolymphatic sac signal heterogeneity	Presence of both hyperintense and hypointense signal in the sac
Endolymphatic duct midpoint	The midpoint between the common crus and the operculum
Endolymphatic sac width and length	Maximal widths and lengths of the sac that can be appreciated on axial imaging
Endolymphatic sac types	
Type 1 (dumbbell)	Type 1 = sac with pinched point/narrowing with 2 dilated areas on either side
Type 2 (laterally dilated)	Type 2 = sac is dilated with a large lateral bulge and narrower medial component
Type 3 (medially dilated)	Type 3 = sac is wider medially than it is laterally
Type 4 (short and or narrow)	Type 4 = sac is relatively short/narrow and does not meet other type criteria
Male sex	NA
Female sex	
Genetic variant type (*SLC26A4* gene)	
No variant detected	
Missense	Missense = a single base pair substitution
Splice site	Splice site = an alteration of the DNA sequence that occurs at the boundary of an intron and an exon
Other	Other = deletions, frameshifts, and noncoding transcripts

A description for each covariate is provided. Additional schematics for each shape type of the endolymphatic duct and sac are provided in the supplemental online content. Choice of covariate was based upon their previous reported role in EVA hearing loss.[8,14] However, the endolymphatic duct shapes, sac shapes, and genotype were chosen as novel exploratory covariates (their role in EVA hearing loss has not yet been previously explored). EVA, enlarged vestibular aqueduct.

recorded in months. We did not include measurements from 3000 to 8000 Hz, as these frequencies were not uniformly tested over the period of data entry and so were subject to large amounts of missing data. In patients with profound HL in which thresholds could not reliably be obtained (signposted by a ">" symbol on the audiogram), it was decided that such values should be inputted as 120 dB loss to aid the uniform analysis of continuous outcomes. Data modeling steps are outlined in the following sections.

Statistical Analysis
All analyses were performed using Jupyter notebook in Python coding language and incorporated statistical packages (© Copyright 2015, Jupyter Team USA https://jupyter.org.):

Step 1: Utilize linear regression to characterize frequency-specific HL over time.

We modeled change in average *frequency-specific* HL from baseline (first available accurate hearing test) over time (years). The purpose of this subanalysis was to obtain a general initial understanding of HL trajectories at ear level over time. Regression was performed upon frequency-specific data from 1526 audiograms.

Step 2: Identify prognostic covariates for HL using multivariate linear regression.

This second subgroup analysis enabled us to identify prognostic covariates associated with baseline average HL severity *and* severity of future average HL progression (based on initial hearing test results). For the assessment of future HL progression, data were modeled longitudinally by modeling the average hearing level across the frequency bands as the regression target. The baseline hearing level, corresponding to the patient's first recorded hearing test, was included as a predictor variable along with the time since the first test (months) and the covariates measured at baseline. For each covariate, an interaction term with the time since the first test was added to allow for the measurement of whether a feature corresponded with increased or decreased HL progression.

We used the HC3 method to adjust for heteroskedasticity in the linear regression analyses, ensuring that standard errors were appropriate for the longitudinal modeling process. This improved confidence in our reporting of significant covariates.[15,16] We used up to the first 100 months of audiological data available to mitigate the effect of outliers having undue influence on model coefficients.

Step 3: Explore the ability to construct predictive models for average future HL progression.

Using the same data preparation process as in step 1, we explored various modeling techniques to predict average HL over time given the baseline covariates and hearing test data. The elastic-eet linear regression model was used, which combines L1 and L2 regularization to limit model overfitting.[17,18] This regularization technique would potentially limit the overfitting of a relatively small data set.

We then compared elastic-net model performance with that of the 2 supervised ML algorithms (histogram gradient boosting and random forest regression) to model future average HL. We chose these

algorithms because they potentially require less assumptions about data, can handle errors pertaining to data imbalance, and are less reliant upon linear relationships between the variables.[19-21]

Cross-validation across the entire data set was used to obtain an estimate of the model performance on unseen data. We opted for this strategy, rather than an independent validation cohort, due to the limited number of patients in the overall dataset.

