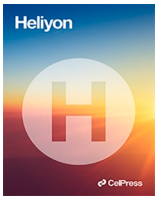
Heliyon 10 (2024) e32733

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

# Preliminary results of classifying otosclerosis and disarticulation using a convolutional neural network trained with simulated wideband acoustic immittance data

Michael Lauxmann [a,*], Felix Viehl [b], Barbara Priwitzer [c], Benjamin Sackmann [b]

[a] Doctor of Engineering, Faculty of Engineering, Reutlingen University, Alteburgstr. 150, 72762, Reutlingen, Germany
[b] Master of Science, Reutlingen Research Institute, Reutlingen University, Alteburgstr. 150, 72762, Reutlingen, Germany
[c] Doctor of Natural Sciences, Faculty of Engineering, Reutlingen University, Alteburgstr. 150, 72762, Reutlingen, Germany

ARTICLE INFO

*Keywords:*
Convolutional neural network
Finite-element model
Monte-Carlo simulation
Middle-ear pathologies
Hearing diagnostic
Wideband impedance
Acoustic stapedius reflex

ABSTRACT

Current noninvasive methods of clinical practice often do not identify the causes of conductive hearing loss due to pathologic changes in the middle ear with sufficient certainty. Wideband acoustic immittance (WAI) measurement is noninvasive, inexpensive and objective. It is very sensitive to pathologic changes in the middle ear and therefore promising for diagnosis. However, evaluation of the data is difficult because of large interindividual variations. Machine learning methods like Convolutional neural networks (CNN) which might be able to deal with this overlaying pattern require a large amount of labeled measurement data for training and validation. This is difficult to provide given the low prevalence of many middle-ear pathologies. Therefore, this study proposes an approach in which the WAI training data of the CNN are simulated with a finite-element ear model and the Monte-Carlo method. With this approach, virtual populations of normal, otosclerotic, and disarticulated ears were generated, consistent with the averaged data of measured populations and well representing the qualitative characteristics of individuals. The CNN trained with the virtual data achieved for otosclerosis an AUC of 91.1 %, a sensitivity of 85.7 %, and a specificity of 85.2 %. For disarticulation, an AUC of 99.5 %, sensitivity of 100 %, and specificity of 93.1 % was achieved. Furthermore, it was estimated that specificity could potentially be increased to about 99 % in both pathological cases if stapes reflex threshold measurements were used to confirm the diagnosis. Thus, the procedures' performance is comparable to classifiers from other studies trained with real measurement data, and therefore the procedure offers great potential for the diagnosis of rare pathologies or early-stages pathologies. The clinical potential of these preliminary results remains to be evaluated on more measurement data and additional pathologies.

## 1. Introduction

The human sense of hearing enables the recognition of sounds, speech, tones and noises and thus plays an essential role in interpersonal communication. If a person's ability to hear is compromised, this can have a restrictive effect on everyday life and make

* Corresponding author.
*E-mail addresses:* michael.lauxmann@reutlingen-university.de (M. Lauxmann), felix.viehl97@gmx.de (F. Viehl), barbara.priwitzer@reutlingen-university.de (B. Priwitzer), benjamin.sackmann@reutlingen-university.de (B. Sackmann).

https://doi.org/10.1016/j.heliyon.2024.e32733
Received 6 March 2023; Received in revised form 29 May 2024; Accepted 7 June 2024
Available online 8 June 2024

social interactions considerably more difficult [1]. Given the widespread occurrence of middle-ear disease and the difficulty of accurately identifying the cause of conductive hearing loss [2,3], a noninvasive and cost-effective diagnostic tool that reliably distinguishes among middle-ear pathologies would be valuable. This would aid in better selection of cases for surgery, surgical preparation, patient counseling, as well as help prevent unnecessary surgeries, for example, avoid exploratory middle-ear surgery. Common non-invasive diagnostic procedures include audiometry, otoscopy, and computerised tomography scans, although the latter is associated with a certain level of risk. However, these procedures either rely on the subjective judgment of the patient, are very time-consuming or costly or do not provide enough information to make an accurate diagnosis.

Wideband acoustic immittance (WAI) and its extension the wideband tympanometry (WBT), which additionally uses varying static air pressures in the ear canal to obtain more information, are very promising, objective diagnostic procedures enabling rapid and noninvasive differentiation of middle-ear diseases. Previous studies have shown that characteristics of various middle-ear diseases are visible in WAI measurements allowing identification of these pathologies [4–10].

However, a purely visual evaluation of the measurement data through comparison with standardized data of normal and pathological ears is difficult due to large inter-individual differences. Additionally, the effects of pathological changes in the middle ear on WAI measurements are not yet fully understood [11,12]. Therefore, WAI measurements are currently used only to a very limited extent in clinical practice.

A promising approach to better evaluate WAI measurement data is the use of machine learning. Nie et al. [13] were able to classify WBT measurement data of patients with normal and otosclerotic ears with an area under the curve (AUC) of the receiver operating characteristic (ROC) curve of 97 % by using data augmentation to enrich the measurement data and adapting an already pre-trained neural network via transfer learning. Sundgaard et al. [14] also used a deep learning approach for detecting otitis media based on WBT measurements. An AUC of 97 % was achieved, however it was not possible to distinguish between different types and stages of otitis media.

Even though classification based on WBT measurement data with high accuracy is possible using machine learning, certain problems remain. Despite the use of data augmentation in both papers, a large amount of WBT measurements is still required to sufficiently train a neural network. In Ref. [13] measurements performed on 80 patients suffering from otosclerosis and in Ref. [14] measurement samples of 526 ears with effusion and acute otitis media were used. Conducting clinical studies of this magnitude is costly and time-consuming. In addition, it is even more difficult to obtain measurement data from patients with rare middle ear diseases such as disarticulation or malleus head fixation. At an early stage, the diagnosis of a disease is often uncertain; therefore, the corresponding data can only rarely be included in clinical studies.

