Research *Articles*

A multimodal database for the collection of interdisciplinary audiological research data

M. Amparo Callejón-Leblic[1,2,*] (iD), Sergio Blanco-Trejo[1] (iD), Brenda Villarreal-Garza[1] (iD), Ana M. Picazo-Reina[1] (iD), Beatriz Tena-García[1] (iD), Ana Lara-Delgado[1], Manuel Lazo-Maestre[1] (iD), Francisco López-Benitez[1], Fernando Escobar-Reyero[1] (iD), Marta Álvarez-Cendrero[1] (iD), M. Luisa Calero-Ramos[1] (iD), Cayetana López-Ladrón[1] (iD), Cristina Alonso-González[1] (iD), Francisco Ropero-Romero[1] (iD), Leyre Andrés-Ustarroz[1] (iD), Marta Cuaresma-Giráldez[3], Mercedes Atienza-Ruiz[4,5] (iD), J. L. Cantero-Lorente[4,5] (iD), Alberto Moreno-Conde[3] (iD), Jesús Moreno-Conde[3] (iD), Serafín Sánchez-Gómez[1] (iD)

[1] Otorhinolaryngology Department, Virgen Macarena University Hospital, Seville, Spain / [2] Biomedical Engineering Group, University of Seville, Seville, Spain / [3] Innovation Unit, Virgen Macarena University Hospital, Seville, Spain / [4] Laboratory of Functional Neuroscience, Pablo de Olavide University, Seville, Spain / [5] Neurodegenerative Diseases Network Research Center (CIBERNED), Madrid, Spain

Correspondence:

*M. Amparo Callejón-Leblic

Virgen Macarena University Hospital, Avda. Dr. Fedriani 3, 41009 Seville, Spain.
Biomedical Engineering Group, University of Seville, Cº de los Descubrimientos s/n, 41092, Seville, Spain.
Email: mcallejon@us.es

Abstract

Hearing loss constitutes a major disability that hinders communication and quality of life. Recent evidence has uncovered its impact on cognitive decline, thus highlighting its multifactorial dimension and the outstanding need for creating multimodal hearing datasets that further cover clinical data across different health domains. The aim of this study is to develop a multi-collaborative database to systematically collect and analyze interdisciplinary data for audiological research, including auditory thresholds, speech tests, auditory evoked potentials, cognitive and quality-of-life tests, and medical images, among others. The database has been implemented in the Otorhinolaryngology Service of the Virgen Macarena University Hospital in Seville, integrated in the Intranet of the Andalusian Health Service, connected to the electronic patients' medical records. This database relies on open-source software and complies with national and international guidelines on data protection. A specific registry module has been designed to automatically import auditory thresholds and auditory evoked potentials from clinical devices into the platform. A mobile app has

Clinical implications

This work proposes a multimodal digital platform for collecting hearing data from various sources, including auditory thresholds, speech tests, auditory evoked potentials, and cognitive, health and quality-of-life tests, among others. The aim of our platform is a multi-collaborative and unified technical framework for clinicians and researchers in the audiology field. This will hopefully promote a joint interdisciplinary collaboration among researchers and clinicians to collect high-quality multimodal hearing research data. The availability of open and comprehensive hearing datasets may help better delineate hearing phenotypes and profiles, as well as identify further interdisciplinary associations.

also been implemented to collect questionnaires from patients remotely. A demo web version of the platform is freely available to the audiology community. The multimodal platform developed paves the way towards a multi-collaborative and unified framework for audiology research in Spain. Nevertheless, support from clinicians and healthcare stakeholders remains critical to develop more evidence and high-quality multimodal open datasets in hearing research.

Keywords

Introduction

According to the WHO, more than 1.5 billion individuals worldwide suffer from some type of hearing disorder, and at least 430 million will need medical attention (WHO, 2021). Hearing loss hinders the communication capability of a person, vital for an active and independent lifestyle (Dawes & Völter, 2023; Mosnier *et al.*, 2018). If left untreated, hearing loss may lead to loneliness and social isolation, as well as depression, reduced self-efficacy, poorer physical activity, and/or limited cognitive stimulation (Calvino *et al.*, 2022; Martinez-Amezcua *et al.*, 2021; Powell *et al.*, 2021). Also according to the WHO, the number of individuals suffering from dementia is set to increase from 55 million in 2019 to 139 million in 2050 (Gauthier *et al.*, 2022). Recent evidence has revealed that these two conditions are closely related, with hearing loss being a modifiable risk factor that may prevent or delay the onset of approximately 8% of dementia cases. Other contributing factors include comorbidities, low levels of education, and lifestyle habits (Livingston *et al.*, 2020; Uchida *et al.*, 2019a; Loughrey *et al.*, 2018).

Hearing devices, such as hearing aids and cochlear implants, have been identified as protective factors by delaying the onset of dementia and maintaining independence, social engagement, and quality of life (Völter *et al.*, 2022a; Castiglione *et al.*, 2019; Wayne & Johnsrude, 2015; Lin *et al.*, 2013). Given the complexity of hearing loss, interdisciplinary approaches and interventions are necessary to delineate both the audiological and neurocognitive profiles of patients (Andries *et al.*, 2023), as well as their intrinsic interdependence (Giallini *et al.*, 2023a; Dazert *et al.*, 2020). It is well known that hearing loss is related to a variety of genetic and clinical risk factors including viral infections or pathogens, chronic disease (e.g., hypertension, cardiovascular disease, diabetes), ototoxic drugs, smoking, head trauma, noise exposure, ear related conditions (e.g., Meniere's disease), and nutritional deficiencies, among others (Choi *et al.*, 2024; Daniel, 2007; Nieman & Oh, 2020). While some of these factors can be partially managed through prevention measures (Lin *et al.*, 2011; Wang *et al.*, 2020), hearing loss has also proved to be a multifactorial condition linked to environmental, socioeconomic, and lifestyle factors such as physical activity and frailty, smoking, malnutrition, and social isolation (He *et al.*, 2019; Lawrence *et al.*, 2020; Rutherford *et al.*, 2018). Thus, there is a need to evaluate clinical and behavioral variables that impact a patient's quality of life, along with the benefits of developing multifactorial treatment approaches (Brewster *et al.*, 2022; Goodwin *et al.*, 2023b).

Over the past decade, advances in audiological innovations have been coupled with the implementation of digital health techniques, including big data analysis and artificial intelligence (AI) algorithms (Gajecki & Nogueira, 2022; Rajpurkar *et al.*, 2022; Sanchez-Lopez *et al.*, 2020; Waring *et al.*, 2020) as well as mobile health (m-health) resources (Almufarrij *et al.*, 2023; Swanepoel, 2023). According to Wasmann and colleagues (Wasmann *et al.*, 2021), computational audiology has the potential to set up a new standard paradigm for modern audiological care delivery. With the aim of systematically gathering hearing data, various research groups are actively working on the definition of audiological registries and datasets, paving the way for more comprehensive and personalized hearing therapies (Huang *et al.*, 2021; Sánchez-López *et al.*, 2019; Chen *et al.*, 2017; Theunisse *et al.*, 2014; Berrettini *et al.*, 2011). However, there is an outstanding need to create multimodal digital platforms and datasets for collecting a wide range of variables across different health domains within hearing research. The main objective of this work is to develop a multi-collaborative database to systematically collect and analyze multimodal hearing data. This database aims to offer support to clinical and translational audiological research, both nationally and internationally.

Project objectives and requisites

This project aims to develop and validate a multimodal database for information management in the fields of audiology and hearing research. The database will facilitate the integration of data from diverse sources, including auditory thresholds, speech tests, auditory evoked potentials (AEPs), cognitive, health and quality-of-life tests, and medical images, among others. The objectives of this database are:

- To facilitate systematic collection and data recruitment based on real-world clinical conditions
- To provide researchers with access to and management of multimodal information based on relevant state-of-the-art variables on hearing research
- To offer add-on tools for the advanced analysis, visualization and exploitation of the data, as well as their integration with available clinical information
- To provide mHealth solutions, defined as medical and public health practice supported by mobile devices (Ryu, 2012), for remote follow-up and data collection from hearing loss patients
- To promote collaboration between researchers across different organizations in multicenter studies by complying with regulations, security and confidentiality issues outlined in international guidelines on data protection, *datasets for datasheets* specifications (Gebru *et al.*, 2021) and FAIR principles for effective data reuse (Wilkinson *et al.*, 2016).

Platform system and requirements

The multimodal database is based on the ITCBio platform system (*HUVM Website*, n.d.; Moreno-Conde *et al.*, 2019), connected to the patient's electronic health records (EHR) through the Intranet of the Andalusian Health Service (SAS) named "Diraya" (Muñoyerro-Muñiz *et al.*, 2020). ITCBio platform implements a context management system for research, based on open-software and interoperability standards such as ISO 13606 and CDISC (Moreno-Conde *et al.*, 2022). It also complies with national and international guidelines on data protection GDPR (EU 2016/679; Publications Office of the European Union, 2016). The use of the platform and recruitment of

health data has been approved by an Ethics Committee under the studies PEIBA 1735-N-22 and 1752-N-22.

Platform Implementation and Validation

Platform Implementation: Architecture and Modules

A new multimodal platform combining different types of data has been implemented in the Otorhinolaryngology (ORL) Service of the Virgen Macarena University Hospital (VMUH), in Seville. The platform incorporates three main modules: i) a registration module, where the information is recorded, pseudo-anonymized, and followed up; ii) a data analysis module, where different subsets of patients can be selected and subjected to various analytical and statistical tools, ranging from basic descriptive statistical analysis to more complex signal processing pipelines and ML algorithms, and iii) a mobile app that allows communication with the patient, who can complete questionnaires remotely. **Figure 1** shows the conceptual architecture of the platform, along with the main data sources collected and the main modules deployed. A demo web version of the platform has been implemented and can be freely accessed through the links in Appendix A.