For the elastic-net model, the hyper-parameters controlling the level of regularization and L1-ratio were selected using nested cross-validation (10 inner and 10 outer folds). Default hyper-parameters were used for the histogram gradient boosting and random forest models as experiments showed no benefit by performing a hyper-parameter search for these models. A further explanation of the algorithms and statistical terminology can be found in the supplemental content.

RESULTS

The results of the initial linear regression are provided in Table 2. The analyses show that baseline average HL (taken at the first available hearing test) is *significantly* worse in higher frequencies compared to lower frequencies in which the confidence intervals for a given frequency baseline average do not overlap with the coefficient (average baseline HL) of another frequency. For example, the baseline average HL at 500 Hz (75.7 dB) was significantly lower than the baseline average HL at 4000 Hz (91.7 dB).

Similarly, the linear regression analysis provides an estimate for the degree of HL to be expected per year, per frequency and shows there is a significant difference in the rate of HL in higher frequencies (rate of HL is more severe) compared to lower frequencies. Significance is demonstrated when the CI for a given frequency does not overlap with the coefficient for the rate of HL per year in another frequency. For example, the rate of HL at 4000 Hz (0.83 dB per year) was significantly higher than the rate of HL at 500 Hz (0.36 dB per year).

For the multivariate analysis, our dataset consisted of 229 patients, 94 males, and 133 females. Of these, 170 patients had imaging measurements available for further analysis. Modeling data pertaining to the identification of prognostic covariates for HL are shown in Table 3.

Data from 148 ears (148 hearing tests) were used to identify prognostic covariates for baseline average HL (72 dB, model intercept). Incomplete partition type 2 was a *significant* predictor for worse baseline average HL (coefficient 12.95 dB, standard error 5.0 dB, 95%

CI 3.0-22 dB, $P = .011$). This suggests that in ears with EVA and IP-2, baseline average HL is worse by 12.9 dB. All other imaging covariates, gender, and genetic variant type did not significantly impact HL. Goodness of fit (R^2) was 0.168, suggesting approximately 16.8% of the variance in baseline average HL is explained by the model.

Data from 120 ears (384 hearing tests) were used to model future average HL progression from initial hearing test result. On average HL progressed from baseline by 36 dB (model intercept) per ear. When accounting for the interaction of time since the first hearing test, the presence of ES signal heterogeneity was significantly predictive for worse HL progression (coefficient = 0.24 dB, 95% CI 0.062-0.429, $P = .009$). Other clinical and genetic covariates did not show significant associations with HL. Goodness of fit (R^2 0.43) suggested that 43% of the variance in future average HL from baseline, up to 100 months post initial testing, is explained by the model.

The results of the predictive models for future average HL progression (from baseline hearing test) are shown in Table 4 and are based on the analysis of 120 ears (384 hearing tests). The elastic-net regression provided the best model performance (R^2 0.32) and lowest mean absolute error (11.05 dB) when evaluated using nested cross-validation. The elastic-net model utilized the coefficients for baseline average HL (0.45 dB) and presence of sac heterogeneity (0.143 dB) preferentially for model prediction, with other coefficients having very small impacts on model performance (coefficients <0.0). The ML algorithms performed slightly lower (R^2 0.26 and 0.27) than the elastic-net model. A visual representation of elastic-eet prediction model performance is shown in Figure 1.

DISCUSSION

This study has used HL in EVA as an exemplar to highlight recommended strategies for the hearing health academic community to improve prognosis research in progressive HL.[4] The main principles deployed were (i) to determine exploratory PFs of interest a priori based upon robust research in the wider literature, (ii) to generate the largest sample size possible (in this case from an international collaboration), and (iii) to avoid dichotomization of continuous outcomes.[4,6] We adopted a multivariate analysis approach to confirm the prognostic effects of individual covariates using different outcome measures and chose the best data start point possible (baseline average hearing test result) within the constraints of retrospective research design.