One way out of these problems might be virtually generated data for the AI training. By using a mathematical model of the auditory organ WAI data can be simulated and as many data sets as needed can be generated at reasonable time and cost even for rare conditions and early stages of diseases. These can subsequently be used to train a neural network which can classify measurement data collected from patients. Such a mathematical model was developed by Ref. [15]. It aims to simulate the data recorded during measurements and predict characteristic changes caused by middle ear pathologies. Local physical model parameters are adjusted to reproduce the anatomical causes of the respective pathologies. The predictive accuracy of this model was demonstrated using temporal bone measurements [12].

The aim of this work is to train a convolutional neural network (CNN) based on virtually generated WAI data that were simulated with the mathematical model from Ref. [12]. The CNN will be used to classify WAI measurements published in the literature and performed on patients with normal ears as well as patients suffering from otosclerosis and disarticulation. The classification is based on WAI data being cheaper and easier to obtain compared with diagnostic tools like umbo velocity measurements and high resolution computed tomography as described in Ref. [8]. Further, we assess the potential of confirming CNN-based WAI diagnosis with measurements of acoustic stapedius reflex threshold. The CNN-based procedure is supposed to be fully automatic, without the need to extract specific features manually. To our knowledge, there is no published literature using simulated data for the training of a CNN to classify measured WAI samples in normal, otosclerotic, and disarticulated ears. But there is previous work generating synthetic data for usage in applications like middle-ear diagnosis [16–19].

## 2. Materials and methods

To bypass the challenge of limited measurement data on pathologies simulated data were used for the training of the neural network. WAI measurement data published in literature and accessed via the WAI database [20] were used as classification test data to assess the performance of the CNN on real data. The neural network in this study shall be provided with the maximum available information from WAI measurements. Therefore, in this study, impedance magnitude, impedance phase, energy reflectance, pressure reflectance phase, assumed ear canal area at the probe position, and frequency of the first ear canal antiresonance in the complex impedance as a measure for the distance from the probe to the eardrum were considered. In the following the WAI measurement data and simulated training data are described in detail and an overview of the CNN architecture is presented.

### 2.1. Definition of WAI measures

The classification is based on the quantities of ear canal impedance (ZEC), pressure reflectance (R) and energy reflectance (ER), which fall under the rubric of wideband immittance measures [21]. To measure these quantities, a dynamic sound pressure is generated using a probe inserted into the ear canal. The sound pressure propagates through the ear canal and causes the tympanic membrane to vibrate. Part of the sound pressure is absorbed, and the rest is reflected. The reflected sound pressure and the sound

pressure emitted by the probe interfere. The resulting pressure can be measured by a microphone fitted in the probe. From the measured sound pressure, the impedance $ZEC_M(x)$ at a distance $x$ from the tympanic membrane can be determined using the Thevenin-equivalent of the probe derived from a calibration procedure. The pressure reflectance $R(x)$ at the tympanic membrane can be quantified by equation (1) in terms of the acoustic ear canal impedance $ZEC_M$ and the characteristic impedance of the ear canal $Z_C = \rho\, c\, /\, A_M(x)$, where $Z_C$ depends on the density of air $\rho$, the speed of sound in air $c$, and the cross-section area of the measurement position $A_M(x)$ such that:

$$R(x) = \frac{ZEC_M(x) - Z_C(x)}{ZEC_M(x) + Z_C(x)}. \tag{1}$$

The accuracy of the measured pressure reflectance $R(x)$ depends on the underlying approximations that (i) the viscosity of air is negligible, and (ii) that the ear canal is a rigid-walled tube of constant cross-section. With these assumptions, which do not hold true above 5 kHz due to increasing influence of cross-section variations [22], the magnitude of the pressure reflectance along the tube is approximately constant and equals the magnitude of the pressure reflectance at the termination (medial end of the ear canal) with $x = 0$. The angle of the pressure reflectance measured along the tube equals the angle of the reflectance at the termination plus a factor related to the round-trip travel time of the pressure wave between the measurement point and the termination at the tympanic membrane. The impedance $ZEC_T$ at the target position $x = L_T$ can be calculated from the impedance $ZEC_M$ at the measurement position $x = L_M$ by the following equation [11]

$$ZEC_T = Z_C \cdot \frac{ZEC_M + i\, Z_C \tan(k(L_T - L_M))}{Z_C + i\, ZEC_M \tan(k(L_T - L_M))} \quad with \quad Z_C = \frac{\rho\, c}{A_M}. \tag{2}$$

The wavenumber k is calculated from the quotient of the angular frequency and the speed of sound: $k = \omega\, /\, c$. From the pressure reflectance the so-called power/energy reflectance ER can be derived as the ratio of the power or energy in the reflected wave to the power or energy in the incident wave. It is defined by:

$$ER = |R(x)|^2 = \left|\frac{ZEC_T - Z_C}{ZEC_T + Z_C}\right|^2. \tag{3}$$

Note that the energy reflectance ER is a real number and contains no phase information, whereas the pressure reflectance R and the impedance values $ZEC_T$ and $ZEC_M$ are complex numbers. Energy absorbance is simply $1 -$ ER, and therefore, we use these terms interchangeably when referring to published literature.

### 2.2. Selection and preparation of WAI measurement data

The measurement data used as test data for the classification come from a freely accessible WAI database described in Ref. [20], which is funded by the NIDCD (National Institute on Deafness and other Communication Disorders) and hosted on a server at Smith College (USA). The database contains the individual WAI measurement data from peer-reviewed papers. Most datasets are from normal hearing infants, children, and adults, but there are also some samples from subjects with pathologies such as otosclerosis, disarticulation, and semicircular canal dehiscence.