Given that the main objective of this work is to create a multimodal database to collect comprehensive audiological datasets, various registration modules have been developed, including a specific module dedicated to automatically collect audiometric thresholds and speech perception tests. Measures of spectral and temporal resolution, reported to be relevant for speech intelligibility (Relaño-Iborra & Dau, 2022; Archer-Boyd *et al.*, 2018) have also been incorporated. A test battery has been designed to quantify and identify the main cognitive domains involved in patients with different hearing profiles and rehabilitation strategies. This battery has been selected with the help of experts in the fields of cognitive impairment and Alzheimer's disease (Fernandez-Alvarez *et al.*, 2023). An extensive collection of questionnaires evaluating self-reported hearing difficulties, general health, personality, and lifestyle has been included to further understand the overall impact and association of hearing loss in real-life scenarios, which are difficult to emulate in clinical settings. A module has been designed to automatically import
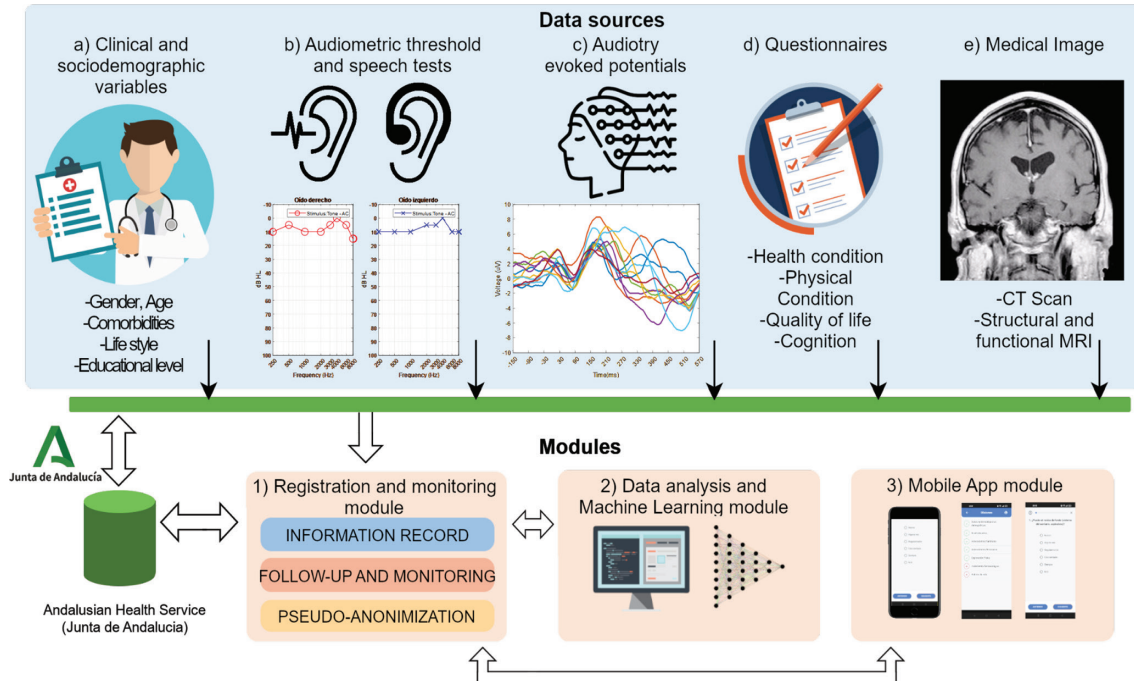
Figure 1. Functional architecture of the platform. The platform consists of three main modules: 1) a registration and monitoring module, where the information is recorded, followed-up and pseudo-anonymized; 2) a data analysis and machine learning module, and 3) an app module that allows the patient to remotely complete questionnaires through their mobile phone. The data sources of the multimodal platform include: a) clinical and sociodemographic variables, b) audiology data, such as audiometric thresholds and speech tests, c) auditory evoked potentials, d) health and cognitive questionnaires and e) medical image data.

electrophysiological auditory brainstem and cortical responses, providing information about the functioning of both peripheral and central auditory pathways. Additionally, a dedicated module and pipeline for processing medical images of the ear and the brain has also been implemented. Further details of the variables, questionnaires, and tests employed in the platform are described below. The three main modules of the platform, shown in **Figure 1**, are explained in detail as follows:

1. **Registration and monitoring module**
   The primary registration and monitoring module facilitates the recording of new data into the platform and its integration with the patient's clinical records (e.g., demographic, diagnosis, laboratory, prescriptions, etc.). S1_Figure 1A in Supplementary Material 1 shows a screenshot of the main page of the database registration module. The data sources that feed the registration and monitoring module in the database are explained below.
   a. *Clinical and sociodemographic data*
      Both clinical and demographic data can be registered, including variables such as age, gender, height, weight, lifestyle habits (smoking and alcohol intake),

educational level, urban or rural habitat, etc. Relevant clinical data, such as previous comorbidities and diagnostics, surgical interventions, blood tests, and pharmacological history, are obtained from a patient's EHR by authorized professionals. These are based on predefined terms according to the International Classification of Diseases (CIE 10a; *eCIE-Maps - Documentación (Sanidad.Gob.Es); ICD- 10 - CM International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) (Cdc.Gov)*). Additionally, socioeconomic variables are related to a patient's access to appropriate hearing rehabilitation (Zhan *et al.*, 2020). S1_Figure 1B shows a view of the clinical and demographic data section. As part of surgical information, we have also codified the SAMEO-ATO system for categorizing tympanomastoid interventions in our platform (Yung *et al.*, 2018).
   b. *Audiometric thresholds and speech tests*
      This section allows clinicians to gather information regarding the type of hearing loss, etiology, duration of hearing deprivation, and details of the hearing device used by the patient (see S1_Figure 1C). Audiological data are collected and represented on the platform, including hearing thresholds for air and bone conduction and speech tests in quiet

and in noise (**Figure 2**). A dedicated data registration module has been developed to automatically import and read the XML files exported from the Otoaccess software and audiometers available in our clinical setting (Interacoustics A/S, Middelfart, Denmark; sS1_Figure 1D). The codes implemented for automatically importing the XML files with the raw audiometry and speech tests are accessible as described in Appendix B. Additionally, temporal and spectral resolution is measured through SMRT test and is also collected in the platform (Zhang *et al*. 2023, Aronoff & Landsberger, 2013). Categories of auditory performance (CAP) scale has also been implemented (Archbold *et al.*, 1998).

c. *Auditory evoked potentials (AEPs)*
The digital platform also allows the uploading of different types of auditory evoked potentials (AEPs), such as brainstem and cortical AEPs. A dedicated data registration module has been designed to specifically import the XML files containing the epoch signals exported by the Eclipse EP25 system (Interacoustics A/S, Middelfart, Denmark). In addition to the raw EEG data, this module automatically reads and collects specific information related to protocol stimuli and setup, such as frequency, intensity, noise filters or gain (Callejón-Leblic *et al.*, 2023). The source codes implemented are available in Appendix B.

d. *Tests and questionnaires*
An extensive collection of validated tests and questionnaires (see Supplementary Material 2) has been implemented in the platform, covering different health domains including hearing quality, tinnitus, balance and dizziness, cognition, mood disorder, quality of life, physical activity, sleep quality, and nutrition. The selection of these tests is based on the

research aim of evaluating different health domains and their intrinsic interdependence or impact on hearing (Uchida *et al.*, 2019b; Viergever *et al.*, 2021). To select them, a comprehensive review of state-of-the-art of tests and questionnaires used in the literature has been conducted.

• Hearing-related questionnaires
Different questionnaires were selected to assess hearing quality. For example, SSQ-12 (Cañete *et al.*, 2022) and HISQUI-19 (Calvino *et al.*, 2016) were selected to assess spatial audition, sound perception, and hearing quality. Specific tests, such as NCIQ (Sanchez-Cuadrado, Gavilan, *et al.*, 2015; Sanchez-Cuadrado, Lassaletta, *et al.*, 2015) and APSQ (Billinger-Finke *et al.*, 2020) were selected to evaluate quality of life and satisfaction with the sound processor in cochlear implant users, respectively. However, APHAB (Solarte *et al.*, 2016) was selected to evaluate subjective hearing in users with hearing aids. On the other hand, tinnitus and dizziness constitute relevant variables in patients with hearing disorder (Wang *et al.*, 2020), therefore THI (Newman *et al.*, 1996) and DHI (Formeister *et al.*, 2020) questionnaires were selected for this purpose.