To contextualize our results, prior to this study, confidence in which clinical factors (if any) were prognostic for HL trajectory in EVA was

Table 2. Frequency-Specific Hearing Loss Demonstrated by Linear Regression Analysis

Frequency (Hz)	Average Baseline HL	Standard Error of Baseline Average	95% CI of Baseline Average	Rate of HL per Year	Standard Error of HL Rate	95% CI of HL Rate
250	69.9	0.79	68.3-71.4	0.21	0.09	0.038-0.39
500	75.7	0.65	77.4-77	0.36	0.08	0.18-0.53
1000	80.1	0.66	78.7-81.3	0.55	0.091	0.37-0.73
2000	86.3	0.64	85-87.5	0.70	0.83	0.53-0.87
4000	91.7	0.66	90.4-93.0	0.83	0.09	0.64-1.01

Both initial baseline average HL per frequency and the rate of HL per year are shown. The values in all columns are in decibel apart from the frequency column.
 HL, hearing loss; Hz, hertz.

Table 3. Exploring the Prognostic Effect of Clinical Covariates on Hearing Loss

Covariate (Candidate Prognostic Factor)	Modeling Future Change in Average HL from Baseline Hearing Test			Modeling Baseline Average HL Severity		
	Coefficient	P	95% CI	Coefficient	P	95% CI
IP-1	0.51	.45	−0.8 to 1.85	−36.3	.45	−131 to 58.8
IP-2	0.1	1.2	−0.06 to 0.27	12.95	.01	3.0 to 22.9
Type 2 ED	−0.09	.34	−0.28 to 0.1	−0.23	.96	−10.8 to 10.4
Type 3 ED	−0.004	.96	−0.22 to 0.21	1.81	.72	−8.5 to 12.2
Type 4 ED	0.10	.25	−0.07 to 0.27	−1.65	.75	−12.3 to 8.9
ES heterogeneity	0.24	.009	0.062 to 0.43	−5.9	.12	13.53 to 1.6
ES type 2	0.28	.057	−0.009 to 0.58	3.3	.5	−7.9 to 14.6
ES type 3	0.13	.41	−0.19 to 0.46	6.6	.25	−4.8 to 18.1
ES type 4	0.26	.055	−0.05 to 0.53	1.38	.8	−10.6 to 13.4
ES mild enlargement	−0.15	.43	−0.52 to 0.23	−6.34	.4	−21.3 to 8.5
ES moderate enlargement	−0.107	.38	−0.34 to 0.13	−4.94	.42	−17.08 to 7.2
Male gender	−0.002	.98	−0.19 to 0.18	4.9	.208	−2.76 to 12.56
Missense variant	0.01	.93	−0.23 to 0.29	5.6	.25	−4.04 to 15.2
Splice site variant	−0.22	.12	−0.5 to 0.06	−6.8	.28	−19.52 to 5.7
"Other" variant	0.08	.5	−0.16 to 0.33	−4.2	.429	−14.6 to 6.2
Baseline average HL	−0.003	.255	−0.009 to 0.002	NA	NA	NA
ED midpoint width	−0.91	.66	−5.0 to 3.2	2.7	.38	−3.5 to 9.04
ES length	−0.02	.18	−0.05 to 0.01	0.27	.66	−0.98 to 1.5
ES width	−0.02	.51	−0.1 to 0.05	−1.9	.17	−4.7 to 0.84

Two outcome measures were used (baseline average hearing loss and future change in average hearing loss) with multivariate regression. All numerical values (apart from the *P* values) are in decibels. For future change in average hearing loss the reported coefficients are those with the "time since first hearing test" interaction term to aid readability. A full report can be found in the supplemental content (including coefficients without the interaction term applied).
Interpretation: a positive coefficient (greater than 0) implies worse HL, whereas a negative (−) coefficient implies protection against HL. For example, the presence of a splice site variant predicts an average of 6.8 dB less hearing loss at baseline (not significant as *P* > .05), whereas the presence of sac signal heterogeneity predicts a 0.24 dB increased progression of HL from baseline average. Note: default variants (female gender, type 1 ED, type 1 ES type, no variant, etc.) are not shown in the table.
ED, endolymphatic duct; ES, endolymphatic sac; HL, hearing loss; IP-1, incomplete partition type 1; IP-2, incomplete partition type 2.

limited—only a handful of low risk of bias studies with adequate PF design and smaller cohorts.[3,9] The frequency HL analysis statistically confirms that higher frequencies of hearing are lost more severely at baseline and continue to be lost more severely over time compared to lower frequencies of hearing. This study therefore provides novel and useful clinical information to be passed on to individuals with a new diagnosis of EVA pertaining to the general anticipated HL trajectory.