The database was filtered for measurements from normal, otosclerosis, and disarticulation adult subjects. The measurement data had to include all information needed for the calculation of complex R and ZEC at ambient pressure in the frequency range 226–6000 Hz. Individual samples from all the studies with appropriate data were further checked for plausibility and preprocessed for usage as test data for the neural network. First, ER was calculated from the complex impedance given in the database and compared to the given ER. Secondly, the first ear canal antiresonance was identified where possible. Depending on the age of the probands appropriate ear canal areas were chosen and finally with this ear canal area the R-phase was calculated, and ER was recalculated with age dependent ear canal areas.

ER can be calculated directly from ZEC and the characteristic impedance, as shown in equation (3). As a first estimate, the characteristic impedance was determined with assumption of an average ear canal area of 44 $mm^2$ as chosen for adults by the HearID system [23], density of air of 1.21 $kg/m^3$ and a sound velocity of 343 m/s. ER was calculated from the given ZEC magnitude and phase using equation (3) and compared to the given ER in the WAI database. If the maximum absolute difference or the frequency averaged mean absolute difference between the reflectance from the database and the recalculated reflectance in the frequency range 226–1600 Hz was larger than 0.2 or 0.15 respectively, the sample was discarded. Such deviations could be attributed to the measurement protocol and in particular to the probe calibration. The effects of probe impedance on reflectance measurements are discussed e.g. in Ref. [24].

Since probe insertion depth affects impedance, a further criterion in the classification to improve comparability between simulation and measurement data is the frequency value of the first acoustic ear canal antiresonance. Providing the frequency of antiresonance instead of an estimated ear canal length avoids errors in its estimation, which is also slightly affected by the impedance of the eardrum. The detection algorithm for the first antiresonance checked the following criteria: 1) The ZEC phase must increase in two consecutive frequencies and change its sign or pass zero degrees. 2) The ZEC amplitude at these frequencies must have a local minimum. For a unique determination of the first antiresonance, and not to detect by chance higher antiresonances, ZEC phase was restricted not to surpass $+45°$ in the frequency range from 226 Hz to the frequency candidate detected by the algorithm. If the algorithm was not able to detect this antiresonance or if the frequency of the antiresonance was below 2144 Hz, the related sample was excluded. Frequencies

below 2144 Hz correspond to an acoustic ear canal length of more than 40 mm if a cylindrical canal and rigid termination is assumed. This ear canal length is already a very conservative assumption given the longest measured ear canal length by Ref. [25] of 35 mm and the assumption from Ref. [23] that the measurement probe could sit about 12 mm from the ear canal entrance.

Ear canal cross-section also affects impedance. Some studies provided acoustic estimates of the ear canal area in the WAI database, but [26] reported that constant area estimates are often closer to the real area and more robust than acoustic estimates of the area. Voss et al. [23] showed that the area depends significantly on the age cohort. Therefore, the area of the samples was estimated from the median age of the age groups given in Ref. [23]. All data remaining after the previous checks contained either sample age or related descriptions in the paper that allowed to categorize all samples in the decades of age given in Ref. [23], which are 18–29, 30–39, 40–49, 50–59, 60–69, and 70–79 years. R phase was not given in the database and had to be calculated from the complex ZEC. The area at the 12 mm insertion depth measured in Ref. [23] was taken as an assumption for the subjects areas to calculate the characteristic impedance. ER was recalculated with this area assumption for all samples except the disarticulation samples with implausible phase, as discussed later. The recalculated ER was used for classification. Usually, ER is said to be mostly insensitive to ear-canal length and ear canal area [21,26,27]. However, this strictly only holds true if for the calculation of ER from measured impedance a correct ear canal cross-section is chosen.

### 2.3. Generation of simulated WAI training data

The virtual WAI training data for the CNN was gathered through Monte Carlo simulations using the finite element (FE) model published in Ref. [12], which is based on [15,28,29]. The FE model is executed once for each parameter sample. The FE model was preprocessed in HYPERMESH (Altair Engineering, Inc., Troy, MI) and solved with the structural solver OPTISTRUCT (Altair Engineering, Inc., Troy, MI). The solver-input-files were parametrized in MATLAB. In the FE model the ear canal, tympanic cavity, and tympanic membrane geometry are based on microcomputed tomography data of post-mortem temporal bones. The air in the ear canal and tympanic cavity is modeled with hexahedral (HEXA8) elements with a compressible inviscid fluid material (MAT10). The ear canal and tympanic cavity walls are assumed to be rigid and lossless. The tympanic membrane is meshed with second-order shell elements (QUAD8) with isotropic material (MAT1). For the thickness distribution we refer to Ref. [12]. The ossicles are modeled as rigid bodies, characterized by their mass and moments of inertia (CONM2). All ligaments, tendons, and joints are represented by massless generalized spring-damper structural elements (CBUSH). The inner ear load is represented by a viscose damper element. Further details about the model are published in Ref. [12]. In custom MATLAB scripts, stapes transfer function (STF), umbo transfer function (UTF), and the WAI quantities ZEC, R, and ER are calculated as described in Ref. [12] from the FE outputs. The model is validated on data from a temporal bone study [30] and shows good agreement with ER, ZEC, UTF, and STF for normal ears and ears with otosclerosis, malleus fixation, and disarticulation. For more details about the validation procedure, we refer to Ref. [12].

The process for creating virtual training data through Monte Carlo simulations involved drawing a large number of samples and simulating them with the model. The resulting data was then filtered through literature-based measurement limits, without the use of iterative sampling algorithms. The detailed procedure was as follows.