• Cognitive tests
A comprehensive test battery has been designed to evaluate various cognitive domains based on previous studies that had significantly associated hearing loss with a poorer cognitive performance (Andries *et al.*, 2023; Calvino *et al.*, 2022; Castiglione *et al.*, 2016; Claes *et al.*, 2018; Giallini *et al.*, 2023b; Huber *et al.*, 2020; Völter *et al.*, 2022b), particularly affecting cognitive domains such as processing speed, cognitive flexibility (Huber *et al.*, 2020; Völter *et al.*, 2017), non-verbal reasoning (Kramer *et al.*, 2018;

Figure 2. Visualization of the imported patients' audiometries in the database.
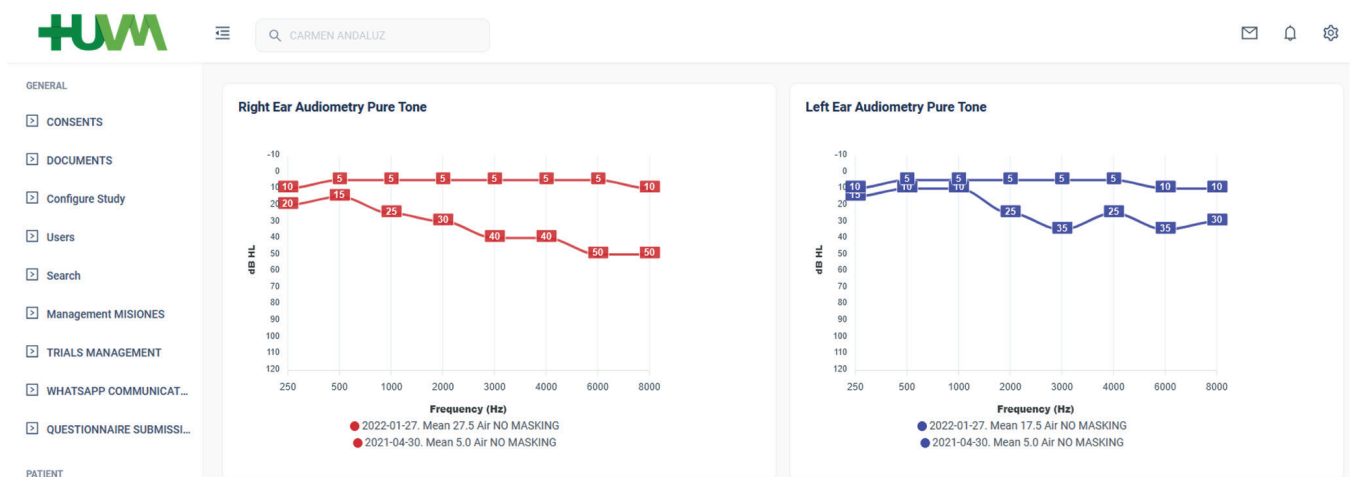
Zhang *et al.*, 2023), spatial and episodic memory (Jayakody *et al.*, 2018) and working memory (Kramer *et al.*, 2018). Therefore, the cognitive battery developed in this platform consists of a compilation of tests that evaluate general cognition (e.g. MoCA), memory (e.g. DSST), language (eg. denomination), verbal fluency (e.g. semantic and phonemic), attention (e.g. N-back), and different executive functions (e.g. TMT A&B, Matrix, Stroop Visual).

- Psychophysical and general health questionnaires
Hearing loss has also been related to mood disorders and personality changes, especially in older adults (Brewster *et al.*, 2022; Cuda *et al.*, 2024; Goodwin *et al.*, 2023a). Through specific tests, our platform also evaluates the mood disorders most commonly reported in patients with hearing loss: depression (GDS-15), loneliness (JLS), apathy (DAS) and anxiety (BAI). These psychosocial consequences of hearing loss have also been shown to indirectly contribute to cognitive impairment in these patients (Babajanian & Gurgel, 2022; Calvino *et al.*, 2022). Furthermore, other physical and psychosocial conditions and health variables associated with hearing loss (Holman *et al.*, 2019; Wells *et al.*, 2020), such as quality of life (e.g., HUI3-S and GBI), physical activity (e.g. PASE, AFAI), activities of daily living (e.g. IADL, BI), sleep quality (e.g. PSQI, ESE) and nutrition (e.g. ADM) have also been taken into consideration in our platform.

- *e. Clinical image data*
A module for uploading medical images, including computed tomography (CT) scans of the ear, and structural and functional magnetic resonance images (MRI) of the brain, is being developed (Callejón-Leblic *et al.*, 2024; Lazo-Maestre *et al.*, 2024; Callejón-Leblic & Miranda, 2020). The objective is to collect data regarding cochlear anatomy to analyze its relation to hearing performance, especially relevant in cochlear implant recipients (Fitzhugh & Pa, 2023; Giroud *et al.*, 2021; Ha *et al.*, 2020; Rosemann & Thiel, 2020; Wang *et al.*, 2022). Additionally, neuroimaging biomarkers will help us analyze possible associations between hearing status and cognition (Cantero *et al.*, 2018; Ha *et al.*, 2020; Rosemann & Thiel, 2020; Giroud *et al.*, 2021). Future implementations will explore the use of standardized architectures, such as BIDS (Gorgolewski *et al.*, 2016), to process, store, and share neuroimaging data.

Further information, including the definition of all the variables implemented in the platform, can be found in Supplementary Material 3.

2. Data analysis and machine learning module

The collected data are properly labelled and pseudo-anonymized. Basic analytical tools have been implemented in the platform to display descriptive statistics and relevant trends in data. Some of the add-ons already implemented in the platform include tools for visualizing the distribution and evolution of collected variables, such as audiometries and speech tests over time, depending on parameters such as the audiological profile of patients (type of hearing loss, etiology, etc.). In addition, a dedicated exportation data module has been implemented to export CSV data for external analysis. Expert researchers can also access the database with secure protocols through SQL queries, or process data directly using different languages based on Python, Matlab/Octave or R. Based on the multimodal data gathered, a dedicated pipeline has been coded to conduct various data analysis including descriptive statistics, correlation analysis, regression modelling, and ML algorithms, among others. Post-processing pipelines have been developed to monitor data quality by evaluating different metrics such as integrity, completeness, consistency, and accuracy (Schmidt *et al.*, 2021; Kahn *et al.*, 2016; Chen *et al.*, 2014).

3. Mobile App

The app module systematically collects information on health status and quality of life through validated questionnaires, which can be personalized and remotely completed by patients on their mobile phones. It also provides access to educational information and content provided by clinical specialists. A screenshot of some of the different facilities provided by the mobile app can be seen in **Figure 3**.

Platform Validation

From a functional perspective, the multimodal database provides advanced capabilities that support the primary research processes present in our audiological clinical setting. Currently, the platform database is being validated by specialized clinicians and researchers by handling and collecting data from patients with hearing loss in the context of a multicentric project funded by the Spanish Ministry of Economy and Digital Transformation under project AI4HA from Misiones I+D in Artificial Intelligence 2021, which aims to develop open databases and IA algorithms to early
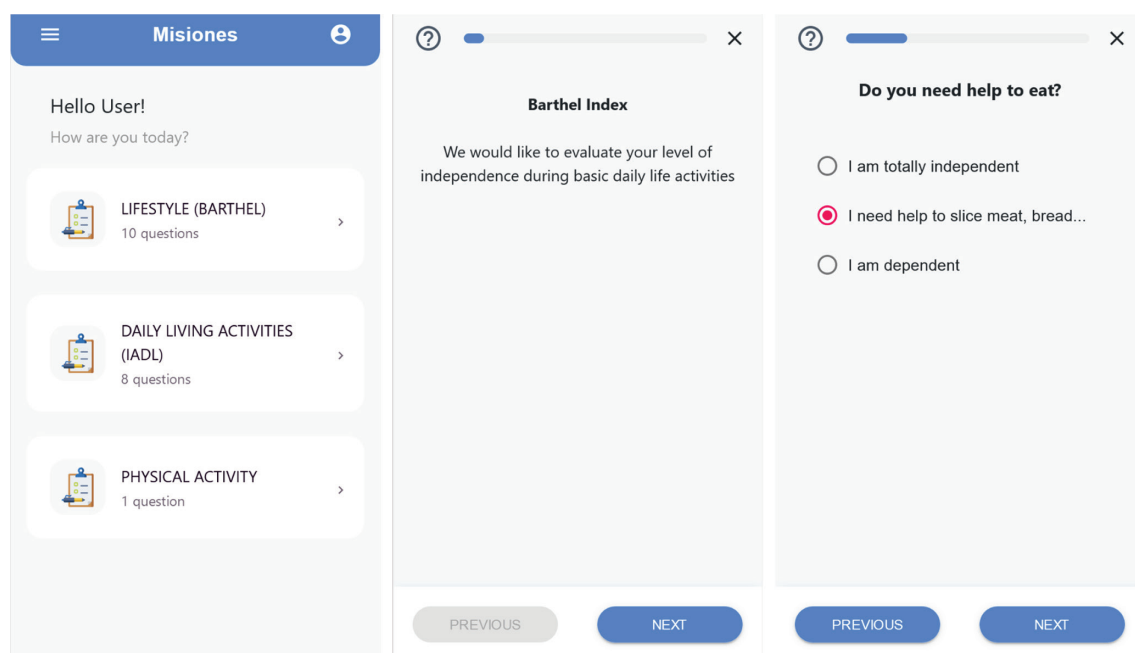
Figure 3. Mobile app developed to collect health information and questionnaires remotely

detect and treat diseases with high prevalence in aging. At present, data from more than 200 patients and healthy controls has been collected.

Discussion

In this project, a unified framework integrating information from various audiological and health data sources relevant in both clinical and translational research fields in audiology has been developed. Within this process, a dedicated multimodal database oriented towards the integration of cross-sectional data from different health domains and devices has been implemented, thereby facilitating the customization of various add-ons and data analytics tools. The clinical suitability and applicability of the implemented tests and measures, as well as the possibilities that this platform brings to further standardization, are discussed below.