We provide evidence that covariates from routinely measurable clinical data account for up to 43% of HL progression from baseline

Table 4. Modeling Techniques

Model Type	Goodness of Fit R^2 (0-1)	Mean Absolute Error (dB)
Elastic-net regression	0.32	11.05
Random forest regression (machine learning)	0.26	11.38
Histogram gradient boosting regression (machine learning)	0.27	11.25

Each modeling technique type is outlined. The R^2 value indicates goodness of fit, that is, the proportion of the variance in hearing loss prediction accounted for by the covariates of the model. The mean absolute error indicates the average number of decibels the model could over/underpredict future average hearing loss progression by.

hearing and 16.8% of initial average HL. Morphology of the membranous labyrinth, specifically the presence of IP-2, and signal intensity within the enlarged ES have independent predictive effects on baseline average HL and degree of future average HL progression, respectively. As these PFs are corroborated in other low risk of bias studies, we are confident they should be used in future prognostic model development.[9,10,22] The degree of initial baseline average HL and the presence of sac signal heterogeneity appear to be important coefficients for predicting average HL progression (elastic net).

The clinical implication of these findings is 3-fold. First, given that ES signal heterogeneity can only be assessed on MR imaging, we recommend that this modality should be acquired as early as possible in patients suspected of EVA, as our data support that they are at increased risk of HL progression. Second, these findings support that the progressive component of HL in EVA could be exacerbated by reflux of hyperosmolar proteinaceous sac content causing damage to the sensory organ.[23] Third, patients with these PFs should undergo rigorous audiological surveillance such that timeframes for surgical management with CI can be optimized especially if they have a significant degree of HL at baseline testing.[24,25]

We did not find male sex or endolymphatic midpoint width to be PFs for HL. This is contrary to previous low risk of bias studies but may be explained by the fact different outcome measures have been used
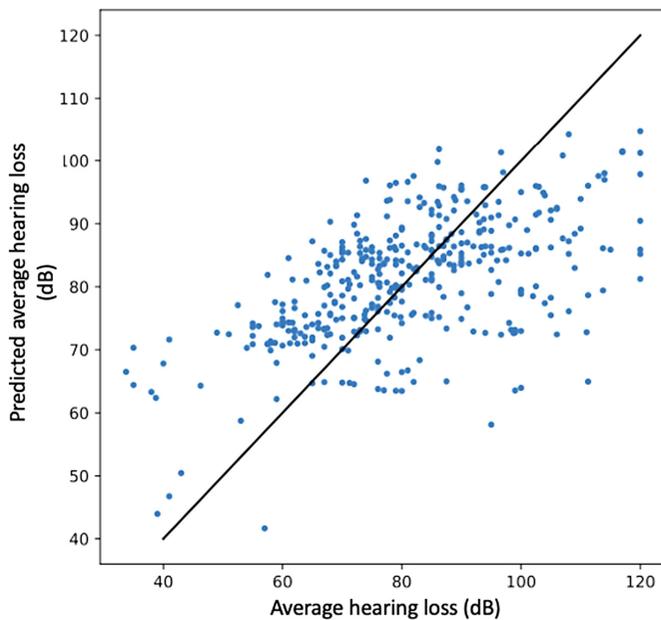
Figure 1. A graphical representation of the elastic-net regression performance. The slope represents a fitted 1 to 1 line (e.g., the point of intercept on the slope for 80 dB on the *x*-axis is also the point of intercept on the slope for 80 dB on the *y*-axis). The data plots are predicted values for future average HL per ear. If the predictive capability of the model was 100% ($R^2 = 1$) then all the data points would line up along the slope. The scatter of data points represents the model performance of $R^2 = 0.32$. In other words, 32% of variation in future average HL can currently be predicted by the model.

between studies, and an interplay between multiple factors determines timeframes for CI surgery.[3,9,14]

Our best-performing clinical prediction model for future average HL progression from baseline test was the elastic net. It suggests that using the current candidate predictors for EVA HL, the model performs 32% better than chance alone but may give predictions that are on average 11 dB better or worse than the true value of average HL. Clearly, there is a degree of inherent predictive capability within readily accessible health-care data and this study is therefore a step in the right direction toward a clinical model for EVA HL progression. These findings should rally the EVA research community to further explore biomarkers in large collaborative clinical data sets to identify other novel PFs. This will improve R^2 (goodness of fit) and reduce mean absolute error, therefore improving the capability of generating a clinically translatable model.