1) Uniform parameter distributions were defined on log10-transformed parameter ranges, which were reported in Ref. [12] as initial parameters for the identification case of the normal ear.
2) A Halton pseudorandom sequence of 400,000 samples was drawn once before performing the simulations from the uniform distributions using UQLAB-Toolbox [31].
3) With the FE model values for STF, UTF, ZEC, and ER were simulated for all drawn 400,000 samples according to the definitions in Ref. [12]. All responses were calculated at 14 logarithmically distributed frequencies between 200 and 6000 Hz. Otosclerosis was simulated by increasing the annular ligament stiffness parameters of all simulated individuals by a factor of 10,000. Disarticulation was simulated by setting stiffness and damping parameters of the incudostapedial joint for all simulated individuals to zero. The pathologies were simulated by repeating the model evaluation and by only changing these specific pathology-related parameters, compared to the samples' values of the initial normal sample, drawn from the uniform distributions.
4) To constrain the simulation to plausible data, limits were defined based on averaged literature data for the simulated STF, UTF, ER, and relative changes to the normal condition due to the pathology. Only simulations of individuals that were within all defined limits both for normal and all pathological cases were retained and used as training and validation data of the CNN. All limits are described below and are shown in Figs. 4 and 5.

The following references were considered for the normal ear limits. In Ref. [32] the grand mean and twice the standard deviation (SD) of 13 study means of STF was published. Since recent research showed that the variation within the normal ears' population is underestimated [33], samples exceeding ± 2 SD were also allowed within a 6 dB tolerance band. Rosowski et al. [34] published the mean ± SD of the umbo transfer function measured on 58 normal ears (29 subjects) and averaged ER data from the same population as well as the two most extreme samples 23R and 6R with the largest and smallest ER at frequencies below f = 2000 Hz out of the 58 normal ears, see Fig. 4. Accounting for model and measurement uncertainties and allowing violation of the limits which are based on averaged references at single frequencies the following limits were defined for the normal ear.

(1) $STF_{simulated} \leq STF_{\text{literature [32],mean+2SD}} + 6\ dB,\ \text{for } 200\ Hz \leq f \leq 6000\ Hz,$
(2) $STF_{simulated} \geq STF_{\text{literature [32],mean−2SD}} - 6\ dB,\ \text{for } 200\ Hz \leq f \leq 6000\ Hz,$
(3) $UTF_{simulated} \leq UTF_{\text{literature [34],mean+2SD}} + 6\ dB,\ \text{for } 200\ Hz \leq f \leq 6000\ Hz,$

(4) $\mathrm{UTF}_{\mathrm{simulated}} \geq UTF_{\mathrm{literature\ [34],mean-2SD}} - 6\ dB,$ for $200\ Hz \leq f \leq 6000\ Hz$,
(5) $\mathrm{ER}_{\mathrm{simulated}} \leq \max\left(ER_{\mathrm{literature\ [34],sample\ 23R}} + 0.1, 1\right),$ for $200\ Hz \leq f \leq 1600\ Hz$,
(6) $\mathrm{ER}_{\mathrm{simulated}} \geq \min\left(ER_{\mathrm{literature\ [34],sample\ 6R}} - 0.1, 0\right),$ for $200\ Hz \leq f \leq 1600\ Hz$.

In [30] the mean ± SD of the change in UTF, ZEC, and ER due to pathologies like disarticulation and otosclerosis was studied in 14 temporal bones by performing measurements on each specimen in an initial normal condition and subsequent manipulated conditions. This change in measurements of UTF, ZEC, and ER is referred to as ΔUTF, ΔZEC, and ΔER, respectively, and corresponds in simulated individuals to the change in magnitude of UTF, ZEC, and ER due to adjustment of pathology-relevant model parameters to abnormal value ranges. Due to the small study size, the SD in Ref. [30] might be overestimated. Therefore, limits in this study are based mostly on the single ± SD range. Exceptions were defined for disarticulation to account for model uncertainties from not fully understood effects like pretension loss or change of tympanic membrane shape. Such effects might result as side effects from separating the incudostapedial joint as discussed in Ref. [12]. Furthermore, focus was placed on the minimum relative changes from normal to pathology in the frequency range below 1600 Hz, which has shown to be most sensitive to pathological changes. A maximum relative change from the normal case to a pathologic case in UTF (ΔUTF) was only limited in case of otosclerosis to prevent simulating extreme outliers, which might result from the simplified assumption of a very high stiffening factor of the annular ligament. Since UTF is still very sensitive, even for stiffening factors above 1000, but ZEC and ER become increasingly insensitive, this filter barely affects the distribution of the population, which was used for the CNN training and validation. The following limits were defined for otosclerosis:

(7) $\Delta\mathrm{ZEC}_{\mathrm{simulated}} \geq \Delta\mathrm{ZEC}_{\mathrm{literature\ [30],mean-SD}}$, for $200\ Hz \leq f \leq 1600\ Hz$,
(8) $\Delta\mathrm{ER}_{\mathrm{simulated}} \geq \Delta\mathrm{ER}_{\mathrm{literature\ [30],mean-SD}}$, for $200\ Hz \leq f \leq 1600\ Hz$,
(9) $\Delta\mathrm{UTF}_{\mathrm{simulated}} \leq \Delta\mathrm{UTF}_{\mathrm{literature\ [30],mean+SD}}$, for $200\ Hz \leq f \leq 1600\ Hz$,
(10) $\Delta\mathrm{UTF}_{\mathrm{simulated}} \geq \Delta\mathrm{UTF}_{\mathrm{literature\ [30],mean-SD}} - 10\ dB$, for $200\ Hz \leq f \leq 6000\ Hz$,

and for disarticulation.