Towards the standardization and collection of interdisciplinary audiological datasets
In the field of audiology, an issue remains with comparatively small and disaggregated datasets generated at only few centers. Tonal audiometry still represents, in many cases, the clinical gold standard and the primary measure used to characterize

individual hearing loss. It also serves as the basis for treatments involving hearing solutions such as hearing-aid and cochlear implant fitting. However, hearing loss is also related to poor spectral resolution and reduced intelligibility, especially in noisy real-life scenarios, where difficulties in perception and clarity are not reflected by an audiogram (Sanchez-Lopez *et al.*, 2020). Indeed, the consequences of poor speech understanding may also lead to a decline in cognitive capabilities and quality of life. These factors are relevant variables, not only for fully understanding the impact of hearing loss on a patient's daily life (Raymond *et al.*, 2023; Gurgel *et al.*, 2022), but also for guiding and improving personalized hearing aid and cochlear implant fittings (Cone-Wesson & Wunderlich, 2003; Rader *et al.*, 2023; Van Dun *et al.*, 2016; Van Opstal & Noordanus, 2023). However, these variables are not always comprehensively assessed in daily clinical practice, possibly due to the lack of suitable, clinically applicable tests and materials, as well as standardized protocols across centers. The issue of the coexistence of different data management systems and audiological materials persists nationally and internationally, hindering our ability to provide open and statistically significant datasets for re-use and validation (Ait Abdelouahid *et al.*, 2023).

Initial efforts to provide comprehensive normative audiological data across five international institutions and countries were made by Van Esch *et al.*

(2013), who proposed a battery of tests termed the "preliminary auditory profile". This profile includes various measures of loudness perception, listening effort, speech perception, spatial hearing, spectral and temporal resolution, cognitive capabilities, and self-reported hearing disability. The battery was evaluated in an international multi-center study with two normal-hearing and hearing-impaired subgroups and showed good test-retest variability across centers, although with some differences in language-specific tests across countries. More recently, Buhl *et al.* (2019, 2020) proposed Common Audiological Functional Parameters (CAFPAs) as a compact but statistically meaningful representation of audiological diagnostics and treatments. This aims to provide real-world clinic databases with a universally applicable audiological datasheet based on Bayesian networks, which is partially robust against the choice of different or redundant audiological tests. A first preliminary set of ten CAFPAs representing information on hearing thresholds, supra-threshold deficits, binaural, neural and cognitive properties of the human auditory system, as well as socio-economic variables, was proposed and validated. As concluded by the authors, further collection of data from real clinical databases should be conducted to train and validate the CAFPA concept or other ML-based diagnostic tool approaches.

Motivated by these previous studies, our multimodal platform aims to integrate various health domains related to auditory function into a single dedicated database. This facilitates the systematic collection of heterogeneous data for hearing research. With this purpose, we have met two technical requirements that we considered important for automatization. Firstly, we integrated XML files containing the audiometric thresholds and raw auditory EEG signals from the Interacoustics proprietary medical devices available in our hospital into our platform. Detailed documentation of the source codes used to import and read these XML files is provided in Appendix B, with the intention of aiding other clinics and research centers using the same equipment. Another milestone was achieved through the integration of this multimodal data with clinical information derived from patients' EHRs. This was accomplished by successfully connecting the multimodal platform into the Intranet of the Andalusian Health Service. This has allowed us to use international normalized predefined CIE 10 diagnostic and treatments codes, hopefully enabling data accessibility across primary and tertiary care centers in Andalusia and Spain. Moreover, the platform complies with the international standard ISO 13606 (*ISO 13606 Standard - EHR Interoperability*; Santos *et al.*, 2010), which provides specifications for the structure of clinical information to be shared between different EHR systems. This standard aims to improve interoperability and effective data exchange (Frid *et al.*, 2023). A previous work by Moreno-Conde *et al.* (2022) provides a semantic infrastructure description of certain terminology subsets and forms included in the multimodal platform developed, thus specifying its associated attributes and how they are mapped into ISO 13606.

Clinical applicability and suitability

A wide variety of tests has been implemented for subjects ranging from normal hearing to hearing impaired, including those with different hearing devices such as hearing aids or cochlear implants. In this way, different test implementations have been defined and selectively adapted to different hearing profiles. A critical factor for clinical applicability is the evaluation time. Currently, the complete battery of tests implemented in our platform takes up to three hours (divided into two sessions): one session of 60 minutes for audiological assessment (auditory thresholds, speech in noise test, auditory electrophysiological measures, etc.) and a second session of about 120 min for cognitive evaluation, including breaks to ensure that patients do not become fatigued. According to our experience, patients with bilateral severe hearing loss, low levels of education and/or advanced age, find it more difficult to complete the entire cognitive battery. Additionally, it is essential to ensure the communication and comprehension of instructions, aided by visual information (i.e. slides, dictation-writing tools) for a more fluid communication. We acknowledge that obtaining a full auditory and cognitive profile across different domains is time-consuming and probably not suitable for clinical use. Conversely, the implemented database and mobile app, allow us to partly automate the process and collect data from patients remotely, successfully integrating it into our daily clinical setup. However, we have also found that some subjects, especially older patients, often report difficulties when using their mobile phones, necessitating in-person interviews. Thus, it is important to evaluate and detect the most relevant cognitive domains and variables related to hearing loss, with the aim of developing more

versatile screening tests that could be strategically useful for evaluating patients in daily clinical practice.

Ethics, privacy, and quality issues

As new AI-based clinical decision support systems are increasingly being incorporated into healthcare, remaining challenges such as data quality, bias, ethics, and privacy regulations become more evident and critical. Technical challenges exist to comply with the GDPR while still adhering to FAIR principles. In our platform, pseudonymization and anonymization are addressed at two different levels. Internally, data extraction performed by the researchers is done in a pseudo-anonymized manner using a patient identifier encryption mechanism based on NUHSA codes (Protti, 2007). The patient identifier (NUHSA) is replaced by a code generated using the SHA-512 encryption algorithm (Algredo-Badillo *et al.*, 2012), with a specific seed unknown to the researchers. This mechanism ensures technical and functional separation between the research team and those who carry out the pseudonymization. It includes specific measures to prevent the re-identification of the subjects and unauthorized access to the data, as established in Section d) of Organic Law 3/2018, 5 December, GDPR, for the use of pseudonymized personal data for health research purposes (*Ley Orgánica 3/2018, de 5 de Diciembre, de Protección de Datos Personales y Garantía de Los Derechos Digitales*). Pseudonymization still allows the re-identification of patients, if necessary (only by authorized clinical and technical staff of the hospital), due to health risks or findings uncovered during the research. However, externally, open datasets available to the scientific community will be fully anonymized. For anonymization we are currently implementing hash and k-anonymization techniques in our data processing pipelines to ensure robust anonymization of data, in line with recommendations provided by the Spanish Agency on Data Protection (*Anonimización y Seudonimización | AEPD*). Regarding neuroimaging, further efforts should be made to improve standardization and compatibility across different clinics and centers. Currently, we are planning to implement widely accepted standards like Brain Imaging Data Structure, BIDS, (Boré *et al.*, 2023; Gorgolewski *et al.*, 2016), which would help in the FAIRification process of our neuroimaging dataset. Given the sensible nature of the collected information, our research group must carefully study and analyze neuroimaging privacy and usage policies, as

discussed by other research groups when creating a neuroimage data repository (Jwa & Poldrack, 2022).

High-quality and non-biased datasets are crucial when used as input for IA algorithms that support clinical decisions (Vicente & Matute, 2023). Different methods have been proposed in the literature to quantify the quality of collected data (Bertossi & Geerts, 2020; Holzinger *et al.*, 2020). Some of these methods are based on rigorous methodological queries to the database, safeguarding data consistency (Levman *et al.*, 2023). Other relevant aspects, especially those recommended for ML models, rely on concepts such as causality or explainability (Livshits *et al.*, 2021; Linardatos *et al.*, 2020). Bias analysis and mitigation must be carried out during the development of ML models to guarantee the quality and accuracy of the predictions (Siddique *et al.*, 2023). All these concepts will be properly taken into consideration to quantify the quality of our data when developing new updates for our multimodal platform.

Limitations and future developments

We provide in Supplementary Material 2 and 3 a description of the variables implemented in our platform, offering an overall view of its capability to collect multimodal data from different sources. Our objective is to make such data accessible for research purposes by the entire scientific community in audiology. Thus, we plan to provide the data collected on our platform in open repositories. Regarding dataset specifications, we have followed the template "datasheets for datasets" provided by Gebru *et al.* (2021) which aims to facilitate data sharing and communication between dataset creators and consumers. The platform is scalable and permits new questionnaires and data sections to be implemented according to new research goals and projects. Indeed, future developments will implement other relevant variables, such as distortion product otoacoustic emissions (DPOE) or cochlear implant fitting parameters, among others.

Regarding interoperability, we must acknowledge that our platform is still in the initial phase of development, with heterogeneity being only partially addressed. Additional efforts will be required to evolve towards the adoption of other specifications such as OpenEHR (Haarbrandt *et al.*, 2018; openEHR website) or HL7 (HL7 Diagnostic Audiology Reporting Implementation Guide), which will hopefully make our developments more versatile and accessible. Finally, we also aim to incorporate FAIR-compliant metadata

to generate research data collections, which would allow universal reusability of datasets for open science research (Wilkinson *et al.*, 2016).

In summary, we present the multimodal platform developed as a step towards a multi-collaborative database in Audiology. Nevertheless, in addition to the need for a common framework and technical platform, support from clinicians in audiology and healthcare stakeholders is crucial. The database could be increasingly used by various agents and organizations at a national level, including researchers and industrial partners, to develop more evidence and high-quality multimodal open datasets for audiological research.

Conclusion

In this work, a multimodal digital platform and database have been developed and implemented in the Otorhinolaryngology Service of the Virgen Macarena University Hospital in Seville to systematically collect audiological and health data from different sources, including auditory thresholds, speech tests, auditory evoked potentials, medical images, quality of life, and cognitive tests, among others. A demo web version of the platform, along with all variables and sections implemented, and the coded scripts to automatically import and read audiological data, are freely available for the clinical and scientific community in audiology. Further joint efforts are still required to generate larger, comprehensive, and open datasets for hearing research.