This is the first study to explore the role of ML algorithms to predict the progression of future average HL in EVA. Machine learning is an emerging field within prognosis research and modeling HL.[2] There is a growing impetus to harness the power of ML as we enter an era of digital transformation in health care.[21,26] Many ML algorithms are "data hungry" and there is a paucity of research to determine if they can be effectively applied to relatively rarer diseases (such as EVA) with smaller data sets.[27] Overall predictive performance was modest (best R^2 0.27) and showed similar mean absolute errors to the elastic-net regression. However, putting these results into context, we show for the first time, that based on a relatively small data set in a rare disease (in which limited data on priori predictors has been established), the ML algorithms provide a degree of predictive capability. The wider implication of this study is that the prognosis

analyses used, and implementation of novel ML algorithms, serve as a blueprint for other researchers to draw upon across the domain of progressive HL prediction.

Limitations
This study is limited inherently by its retrospective study design. This resulted in a significant proportion of data parameters with missing data, for example, unavailable imaging data and lack of genetic testing. Although this can introduce a bias within our reported results, we provide the largest subgroup sample sizes to date in EVA research. Our data recruitment sites were all tertiary referral CI centers and so a selection bias may impact upon our results—we may not have captured EVA patients with less severe HL across the community yet to be referred. We aimed to reduce the impact of this by ensuring the earliest possible hearing test data was used prior to referral to CI centers.

Analyses from this large international collaboration have confirmed that HL in EVA and its associated inner ear abnormalities are more severe in higher frequencies. The presence of endolymphatic sac signal heterogeneity and associated IP-2 are prognostic factors for HL in EVA and so early MR imaging is essential in the workup of children following diagnosis of permanent childhood hearing impairment of the sensorineural type. Audiological assessments in such patients should be frequent in anticipation of earlier HL progression and the need for CI surgery, especially if there is significant baseline HL detected from the outset. We have shown that the principles of prognosis research can be applied to the relatively rare disease entity that is progressive HL to generate preliminary clinical prediction models. Further PFs should be identified to improve model performance and work toward clinical translation. Machine learning models showed less predictive capability than penalized regression analysis but provided an alternative approach to model HL data in the wider context of data science in progressive HL research.

REFERENCES

1. Barreira-Nielsen C, Fitzpatrick E, Hashem S, Whittingham JA, Barrowman N, Aglipay M. Progressive hearing loss in early childhood. *Ear Hear*. 2016;37(5):e311-e321. [CrossRef]
2. Riley RD, Windt D van der, Croft P, Moons KGM. Prognosis research in healthcare: concepts, methods, and impact [internet]. https://global.

oup.com/academic/product/prognosis-research-in-healthcare-9780198 796619?cc=pt&lang=en; 2019&.

3. Saeed HS, Kenth J, Black G, Saeed SR, Stivaros S, Bruce IA. Hearing loss in enlarged vestibular aqueduct: a prognostic factor systematic review of the literature. *Otol Neurotol*. 2021;42(1):99-107. **[CrossRef]**

4. Riley RD, Hayden JA, Steyerberg EW, et al. Prognosis research strategy (Progress) 2: prognostic factor research. *PLOS Med*. 2013;10(2):e1001380. **[CrossRef]**

5. Valvassori GE, Clemis JD. The large vestibular aqueduct syndrome. *Laryngoscope*. 1978;88(5):723-728. **[CrossRef]**

6. Steyerberg EW, Moons KGM, van der Windt DA, et al. Guidelines and guidance prognosis research strategy (Progress) 3: prognostic model research. 2013;10(2):e1001381.