(11) $\Delta\mathrm{ZEC}_{\mathrm{simulated}} \leq \Delta\mathrm{ZEC}_{\mathrm{literature\ [30],mean+SD}} + 2\ dB$, for $200\ Hz \leq f \leq 600\ Hz$,
(12) $\Delta\mathrm{ER}_{\mathrm{simulated}} \leq \Delta\mathrm{ER}_{\mathrm{literature\ [30],mean+SD}} + 0.15$, for $200\ Hz \leq f \leq 1000\ Hz$,
(13) $\Delta\mathrm{UTF}_{\mathrm{simulated}} \geq \Delta\mathrm{UTF}_{\mathrm{literature\ [30],mean-SD}} - 2\ dB$, for $200\ Hz \leq f \leq 1000\ Hz$.

The measured ear canal impedances include the variation of ear canal length and cross-section, with variations mostly coming from the acoustic length of the ear canal. The acoustic ear canal length corresponds to the anatomical length of the ear canal minus the probe insertion depth during the measurement and can thus also vary between different measured sessions at the same patient. Furthermore, interindividual variations of anatomical ear canal length are known [25]. Since the FE simulation uses a nominal ear canal length of 9 mm, the simulated population was augmented to further include this variation. As an estimate of the variation of ear canal lengths of the patient population, the automatically detected frequencies of the first acoustic antiresonance from the test population in this study was used. Assuming rigid termination and plane wave sound transmission led to a mean ear canal length of 25.3 mm with a standard deviation of 5.6 mm. This is a plausible assumption when comparing to Ref. [35], who estimated the acoustic lengths, i.e. length from probe to eardrum, based on the reflectance group delay. The average acoustic length of their smaller population of 54 samples from 27 adult subjects with an average age of 44.1 years is 25.3 mm with a standard deviation of 6.2 mm. Therefore, for each of the 79 simulated individuals, 40 samples from a Gaussian distribution $\mathcal{N}(25.3\mathrm{mm}, 5.6\mathrm{mm})$ are drawn for the ear-canal length and a cross-section is chosen randomly from the age groups' median values reported in Ref. [23] at the 12 mm insertion depth. Simulations at the plausible parameter samples gathered from the Monte-Carlo method are redone at 126 logarithmic distributed frequencies from 200 to 15000 Hz. Using equation (2) and the nominal ear canal length and cross-section, firstly the impedance at the eardrum is calculated for all simulated samples. From this position using equation (2), and the sampled ear-canal lengths and cross-sections, the propagated ear-canal impedance for the training population is calculated. This procedure finally yielded 3160 virtual samples for each of the groups normal, otosclerosis, and disarticulation, resulting in a total of 9480 virtual samples for training and validation of the CNN.

The higher resolution and wider frequency range was required for an automatic detection of the frequencies of the first acoustic antiresonance which is done equivalent to the procedure applied to the measurement data. This information is provided to the neural network besides ZEC, R, and ER data to account for ear canal variations. Providing the frequency of antiresonance instead of an estimated ear canal length avoids errors in its estimation, which is also slightly affected by the termination impedance of the eardrum.

### 2.4. Network architecture

In the context of this work, we developed a so-called lightweight CNN, which is supposed to avoid overfitting [13]. It is designed based on the classical LeNet architecture [36], which has been shown to be effective for applications with small data sets [37]. However, regarding the application described in this paper, some necessary modifications have been made. Like common in LeNet architecture, the developed neural network consists of two convolution layers, each followed by a nonlinear activation layer. Classical Rectified Linear Unit (ReLU) layers were used as nonlinear activation layers. A batch normalization layer was inserted between each of the convolution layers and nonlinear activation layers. The use of this type of layer is common in CNNs and it is intended to speed up

the training process. In a LeNet architecture, the ReLU layer is typically followed by a pooling layer. However, this type of layer was not implemented in this CNN due to the small size of the input argument. Since the use of a pooling layer would reduce this pixel number even further, it might have a negative impact on the accuracy of the neural network in this application.

The CNN is implemented in Matlab R2020a using the trainNetwork and classify functions for image classification. Therefore, the simulated training data and measured test data is stored in an ImageDatastore object, which contains the following information of each sample stored as image data: ZEC magnitude and phase with 96 frequency steps from 226 to 6000 Hz, ER magnitude and R phase with 76 frequency steps from 226 to 3000 Hz, and the scalar quantities age-related ear-canal area and the determined frequency of the first ear canal antiresonance. These data are arranged for each sample to a 2D-Image. Since the data vectors of the different measured variables have different lengths, zero-padding is performed. Data were arranged in a way that frequency axes of different data sequences are aligned. Furthermore, 1D-Information like cross-section and length was extended to a complete data column preceding the other data by repeating the scalar values. Since there are four frequency-dependent quantities, with a maximum vector size of 1x96, resulting in a stacked matrix size of 4x96, and two scalar quantities, each extended to a size of 4x1, the total image size generated from the data points is 4x98 pixels. The numerical values of the data are saved as indexed image of data type 8-bit unsigned integer using imwrite to folders named by their label. Training and test data are provided to the network through Matlab ImageDatastore objects.

The first layer of the CNN is an image input layer with an input size of 4 x 98 x 1. It is followed by the first convolution layer with a rectangular kernel of 4 x 30, because the length and width of the image were not equal. The convolution layer was employed with 45 filters. It is followed by a batch normalization layer and a ReLU layer. The second convolution layer was defined with a kernel size of 1 x 10 and 45 filters. The stride was defined for both convolution layers as (1,1), which means that the filters are applied to the respective input matrices with a step size of 1 in both vertical and horizontal directions. No padding was used.