References

AEPD. (October 6th 2021). *Anonimización y seudonimización*. Retrieved April 18, 2024, from https://www.aepd.es/prensa-y-comunicacion/blog/anonimizacion-y-seudonimizacion

Ait Abdelouahid, R., Debauche, O., Mahmoudi, S., & Marzak, A. (2023). Literature Review: Clinical Data Interoperability Models. *Information*, *14*(7), 364. https://doi.org/10.3390/info14070364

Algredo-Badillo, I., Morales-Sandoval, M., Feregrino-Uribe, C., & Cumplido, R. (2012). Throughput and efficiency analysis of unrolled hardware architectures for the sha-512 hash algorithm. *2012 IEEE Computer Society Annual Symposium on VLSI*, 63–68. https://doi.org/10.1109/ISVLSI.2012.63

Almufarrij, I., Dillon, H., Dawes, P., Moore, D. R., Yeung, W., Charalambous, A.P., Thodi, C., & Munro, K. J. (2023). Web-and app-based tools for remote hearing assessment: A scoping review. *International Journal of Audiology*, *62*(8), 699–712. https://doi.org/10.1080/14992027.2022.2075798

Andries, E., Bosmans, J., Engelborghs, S., Cras, P., Vanderveken, O. M., Lammers, M. J. W., Van de Heyning, P. H., Van Rompaey, V., & Mertens, G. (2023). Evaluation of Cognitive Functioning Before and After Cochlear Implantation in Adults Aged 55 Years and Older at Risk for Mild Cognitive Impairment. *JAMA Otolaryngology - Head & Neck Surgery*, *149*(4), 310-316. https://doi.org/10.1001/jamaoto.2022.5046

Archbold, S., Lutman, M. E., & Nikolopoulos, T. (1998). Categories of auditory performance: Inter-user reliability. *British Journal of Audiology*, *32*(1), 7–12. https://doi.org/10.3109/03005364000000045

Archer-Boyd, A. W., Southwell, R. V., Deeks, J. M., Turner, R. E., & Carlyon, R. P. (2018). Development and validation of a spectro-temporal processing test for cochlear-implant listeners. *The Journal of the Acoustical Society of America*, *144*(5), 2983–2997. https://doi.org/10.1121/1.5079636

Aronoff, J. M., & Landsberger, D. M. (2013). The development of a modified spectral ripple test. *The Journal of the Acoustical Society of America*, *134*(2), EL217–EL222. https://doi.org/10.1121/1.4813802

Babajanian, E. E., & Gurgel, R. K. (2022). Cognitive and behavioral effects of hearing loss. *Current Opinion in Otolaryngology & Head and Neck Surgery*, *30*(5), 339–343. https://doi.org/10.1097/MOO.0000000000000825

Berrettini, S., Arslan, E., Baggiani, A., Burdo, S., Cassandro, E., Cuda, D., Dinelli, E., Filipo, R., Mancini, P., Martini, A., & others. (2011). A registry for the collection of data in cochlear implant patients. *Acta Otorhinolaryngologica Italica*, *31*(5), 328.

Bertossi, L., & Geerts, F. (2020). Data quality and explainable AI. *Journal of Data and Information Quality (JDIQ)*, *12*(2), 1–9.

Billinger-Finke, M., Bräcker, T., Weber, A., Amann, E., Anderson, I., & Batsoulis, C. (2020). Development and validation of the audio processor satisfaction questionnaire (APSQ) for hearing implant users. *International Journal of Audiology*, *59*(5), Article 5.

Boré, A., Guay, S., Bedetti, C., Meisler, S., & GuenTher, N. (2023). *Dcm-2Bids (Version 3.1.1)* [Computer software]. https://doi.org/10.5281/zenodo.8436509

Brewster, K. K., Deal, J. A., Lin, F. R., & Rutherford, B. R. (2022). Considering hearing loss as a modifiable risk factor for dementia. *Expert Review of Neurotherapeutics*, *22*(9), Article 9. https://doi.org/10.1080/14737175.2022.2128769

Buhl, M., Warzybok, A., Schädler, M. R., Lenarz, T., Majdani, O., & Kollmeier, B. (2019). Common Audiological Functional Parameters (CAFPAs): Statistical and compact representation of rehabilitative audiological classification based on expert knowledge. *International Journal of Audiology*, *58*(4), 231–245. https://doi.org/10.1080/14992027.2018.1554912

Buhl, M., Warzybok, A., Schädler, M. R., Majdani, O., & Kollmeier, B. (2020). Common Audiological Functional Parameters (CAFPAs) for single patient cases: Deriving statistical models from an expert-labelled data set. *International Journal of Audiology*, *59*(7), 534–547. https://doi.org/10.1080/14992027.2020.1728401

Callejón-Leblic, M., & Miranda, P. C. (2020). A computational parcellated brain model for electric field analysis in transcranial direct current stimulation. In Makarov, S.N., Noetscher, G.M., Nummenmaa, A. (eds) *Brain and Human Body Modeling 2020* (pp. 81). Springer, Cham. https://doi.org/10.1007/978-3-030-45623-8_5

Callejón-Leblic, M. A., Barrios-Romero, M. M., Kontides, A., Sánchez-Gómez, S., & Beynon, A. J. (2023). Electrically evoked auditory cortical responses elicited from individually fitted stimulation parameters in cochlear implant users. International Journal of Audiology, 62(7), 650-658.

Callejón-Leblic, M. A. *et al.* (2024). A full-head model to investigate intra and extracochlear electric fields in cochlear implant stimulation. *Phys. Med. Biol.* 69, 155010. https://doi.org/10.1088/1361-6560/ad5c38

Calvino, M., Gavilán, J., Sánchez-Cuadrado, I., Pérez-Mora, R. M., Muñoz, E., & Lassaletta, L. (2016). Validation of the Hearing Implant Sound Quality Index (HISQUI19) to assess Spanish-speaking cochlear implant users' auditory abilities in everyday communication situations. *Acta Oto-Laryngologica*, 136(1), Article 1.

Calvino, M., Sánchez-Cuadrado, I., Gavilán, J., Gutiérrez-Revilla, M. A., Polo, R., & Lassaletta, L. (2022). Effect of cochlear implantation on cognitive decline and quality of life in younger and older adults with severe-to-profound hearing loss. *European Archives of Oto-Rhino-Laryngology: Official Journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS): Affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery*, 279(10), Article 10. https://doi.org/10.1007/s00405-022-07253-6

Cantero, J. L., Atienza, M., Sanchez-Juan, P., Rodriguez-Rodriguez, E., Vazquez-Higuera, J. L., Pozueta, A., Gonzalez-Suarez, A., Vilaplana, E., Pegueroles, J., Montal, V., & others. (2018). Cerebral changes and disrupted gray matter cortical networks in asymptomatic older adults at risk for Alzheimer's disease. *Neurobiology of Aging*, 64, 58–67. https://doi.org/10.1016/j.neurobiolaging.2017.12.010

Cañete, O. M., Marfull, D., Torrente, M. C., & Purdy, S. C. (2022). The Spanish 12-item version of the Speech, Spatial and Qualities of Hearing scale (Sp-SSQ12): Adap-tation, reliability, and discri-minant validity for people with and without hearing loss. *Disability and Rehabilita-Tion*, 44(8), Article 8.

Castiglione, A., Benatti, A., Velardita, C., Favaro, D., Padoan, E., Severi, D., Pagliaro, M., Bovo, R., Vallesi, A., Gabelli, C., & Martini, A. (2016). Aging, Cognitive Decline and Hearing Loss: Effects of Auditory Rehabilitation and Training with Hearing Aids and Cochlear Implants on Cognitive Function and Depression among Older Adults. *Audiology & Neuro-Otology*, 21 Suppl 1, 21–28. https://doi.org/10.1159/000448350

Castiglione, A., Casa, M., Gallo, S., Sorrentino, F., Dhima, S., Cilia, D., Lovo, E., Gambin, M., Previato, M., Colombo, S., Caserta, E., Gheller, F., Giacomelli, C., Montino, S., Limongi, F., Brotto, D., Gabelli, C., Trevisi, P., Bovo, R., & Martini, A. (2019). Correspondence Between Cognitive and Audiological Evaluations Among the Elderly: A Preliminary Report of an Audiological Screening Model of Subjects at Risk of Cognitive Decline With Slight to Moderate Hearing Loss. *FRONTIERS IN NEUROSCIENCE*, 13. https://doi.org/10.3389/fnins.2019.01279

CDC National Center for Health Statistics (n.d.). *ICD- 10—CM International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) (cdc.gov)*. Retrieved April 18, 2024, from https://www.cdc.gov/nchs/icd/icd-10-cm.htm

Chen, H., Hailey, D., Wang, N., & Yu, P. (2014). A review of data quality assessment methods for public health information systems. *International Journal of Environmental Research and Public Health*, 11(5), 5170–5207. https://doi.org/10.3390/ijerph110505170

Chen, S. Y., Grisel, J. J., Lam, A., & Golub, J. S. (2017). Assessing cochlear implant outcomes in older adults using HERMES: A national web-based database. *Otology and Neurotology*, 38(10), e405–e412. https://doi.org/10.1097/MAO.0000000000001575