7. Arjmand EM, Webber A. Audiometric findings in children with a large vestibular aqueduct. *Arch Otolaryngol Head Neck Surg*. 2004;130(10):1169-1174. **[CrossRef]**

8. Mey K, Bille M, Rye Rasmussen SH, Tranebjærg L, Cayé-Thomasen P. The natural history of hearing loss in pendred syndrome and non-syndromic enlarged vestibular aqueduct. *Otol Neurotol*. 2019;40(3):e178-e185. **[CrossRef]**

9. Farrokhian N, Kocharyan A, Ruthberg J, et al. Predictive modeling and risk stratification of patients with enlarged vestibular aqueduct. *Laryngoscope*. 2022;132(7):1439-1445. **[CrossRef]**

10. Campbell AP, Adunka OF, Zhou B, Qaqish BF, Buchman CA. Large vestibular aqueduct syndrome: anatomic and functional parameters. *Laryngoscope*. 2011;121(2):352-357. **[CrossRef]**

11. Boston M, Halsted M, Meinzen-Derr J, et al. The large vestibular aqueduct: a new definition based on audiologic and computed tomography correlation. *Otolaryngol Head Neck Surg*. 2007 June;136(6):972-977. **[CrossRef]**

12. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology Sue. *Genet Med*. 2015;17(5):405-424.

13. ClinVar [internet]. Available at: https://www.ncbi.nlm.nih.gov/clinvar/. Accessed June 2022.

14. Saeed HS, Rajai A, Nash R, et al. Enlarged vestibular aqueduct: disease characterisation and exploration of potential prognostic factors for cochlear implantation. *Otol Neurotol*. 2022; 1;43(5):e563-e570. **[CrossRef]**

15. Hayes AF, Cai LI. Using heteroskedasticity-consistent standard error estimators in OLS regression: an introduction and software implementation. *Behav Res Methods*. 2007;39(4):709-722. **[CrossRef]**

16. MacKinnon JG, White H. Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *J Econ*. 1985 September 1;29(3):305-325. **[CrossRef]**

17. van Houwelingen JC. Shrinkage and penalized likelihood as methods to improve predictive accuracy. *Stat Neerl*. 2001;155(1):17-34.

18. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. 2010;33(1):1-22. **[CrossRef]**

19. Khalilia M, Chakraborty S, Popescu M. Predicting disease risks from highly imbalanced data using random forest. *BMC Med Inform Decis Mak*. 2011;11:51. **[CrossRef]**

20. Zhang Z, Zhao Y, Canes A, Steinberg D, Lyashevska O, written on behalf of AME Big-Data Clinical Trial Collaborative Group. Predictive analytics with gradient boosting in clinical medicine. *Ann Transl Med*. 2019;7(7):152. **[CrossRef]**

21. Wilson A, Saeed H, Pringle C, Eleftheriou I, Bromiley PA, Brass A. Artificial intelligence projects in healthcare: 10 practical tips for success in a clinical environment. *BMJ Health Care Inform*. 2021;28(1):e100323. **[CrossRef]**

22. Mey K, Muhamad AA, Tranebjaerg L, et al. Association of SLC26A4 mutations, morphology, and hearing in pendred syndrome and NSEVA. *Laryngoscope*. 2019;129(11):2574-2579. **[CrossRef]**

23. Gopen Q, Zhou G, Whittemore K, Kenna M. Enlarged vestibular aqueduct: review of controversial aspects. *Laryngoscope*. 2011;121(9):1971-1978. **[CrossRef]**

24. Hodge SE, Thompson NJ, Park LR, Brown KD. Enlarged vestibular aqueduct: hearing progression and cochlear implant candidacy in pediatric patients. *Otol Neurotol*. 2021;42(1):203-206. **[CrossRef]**

25. Hall AC, Kenway B, Sanli H, Birman CS. Cochlear implant outcomes in large vestibular aqueduct syndrome—should we provide cochlear implants earlier? *Otol Neurotol*. 2019;40(8):e769-e773. **[CrossRef]**

26. Topol E. *Preparing the Healthcare Workforce to Deliver the Digital Future The Topol Review. An Independent Report on behalf of the Secretary of State for Health and Social Care*. Available at: https://topol.hee.nhs.uk/wp-content/uploads/HEE-Topol-Review-2019.pdf. Accessed July 2022.

27. Zhang A, Xing L, Zou J, Wu JC. Shifting machine learning for healthcare from development to deployment and from models to data. *Nat Biomed Eng*. 2022;6(12):1330-1345. **[CrossRef]**