After the features of the datasets are extracted and processed by the dual sequence of convolution layer, batch normalization layer, and ReLU layer, they are passed to the fully connected layer. Its neurons are connected to all neurons of the previous layer, combining local features learned about the image from the previous layers and recognizing higher-level patterns. The fully connected layer has an input size of 2700 and an output size of three, which is equal to the number of the three classes in the dataset. It is followed by a SoftMax layer and a classification layer. The SoftMax layer is used in classifications to assign a decimal numerical value to each class, which can be interpreted as the probability that a sample belongs to the respective class and is used by the classification layer to finally classify the datasets into the different classes. The entire CNN architecture is shown in Fig. 1 and an overview of the individual layers is given in Table 1, showing also the Matlab function names of all layers.

The network was designed and trained provided with all mentioned layer information and the following training options. As solver the stochastic gradient descent with momentum (SGDM) optimizer in MATLAB R2020a was used. The neural network was trained from scratch with 20 epochs using an NVIDIA Quadro P2200 GPU. A validation was performed every 15 iterations. The loss function was the cross-entropy loss, which is defined by equation (4),

$$loss = -\frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{K} w_i\, t_{ni}\, ln(y_{ni}) \tag{4}$$

where N corresponds to the total amount of samples, K is the number of classes, $w_i$ is the weight for class $i$ and $t_{ni}$ is the indicator that the $n^{th}$ sample belongs to class i. The probability $y_{ni}$ associates the $n^{th}$ input with class i, which in this case is the value determined by the SoftMax function [38].

Before training the neural network, the virtually generated dataset was randomly divided into training and validation data. 80 % of the samples were used for training and 20 % for validation. It was ensured that each simulated sample and its propagated derivatives concerning variations in ear canal length and area were only used for either training or validation. To ensure repeatability of the classification and to be able to draw conclusions based on the classification accuracy, a fixed seed was used to divide the dataset in a random but repeatable manner. The performance of the classifications was evaluated based on overall accuracy, precision, recall,

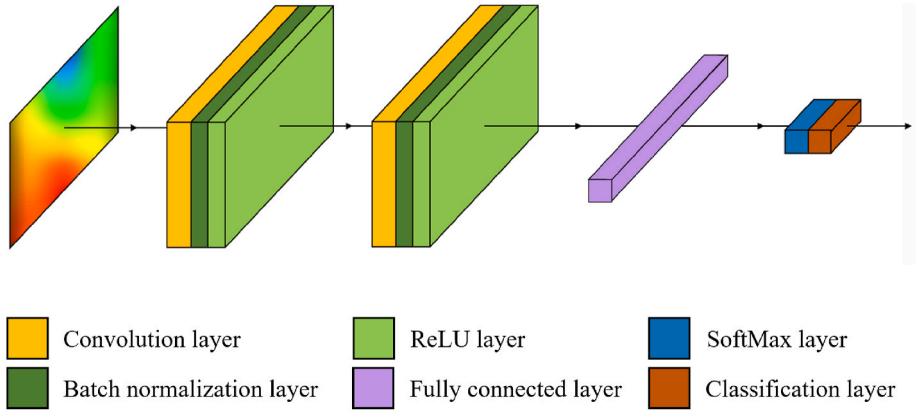



**Fig. 1.** Overview of the CNN architecture.

**Table 1**
Layer arrangement and specifications of the CNN.

| Layer | Matlab function | Channel | Kernel Size | Activations |
|---|---|---|---|---|
| Input layer | imageInputLayer | – | – | 4 x 98 x 1 |
| Convolution layer | convolution2dLayer | 45 | 4 x 30 | 1 x 69 x 45 |
| Batch normalization layer | batchNormalizationLayer | 45 | – | 1 x 69 x 45 |
| ReLU layer | reluLayer | – | – | 1 x 69 x 45 |
| Convolution layer | convolution2dLayer | 45 | 1 x 10 | 1 x 60 x 45 |
| Batch normalization layer | batchNormalizationLayer | 45 | – | 1 x 60 x 45 |
| ReLU layer | reluLayer | – | – | 1 x 60 x 45 |
| fully connected layer | fullyConnectedLayer | – | – | 1 x 1 x 3 |
| SoftMax layer | softmaxLayer | – | – | 1 x 1 x 3 |
| Classification layer | classificationLayer | – | – | 1 x 1 x 3 |

specificity, sensitivity, AUC, and F1-Score [39].

### 2.5. Stapes reflex as a second diagnostic criterion

In addition to the classification of WAI data with the CNN, it was also investigated whether the classification accuracy between normal and pathological ears could be further improved by taking into account the positive or negative result of an acoustic stapes reflex threshold (ASRT) measurement with the help of a decision tree. Since the measurement data from the WAI database do not contain any information on ASRT, the ASRT prevalences are assumed for an initial benefit assessment of the decision tree on the basis of the literature data described below.

Feeney et al. [40] obtained normative data on ASRT for normal-hearing adults. The study included 33 subjects with normal audiometry and clinical immittance. In their study they compared wideband and clinical ASRTs. For the clinical ASRT, they tested for ipsilateral ASRT with a 226 Hz probe tone and broadband noise (BBN) activator in both ears and repeated measurements after 1 month. With the clinical system, 15 of 102 (14.7 %) measurement sessions had absent reflex. Performing subsequently experimental wideband ASRT would result in 7 of 102 (6.9 %) ears without measurable reflex.