Choi, J. S., Adams, M. E., Crimmins, E. M., Lin, F. R., & Ailshire, J. A. (2024). Association between hearing aid use and mortality in adults with hearing loss in the USA: A mortality follow-up study of a cross-sectional cohort. *The Lancet Healthy Longevity*, 5(1), e66–e75. https://doi.org/10.1016/S2666-7568(23)00232-5

Claes, A. J., Van de Heyning, P., Gilles, A., Van Rompaey, V., & Mertens, G. (2018). Cognitive Performance of Severely Hearing-impaired Older Adults Before and After Cochlear Implantation: Preliminary Results of a Prospective, Longitudinal Cohort Study Using the RBANS-H. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 39(9), e765–e773. https://doi.org/10.1097/MAO.0000000000001936

Cone-Wesson, B., & Wunderlich, J. (2003). Auditory evoked potentials from the cortex: Audiology applications. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 11(5), 372–377. https://doi.org/10.1097/00020840-200310000-00011

Cuda, D., Manrique, M., Ramos, Á., Marx, M., Bovo, R., Khnifes, R., Hilly, O., Belmin, J., Stripeikyte, G., Graham, P. L., James, C. J., Greenham, P. V., & Mosnier, I. (2024). Improving quality of life in the elderly: Hearing loss treatment with cochlear implants. *BMC Geriatrics*, 24(1), 16. https://doi.org/10.1186/s12877-023-04642-2

Daniel, E. (2007). Noise and Hearing Loss: A Review. *Journal of School Health*, 77(5), 225–231. https://doi.org/10.1111/j.1746-1561.2007.00197.x

Dawes, P., & Völter, C. (2023). Do hearing loss interventions prevent dementia? *Zeitschrift Fur Gerontologie Und Geriatrie*, 56(4), 261–268. https://doi.org/10.1007/s00391-023-02178-z

Dazert, S., Thomas, J. P., Loth, A., Zahnert, T., & Stöver, T. (2020). Cochlear Implantation. *Deutsches Arzteblatt International*, 117(41), 690–700. https://doi.org/10.3238/arztebl.2020.0690

eCIE-Maps—Documentación (sanidad.gob.es, January 13th 2024). Retrieved April 18, 2024, from https://www.eciemaps.sanidad.gob.es/documentation

Fernandez-Alvarez, M., Atienza, M., & Cantero, J. L. (2023). Cortical amyloid-beta burden is associated with changes in intracortical myelin

in cognitively normal older adults. *Translational Psychiatry*, *13*(1), 1–9. https://doi.org/10.1038/s41398-023-02420-7

Fitzhugh, M. C., & Pa, J. (2023). Exploring gray matter volume atrophy associations with hearing in the UK Biobank. *Alzheimer's & Dementia*, *19*(S23), e074985. https://doi.org/10.1002/alz.074985

Formeister, E. J., Krauter, R., Kirk, L., Zhu, T. R., Rizk, H. G., & Sharon, J. D. (2020). Understanding the Dizziness Handicap Inventory (DHI): A Cross Sectional Analysis of Symptom Factors That Contribute to DHI Variance. *Otology & Neurotology*, *41*(1), 86. https://doi.org/10.1097/MAO.0000000000002438

Frid, S., Duran, X. P., Cucó, G. B., Pedrera-Jiménez, M., Serrano-Balazote, P., Carrero, A. M., Lozano-Rubí, R., & others. (2023). An ontology-based Approach for consolidating patient data standardized with European Norm/International Organization for standardization 13606 (EN/ISO 13606) into Joint Observational Medical Outcomes Partnership (OMOP) repositories: Description of a methodology. *JMIR Medical Informatics*, *11*(1), e44547. https://doi.org/10.2196/44547

Gajecki, T., & Nogueira, W. (2022). An end-to-end deep learning speech coding and denoising strategy for cochlear implants. *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 3109–3113. https://doi.org/10.1101/2021.11.04.467324

Gauthier, S., Webster, C., Servaes, S., Morais, J. A., & Rosa-Nieto, P. (2022). *World Alzheimer Report 2022 – Life after diagnosis: Navigating treatment, care and support*. Alzheimer's Disease International.

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Iii, H. D., & Crawford, K. (2021). Datasheets for datasets. *Communications of the ACM*, *64*(12), 86–92. https://doi.org/10.48550/arXiv.1803.09010

Giallini, I., Inguscio, B. M. S., Nicastri, M., Portanova, G., Ciofalo, A., Pace, A., Greco, A., D'Alessandro, H. D., & Mancini, P. (2023a). Neuropsychological Functions and Audiological Findings in Elderly Cochlear Implant Users: The Role of Attention in Postoperative Performance. *Audiology Research*, *13*(2), 236–253. https://doi.org/10.3390/audiolres13020022

Giallini, I., Inguscio, B. M. S., Nicastri, M., Portanova, G., Ciofalo, A., Pace, A., Greco, A., D'Alessandro, H. D., & Mancini, P. (2023b). Neuropsychological Functions and Audiological Findings in Elderly Cochlear Implant Users: The Role of Attention in Postoperative Performance. *Audiology Research*, *13*(2), 236–253. Scopus. https://doi.org/10.3390/audiolres13020022

Giroud, N., Pichora-Fuller, M., Mick, P., Wittich, W., Al-Yawer, F., Rehan, S., Orange, J., & Phillips, N. (2021). Hearing loss is associated with gray matter differences in older adults at risk for and with Alzheimer's disease. *Aging Brain*, *1*, 100018.

Goodwin, M. V., Hogervorst, E., & Maidment, D. W. (2023a). Hearing difficulties and memory problems: The mediating role of physical health and psychosocial wellbeing. *International Journal of Audiology*, *0*(0), 1–8. https://doi.org/10.1080/14992027.2023.2199443

Goodwin, M. V., Hogervorst, E., & Maidment, D. W. (2023b). Physical activity interventions for adults with hearing loss: A systematic review. *Speech, Language and Hearing*, 1–11.

Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., Flandin, G., Ghosh, S. S., Glatard, T., Halchenko, Y. O., & others. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, *3*(1), 1–9.

Gurgel, R. K., Duff, K., Foster, N. L., Urano, K. A., & deTorres, A. (2022). Evaluating the Impact of Cochlear Implantation on Cognitive Function in Older Adults. *The Laryngoscope*, *132*(S7), S1–S15. https://doi.org/10.1002/lary.29933

Ha, J., Cho, Y. S., Kim, S. J., Cho, S. H., Kim, J. P., Jung, Y. H., Jang, H., Shin, H. Y., Lin, F. R., Na, D. L., Seo, S. W., Moon, I. J., & Kim, H. J. (2020). Hearing loss is associated with cortical thinning in cognitively normal older adults. *European Journal of Neurology*, *27*(6), 1003–1009. https://doi.org/10.1111/ene.14195

Haarbrandt, B., Schreiweis, B., Rey, S., Sax, U., Scheithauer, S., Rienhoff, O., Knaup-Gregori, P., Bavendiek, U., Dieterich, C., Brors, B., Kraus, I., Thoms, C. M., Jäger, D., Ellenrieder, V., Bergh, B., Yahyapour, R., Eils, R., Consortium, H. Gh., & Marschollek, M. (2018). HiGHmed—An Open Platform Approach to Enhance Care and Research across Institutional Boundaries. *Methods of Information in Medicine*, *57*(S 01), e66–e81. https://doi.org/10.3414/ME18-02-0002

He, Z., Li, M., Zou, S., Liao, F., Ding, Y., Su, H., Wei, X., Wei, C., Mu, Y., & Kong, W. J. (2019). Protection and Prevention of Age-Related Hearing Loss. In H. Li & R. Chai (Eds.), *Hearing Loss: Mechanisms, Prevention and Cure* (pp. 59–71). Springer. https://doi.org/10.1007/978-981-13-6123-4_4

Holman, J. A., Drummond, A., Hughes, S. E., & Naylor, G. (2019). Hearing impairment and daily-life fatigue: A qualitative study. *International Journal of Audiology*, *58*(7), Article 7. https://doi.org/10.1080/14992027.2019.1597284

Holzinger, A., Carrington, A., & Müller, H. (2020). Measuring the quality of explanations: The system causability scale (SCS) comparing human and machine explanations. *KI-Künstliche Intelligenz*, *34*(2), 193–198.

Huang, S., Zhao, G., Wu, J., Li, K., Wang, Q., Fu, Y., Zhang, H., Bi, Q., Li, X., Wang, W., Guo, C., Zhang, D., Wu, L., Li, X., Xu, H., Han, M., Wang, X., Lei, C., Qiu, X., ... Yuan, Y. (2021). Gene4HL: An Integrated Genetic Database for Hearing Loss. *Frontiers in Genetics*, *12*. https://doi.org/10.3389/fgene.2021.773009

Huber, M., Roesch, S., Pletzer, B., Lukaschyk, J., Lesinski-Schiedat, A., & Illg, A. (2020). Cognition in older adults with severe to profound sensorineural hearing loss compared to peers with normal hearing for age. *International Journal of Audiology*, *59*(4), 254–262. https://doi.org/10.1080/14992027.2019.1687947

HUVM website. (n.d.). *ITCBio platform*. Retrieved April 18, 2024, from https://www.hospitalmacarena.es/entrada-blog/itcbio-platform/

ISO 13606 Standard—EHR Interoperability (August 10th 2021). Retrieved April 18, 2024, from http://www.en13606.org/information.html

Jayakody, D., Friedland, P., Martins, R., & Sohrabi, H. (2018). Impact of Aging on the Auditory System and Related Cognitive Functions: A Narrative Review. *FRONTIERS IN NEUROSCIENCE*, *12*. https://doi.org/10.3389/fnins.2018.00125

Jwa, A. S., & Poldrack, R. A. (2022). The spectrum of data sharing policies in neuroimaging data repositories. *Human Brain Mapping*, *43*(8), 2707–2721.

Kahn, M. G., Callahan, T. J., Barnard, J., Bauck, A. E., Brown, J., Davidson, B. N., Estiri, H., Goerg, C., Holve, E., Johnson, S. G., Liaw, S. T., Hamilton-Lopez, M., Meeker, D., Ong, T. C., Ryan, P., Shang, N., Weiskopf, N. G., Weng, C., Zozus, M. N., & Schilling, L. (2016). A Harmonized Data Quality Assessment Terminology and Framework for the Secondary Use of Electronic Health Record Data. EGEMS (Washington, DC), 4(1), 1244. https://doi.org/10.13063/2327-9214.1244

Kramer, S., Vasil, K. J., Adunka, O. F., Pisoni, D. B., & Moberly, A. C. (2018). Cognitive Functions in Adult Cochlear Implant Users, Cochlear Implant Candidates, and Normal-Hearing Listeners. *Laryngoscope Investigative Otolaryngology*, *3*(4), 304–310. https://doi.org/10.1002/lio2.172

Lawrence, B. J., Jayakody, D. M. P., Bennett, R. J., Eikelboom, R. H., Gasson, N., & Friedland, P. L. (2020). Hearing Loss and Depression in Older Adults: A Systematic Review and Meta-analysis. *The Gerontologist*, *60*(3), e137–e154. https://doi.org/10.1093/geront/gnz009

Lazo-Maestre, M. *et al.* (2024), An Image-Guided Workflow for Individualized Surgical Planning and Multiphysics Simulation in Cochlear Implant Patients. In: Guisado-Lizar, JL., Riscos-Núñez, A., Morón-Fernández, MJ., Wainer, G. (eds) Simulation Tools and Techniques. SIMUtools 2023. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 519. Springer, Cham. https://doi.org/10.1007/978-3-031-57523-5_17

Levman, J., Ewenson, B., Apaloo, J., Berger, D., & Tyrrell, P. N. (2023). Error consistency for machine learning evaluation and validation with application to biomedical diagnostics. *Diagnostics*, *13*(7), 1315.

Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. *Boletín Oficial del Estado*, *294*, December 6th 2018. Retrieved April 18, 2024, from https://www.boe.es/buscar/doc.php?id=BOE-A-2018-16673

Lin, F. R., Thorpe, R., Gordon-Salant, S., & Ferrucci, L. (2011). Hearing Loss Prevalence and Risk Factors Among Older Adults in the United States. *The Journals of Gerontology: Series A*, *66A*(5), 582–590. https://doi.org/10.1093/gerona/glr002

Lin, F. R., Yaffe, K., Xia, J., Xue, Q. L., Harris, T. B., Purchase-Helzner, E., Satterfield, S., Ayonayon, H. N., Ferrucci, L., Simonsick, E. M., & others. (2013). Hearing loss and cognitive decline in older adults. *JAMA Internal Medicine*, *173*(4), 293–299. https://doi.org/10.1001/jamainternmed.2013.1868

Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2020). Explainable ai: A review of machine learning interpretability methods. *Entropy*, *23*(1), 18.

Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S., Brayne, C., Burns, A., Cohen-Mansfield, J., Cooper, C., & others. (2020). Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *The Lancet*, *396*(10248), 413–446. https://doi.org/10.1016/S0140-6736(20)30367-6

Livshits, E., Bertossi, L., Kimelfeld, B., & Sebag, M. (2021). The Shapley value of tuples in query answering. *Logical Methods in Computer Science*, *17*. https://doi.org/10.46298/lmcs-17(3:22)2021

Loughrey, D. G., Kelly, M. E., Kelley, G. A., Brennan, S., & Lawlor, B. A. (2018). Association of age-related hearing loss with cognitive function, cognitive impairment, and dementia: A systematic review and meta-analysis. *JAMA Otolaryngology–Head & Neck Surgery*, *144*(2), 115–126. https://doi.org/10.1001/jamaoto.2017.2513

Martinez-Amezcua, P., Powell, D., Kuo, P. L., Reed, N. S., Sullivan, K. J., Palta, P., Szklo, M., Sharrett, R., Schrack, J. A., Lin, F. R., & Deal, J. A. (2021). Association of Age-Related Hearing Impairment With Physical Functioning Among Community-Dwelling Older Adults in the US. *JAMA Network Open*, *4*(6), e2113742. https://doi.org/10.1001/jamanetworkopen.2021.13742

Moreno-Conde, J., Moreno-Conde, A., Salas-Fernández, S., & Parra-Calderón, C. L. (2019). ITCBio, a clinical and translational research platform. *AMIA Annual Symposium Proceedings*, *2019*, 673.

Moreno-Conde, J., Salas-Fernandez, S., & Moreno-Conde, A. (2022). MedicalForms: Integrated management of semantics for electronic health record systems and research platforms. *Applied Sciences*, *12*(9), 4322.

Mosnier, I., Vanier, A., Bonnard, D., Lina-Granade, G., Truy, E., Bordure, P., Godey, B., Marx, M., Lescanne, E., Venail, F., & others. (2018). Long-term cognitive prognosis of profoundly deaf older adults after hearing rehabilitation using cochlear implants. *Journal of the American Geriatrics Society*, *66*(8), 1553–1561. https://doi.org/10.1111/jgs.15445

Muñoyerro-Muñiz, D., Goicoechea-Salazar, J. A., García-León, F. J., Laguna-Téllez, A., Larrocha-Mata, D., & Cardero-Rivas, M. (2020). Health record linkage: Andalusian health population database. *Gaceta Sanitaria*, *34*(2), 105–113. https://doi.org/10.1016/j.gaceta.2019.03.003

Newman, C. (January 1st 2020). *HL7 Diagnostic Audiology Reporting Implementation Guide*. https://confluence.hl7.org/display/PHWG/Diagnostic+Audiology+Reporting+Implementation+Guide

Newman, C. W., Jacobson, G. P., & Spitzer, J. B. (1996). Development of the tinnitus handicap inventory. *Archives of Otolaryngology–Head & Neck Surgery*, *122*(2), Article 2.

Nieman, C. L., & Oh, E. S. (2020). Hearing Loss. *Annals of Internal Medicine*, *173*(11), ITC81–ITC96. https://doi.org/10.7326/AITC202012010

openEHR Concept name: Common_audiological_functional_parameters. (May 4th 2024). Retrieved June 3rd , 2024, from https://ckm.highmed.org/ckm/archetypes/1246.145.2160

Powell, D. S., Oh, E. S., Lin, F. R., & Deal, J. A. (2021). Hearing impairment and cognition in an aging world. *Journal of the Association for Research in Otolaryngology*, *22*(4), 387–403. https://doi.org/10.1007/s10162-021-00799-y

Protti, D. (2007). Moving toward a single comprehensive electronic health record for every citizen in Andalucía, Spain. *Healthcare Quarterly (Toronto, Ont.)*, *10*(4), 114–123, 4.

Publications Office of the European Union. (2016, April 27). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016*

on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [Website]. Publications Office of the EU; Publications Office of the European Union. https://op.europa.eu/en/publication-detail/-/publication/3e485e15-11bd-11e6-ba9a-01aa75ed71a1/language-en

Rader, T., Nachtigäller, P., Linke, T., Weißgerber, T., & Baumann, U. (2023). Exponential fitting of spread of excitation response measurements in cochlear implants. *Journal of Neuroscience Methods*, *391*, 109854.

Rajpurkar, P., Chen, E., Banerjee, O., & Topol, E. J. (2022). AI in health and medicine. *Nature Medicine*, *28*(1), 31–38. https://doi.org/10.1038/s41591-021-01614-0

Raymond, M. J., Ma, C., Schvartz-Leyzac, K. C., Camposeo, E. L., Nguyen, S. A., Meyer, T. A., & McRackan, T. R. (2023). Association of Cognitive Impairment Screening Scores With Improvements in Speech Recognition and Quality of Life After Cochlear Implantation. *JAMA Otolaryngology–Head & Neck Surgery*, *149*(4), 344–351. https://doi.org/10.1001/jamaoto.2022.4825

Relaño-Iborra, H., & Dau, T. (2022). Speech intelligibility prediction based on modulation frequency-selective processing. *Hearing Research*, *426*, 108610. https://doi.org/10.1016/j.heares.2022.108610

Rosemann, S., & Thiel, C. M. (2020). Neuroanatomical changes associated with age-related hearing loss and listening effort. *Brain Structure and Function*, *225*(9), 2689–2700. https://doi.org/10.1007/s00429-020-02148-w

Rutherford, B. R., Brewster, K., Golub, J. S., Kim, A. H., & Roose, S. P. (2018). Sensation and psychiatry: Linking age-related hearing loss to late-life depression and cognitive decline. *American Journal of Psychiatry*, *175*(3), 215–224.

Ryu, S. (2012). Book review: mHealth: New horizons for health through mobile technologies: Based on the findings of the second global survey on eHealth (global observatory for eHealth series, volume 3). *Healthcare Informatics Research*, *18*(3), 231. https://doi.org/10.4258/hir.2012.18.3.231

Sanchez-Cuadrado, I., Gavilan, J., Perez-Mora, R., Muñoz, E., & Lassaletta, L. (2015). Reliability and validity of the Nijmegen cochlear implant questionnaire in Spanish. *European Archives of Oto-Rhino-Laryngology*, *272*(7), Article 7.

Sanchez-Cuadrado, I., Lassaletta, L., Perez-Mora, R., Muñoz, E., & Gavilan, J. (2015). Reliability and validity of the Spanish Glasgow Benefit Inventory after cochlear implant surgery in adults. *European Archives of Oto-Rhino-Laryngology*, *272*(2), 333–336.