Hong et al. [41] investigated the relationship between acoustic reflex and conductive disorders with a retrospective review of 212 ears showing evidence for conductive hearing loss. The causes of hearing loss were determined from interoperative findings or computed tomography imaging. Both ipsilateral and contralateral ASRT were measured at tonal activators of 0.5, 1,2 and 4 kHz. In the group of ears with only an ossicular disorder 166 ears had all 8 reflexes completely absent. A number of 1–8 positive reflexes was found in 10 ears with pure ossicular disorder. 94 % of the ears with ossicular disorder had otosclerosis. Kan et al. [42] investigated sensitivity and specificity of a range of diagnostic tests including stapes reflex on patients with surgically confirmed otosclerosis and disarticulation. The study groups consisted of 107 ears of 98 patients with otosclerosis and 20 ears of 19 patients with disarticulated incudostapedial joint. They found no positive acoustic reflexes at ipsilateral stimulus at the study test frequencies 0.5 and 2 kHz in any of the measured ears. They found absent contralateral reflex at 1 kHz in 83 otosclerotic cases and only 1 otosclerotic case with positive reflex. In the disarticulation group 13 ears had absent contralateral reflex and 6 ears positive reflex. Keefe et al. [43] investigated the diagnostic value of a battery of ear-canal tests including acoustic reflex in quantifying risk of otosclerosis. The study sample consisted of 23 normal, 12 otosclerotic ears, and 13 ears tested after surgical intervention. In their study otosclerosis ears were unambiguously classified with clinical and research ASRT as having an absent or highly elevated threshold compared to the normal group ears. Clinical ASRT was performed ipsilateral and contralateral with tonal activators 0.5, 1, and 2 kHz and ipsilateral with a BBN activator.

Based on the findings from Ref. [40], in this study 85.3 % was chosen as a conservative assumption of the positive acoustic reflex prevalence for normal ears. As an optimistic positive prevalence assumption, a value of 93.1 % was chosen. Based on the findings from Ref. [41] a conservative prevalence of positive reflex in otosclerosis and disarticulation cases was chosen as 6 %. This study, however, did not allow to determine a prevalence for ipsilateral reflex alone. The findings from Ref. [42] suggest that ipsilateral reflex was absent in all otosclerosis and disarticulation cases. Keefe et al. [43] support this finding since they found absent or highly elevated thresholds for otosclerotic ears, which led to unambiguous classification of the otosclerotic ears. Therefore, the optimistic choice in this study for ears with otosclerosis or disarticulation is 100 % absent reflex in these cases. The prevalences for the conservative and optimistic scenario are summarized in Table 2.

Fig. 2 shows the decision tree with the two-stage procedure consisting of CNN classification and subsequent review using ASRT. The

**Table 2**
Probability for positive ASRT based on the prevalences given in Refs. [40–43] for a conservative and optimistic scenario.

| true label | scenario | |
|---|---|---|
| | conservative | optimistic |
| normal | 85.3 % | 93.1 % |
| otosclerosis | 6 % | 0 % |
| disarticulation | 6 % | 0 % |

decision rule of the second stage is that a suspicion of the CNN of otosclerosis or disarticulation should be confirmed by a negative ASRT result, a normal finding of the CNN by a positive ASRT. In all other cases, the CNN suggestion is questioned and the test sample is sorted into the category 'unknown'. Due to the lack of information on the ASRT in the test data, literature values about the probabilities of a positive ASRT are used, which are summarized in Table 2 assuming a conservative or optimistic scenario. For example, otosclerotic samples show a negative ASRT with a 94 % probability in the conservative scenario based on the information in the literature. Thus, if for example a true otosclerotic sample is recognized by the CNN as such, it will show a negative ASRT with a 94 % probability and thus be confirmed as a case of otosclerosis. In 6 % of cases, this real otosclerotic sample shows positive ASRT, which means that the CNN suggestion is interpreted as uncertain and the sample is categorized as 'unknown'. The 'unknown' class is also included in the confusion matrix so that the performance value for the diagnosis by means of the decision tree can be calculated correctly. To estimate the performance of the decision tree logic, the test data classified by the CNN is further divided proportionally into the four classes according to the ASRT probabilities and finally the confusion matrix is rounded to whole samples.

## 3. Results

### 3.1. Summary of measurement data

Appropriate data for normal patient ears were found within the WAI database from the investigators [4,22,34,44–50]. Table 5 in the Appendix gives an overview of these studies with basic subject information, inclusion criteria and the used measurement devices. In total 460 normal samples from 261 subjects aged between 18 and 64 years were available for further analysis from the considered studies. Some adult normal subject studies met some but not all criteria and were therefore excluded. For example [6,43], provided only the amplitude of reflectance, i.e. ER. Refs. [51,52] did not completely meet the frequency range criterion. Feeney et al. [40] provided tympanic membrane impedance instead of ear canal impedance. To be comparable to other data, however, this study requires ear canal impedance. Feeney et al. [40] did not provide the estimated ear canal lengths in the database or paper, which would have allowed to reconstruct original measured ear canal impedance.

Nakajima et al. [8] was the only study in the database with suitable data for ears with otosclerosis and disarticulation. The publication contains measurements of 14 ears affected by otosclerosis and 6 ears affected by disarticulation, two of which had partial disarticulation and four of which had complete disarticulation. Patients aged between 22 and 72 years. In Ref. [8], two outliers are described as having atypical tympanometry measurement data for their pathology. The otosclerosis sample (P049L) did not show the expected values of a stiffer ossicular chain in impedance and reflectance. The second outlier (P054R) was a subject with disarticulation, in which the reflectance did not show the characteristic pattern of disarticulation. These two outliers were therefore not used in this work. Further study information is provided in Table 5 in the Appendix.