Sanchez-Lopez, R., Fereczkowski, M., Neher, T., Santurette, S., & Dau, T. (2020). Robust Data-Driven Auditory Profiling Towards Precision Audiology. *Trends in Hearing*, 24. https://doi.org/10.1177/2331216520973539

Sánchez-López, R., Grini-Nielsen, S., Cañete, O., Fereczkowski, M., Wu, M., Neher, T., Dau, T., & Santurette, S. (2019). *A clinical test battery for Better hEAring Rehabilitation (BEAR): Towards the prediction of individual auditory deficits and hearing-aid benefit*. 23rd International Congress on Acoustics, 3841–3848. https://doi.org/10.18154/RWTH-CONV-239177

Santos, M. R., Bax, M. P., & Kalra, D. (2010). Building a logical EHR architecture based on ISO 13606 standard and semantic web technologies. In *MEDINFO 2010* (pp. 161–165). IOS Press.

Schmidt, C. O., Struckmann, S., Enzenbach, C., Reineke, A., Stausberg, J., Damerow, S., Huebner, M., Schmidt, B., Sauerbrei, W., & Richter, A. (2021). Facilitating harmonized data quality assessments. A data quality framework for observational health research data collections with software implementations in R. *BMC Medical Research Methodology*, *21*, 1–15.

Siddique, S., Haque, M. A., George, R., Gupta, K. D., Gupta, D., & Faruk, M. J. H. (2023). Survey on Machine Learning Biases and Mitigation Techniques. *Digital*, *4*(1), 1–68.

Solarte, S. E., Chacón, M. M., & Ortiz, Y. A. (2016). Validez de contenido-escala "abbreviated profile of hearing aid benefit". *Areté*, *16*(1), 39–52.

Swanepoel, D. W. (2023). Advancing Equitable Hearing Care Through Innovations in Technology and Service-Delivery. *Folia Phoniatrica et Logopaedica*, 1–1. https://doi.org/10.1159/000530671

Theunisse, H. J., Mulder, J. J., Pennings, R. J. E., Kunst, H. P. M., & Mylanus, E. A. M. (2014). A database system for the registration of complications and failures in cochlear implant surgery applied to over 1000 implantations performed in Nijmegen, The Netherlands. *Journal of Laryngology and Otology*, *128*(11), 952–957. https://doi.org/10.1017/S0022215114002126

Uchida, Y., Sugiura, S., Nishita, Y., Saji, N., Sone, M., & Ueda, H. (2019a). Age-related hearing loss and cognitive decline—The potential mechanisms linking the two. *Auris, Nasus, Larynx*, *46*(1), 1–9. https://doi.org/10.1016/j.anl.2018.08.010

Uchida, Y., Sugiura, S., Nishita, Y., Saji, N., Sone, M., & Ueda, H. (2019b). Age-related hearing loss and cognitive decline—The potential mechanisms linking the two. *Auris, Nasus, Larynx*, *46*(1), 1–9. https://doi.org/10.1016/j.anl.2018.08.010

Van Dun, B., Kania, A., & Dillon, H. (2016). Cortical Auditory Evoked Potentials in (Un)aided Normal-Hearing and Hearing-Impaired Adults. *Seminars in Hearing*, *37*(1), 9–24. https://doi.org/10.1055/s-0035-1570333

Van Esch, T. E., Kollmeier, B., Vormann, M., Lyzenga, J., Houtgast, T., Hällgren, M., Larsby, B., Athalye, S. P., Lutman, M. E., & Dreschler, W. A. (2013). Evaluation of the preliminary auditory profile test battery in an international multi-centre study. *International Journal of Audiology*, *52*(5), 305–321. https://doi.org/10.3109/14992027.2012.759665

Van Opstal, A. J., & Noordanus, E. (2023). Towards personalized and optimized fitting of cochlear implants. *Frontiers in Neuroscience*, 17, 1183126. https://doi.org/10.3389/fnins.2023.1183126

Vicente, L., & Matute, H. (2023). Humans inherit artificial intelligence biases. *Scientific Reports*, *13*(1), 15737. https://doi.org/10.1038/s41598-023-42384-8

Viergever, K., Kraak, J. T., Bruinewoud, E. M., Ket, J. C., Kramer, S. E., & Merkus, P. (2021). Questionnaires in otology: A systematic mapping review. *Systematic Reviews*, *10*, 1–9. https://doi.org/10.1186/s13643-021-01659-9

Völter, C., Götze, L., Falkenstein, M., Dazert, S., & Thomas, J. P. (2017). Application of a computer-based neurocognitive assessment battery

in the elderly with and without hearing loss. *Clinical Interventions in Aging*, 12, 1681–1690. https://doi.org/10.2147/CIA.S142541

Völter, C., Götze, L., Kamin, S. T., Haubitz, I., Dazert, S., & Thomas, J. P. (2022a). Can cochlear implantation prevent cognitive decline in the long-term follow-up? *Frontiers in Neurology*, 13, 1009087. https://doi.org/10.3389/fneur.2022.1009087

Völter, C., Götze, L., Kamin, S. T., Haubitz, I., Dazert, S., & Thomas, J. P. (2022b). Can cochlear implantation prevent cognitive decline in the long-term follow-up? *Frontiers in Neurology*, 13, 1009087. https://doi.org/10.3389/fneur.2022.1009087

Wang, H. F., Zhang, W., Rolls, E. T., Li, Y., Wang, L., Ma, Y. H., Kang, J., Feng, J., Yu, J. T., & Cheng, W. (2022). Hearing impairment is associated with cognitive decline, brain atrophy and tau pathology. *Ebiomedicine*, 86.

Wang, T. C., Chang, T. Y., Tyler, R., Lin, Y. J., Liang, W. M., Shau, Y. W., Lin, W. Y., Chen, Y. W., Lin, C. D., & Tsai, M. H. (2020). Noise Induced Hearing Loss and Tinnitus—New Research Developments and Remaining Gaps in Disease Assessment, Treatment, and Prevention. *Brain Sciences*, 10(10), Article 10. https://doi.org/10.3390/brainsci10100732

Waring, J., Lindvall, C., & Umeton, R. (2020). Automated machine learning: Review of the state-of-the-art and opportunities for healthcare. *Artificial Intelligence in Medicine*, 104, 101822. https://doi.org/10.1016/j.artmed.2020.101822

Wasmann, J. W. A., Lanting, C. P., Huinck, W. J., Mylanus, E. A. M., Laak, J. W. M. V. D., Govaerts, P. J., Swanepoel, D. W., Moore, D. R., & Barbour, D. L. (2021). Computational Audiology: New Approaches to Advance Hearing Health Care in the Digital Age. *Ear and Hearing*, 42(6), 1499–1507. https://doi.org/10.1097/AUD.0000000000001041

Wayne, R. V., & Johnsrude, I. S. (2015). A review of causal mechanisms underlying the link between age-related hearing loss and cognitive decline. *Ageing Research Reviews*, 23, 154–166. https://doi.org/10.1016/j.arr.2015.06.002

Wells, T. S., Nickels, L. D., Rush, S. R., Musich, S. A., Wu, L., Bhattarai, G. R., & Yeh, C. S. (2020). Characteristics and Health Outcomes Associated With Hearing Loss and Hearing Aid Use Among Older Adults. *Journal of Aging and Health*, 32(7–8), 724–734. https://doi.org/10.1177/0898264319848866

WHO. (2021). *World Report on Hearing*. https://www.who.int/publications/i/item/9789240020481

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., & others. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1), 1–9. https://doi.org/10.1038/sdata.2016.18

Yung, M., James, A., Merkus, P., Philips, J., Black, B., Tono, T., Linder, T., Dornhoffer, J., & İncesulu, A. (2018). International otology outcome group and the international consensus on the categorization of tympanomastoid surgery. *The Journal of International Advanced Otology*, 14(2), 216.

Zhan, K. Y., Lewis, J. H., Vasil, K. J., Tamati, T. N., Harris, M. S., Pisoni, D. B., Kronenberger, W. G., Ray, C., & Moberly, A. C. (2020). Cognitive Functions in Adults Receiving Cochlear Implants: Predictors of Speech Recognition and Changes After Implantation. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 41(3), e322–e329. https://doi.org/10.1097/MAO.0000000000002544

Zhang, Y., Callejón-Leblic, M. A., Picazo-Reina, A. M., Blanco-Trejo, S., & Sánchez-Gómez, S. (2023). Impact of SNR, peripheral auditory sensitivity, and central cognitive profile on the psychometric relation between pupillary response and speech performance in CI users. *Frontiers in Neuroscience*, 17. https://doi.org/10.3389/fnins.2023.1307777

Appendix A

Additional materials can be found on the project's website.

- URL: https://plataforma.innovacionsalud.org (accessed on September 16th 2024).

From this website you will be able to access and test the MISIONES database by using the following user and password:

- User: demo_misiones
- Password: 123456

A user's manual of the demo platform can be found as well in this link.

Appendix B

Additional materials including the source codes implemented to read and import XML files from the Otoaccess software and Eclipse EP25 system (Interacoustics A/S, Middelfart, Denmark) can be found on the Github page:
URL: https://github.com/amparocallejon/MISIONES_IA

Data Availability Statement

Various supplementary materials are freely available as described in the different appendices and in the following links:
Supplementary Material 1
Supplementary Material 2
Supplementary Material 3

Complementary material

A video showing different sections of the platform is available at:
Demo MISIONES IA (accessed on September 16th 2024).