The consistency check between EA from the database and EA calculated from the database led to rejection of in total 5 samples for the normal ear, one each for [45,46,48] and two samples from Ref. [22]. Further, 6 of 13 samples for otosclerosis were rejected. Exemplarily, a comparison between the original ER and the ER calculated from the ZEC is shown for the otosclerotic samples in Fig. 10 in the Appendix. The strict consistency check criterion could not be applied to disarticulation cases as in all available 5 samples impedance phase was unphysical. Reason for the inconsistency is mostly expected in the impedance phase, since reflectance and impedance amplitudes mostly fit into the ranges reported for normal ears [30,34], otosclerotic ears [10,30,53], and ears with disarticulation [5,30,53,54].

Automatic antiresonance detection led to exclusion of further 4 samples from Ref. [34] because estimated acoustic lengths were higher than 40 mm and one further sample was excluded, because antiresonance was not detectable at all. Three samples from Ref. [44] had too high acoustic length and 6 samples of this study failed antiresonance detection. One sample from Ref. [48] had too long acoustic length. Furthermore, two samples from Ref. [8] failed the automatic antiresonance detection. Since inconsistencies in the phase of many samples from Ref. [8] were known from previous proof steps and only a few data are available, antiresonances were manually detected in cases where the automatic detection did not succeed. In these two cases of otosclerosis, where the antiresonance could not be found due to problems with phase, the antiresonances could clearly be detected manually in the amplitude of impedance.

A total of 440 normal samples, 7 otosclerosis samples, and 5 disarticulation samples were left after the consistency check and the

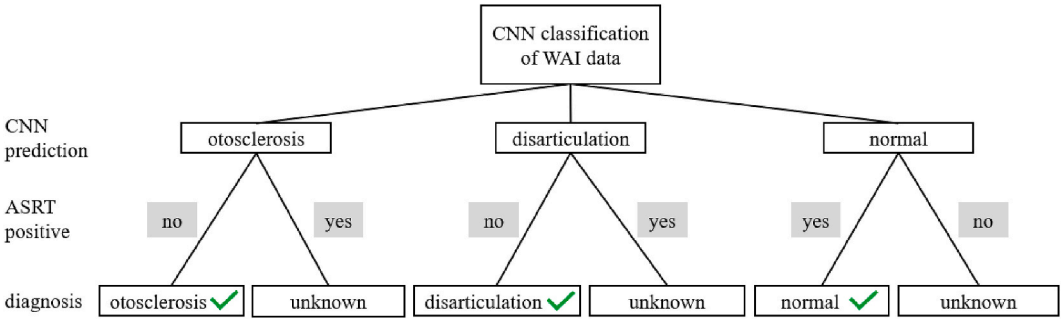



**Fig. 2.** Decision tree for estimating the diagnostic accuracy if the CNN proposal has to be confirmed by ASRT.

detection of antiresonances. Fig. 3 shows the ear canal impedance and reflectance from representative ears from the preprocessed WAI database samples [20]. Three representative ears for each of the cases normal (blue), otosclerosis (red), and disarticulation (green) ear that are closest to the total class median impedance and reflectance are plotted in solid lines. To show the variety, samples with largest and smallest ZEC and ER amplitude below 2000 Hz are plotted with dash-dotted and dotted lines, respectively.

The representative samples are defined as the samples with the least-squares difference summed up over the frequency range 226–2000 Hz of ER and logarithmic ZEC magnitude between sample and the median of these measurands. ER and ZEC magnitude were weighted equally by normalizing all summed differences by the maximum difference across all samples. For the normal ear, samples were required to come from distinctive authors. The low and high impedance samples were defined as the samples with the smallest and largest sum of ER and logarithmic ZEC magnitudes over the frequency range 226–2000 Hz, with the sums of ER and ZEC magnitude equally weighted.

In the frequency range from 226 Hz to 1000 Hz, for most samples shown in Fig. 3, ER is higher in otosclerosis and lower in disarticulations than in normal ears. Furthermore, the first resonance, identifiable by a pronounced minimum in the ER, is in most cases at lower frequencies (500–800 Hz) in disarticulations than in normal ears (800–1200 Hz) and otosclerosis (1100–2500 Hz). However, there are exceptions where samples in parts of the frequency range from 226 Hz to 1000 Hz show a different behavior. The high impedance normal ear sample stands out as having similar ER values to the high impedance otosclerosis sample and having detectable resonances from 1800 Hz to 2200 Hz, but not before. This sample is from Ref. [34] and was recorded using a measurement system that does not monitor static pressure in the ear canal. Thus, it is conceivable that in this sample the eardrum was preloaded by positive pressure created in the ear canal during insertion of the probe, causing the middle ear to stiffen.

The ER changes and resonance shifts compared to the normal ears are caused in otosclerosis by stiffening of the ossicular chain due to fixation of the stapes at the annular ligament and in disarticulation mainly by a reduction of chain stiffness and reduction of damping due to decoupling from the inner ear. For detailed discussions of mechanical cause-and-effect principles, we refer the reader to Refs. [12,30]. Because of the physical dependence between R, ZEC, and ER, the characteristic changes of ER are also seen in the other curves. In the low frequency range up to the first characteristic resonance, ZEC amplitudes are also higher in otosclerosis, analogous to higher ER. The ZEC phase is closer to $-90°$, while the R phase slopes flatter. The opposite is true for disarticulations. At the frequencies where a local minimum in the ER can be detected below 2–3 kHz, a local decrease in ZEC magnitude and R phase and a local increase in ZEC phase can also be detected. We cannot explain the course of the measured ZEC phase in ears with disarticulation. There could be an error in the measurement or evaluation. Plausibly, there would be a local phase increase at the frequency where the first resonance is clearly visible in the ER, as also shown in Ref. [9], Fig. 4.5. For most samples of all classes, there is a prominent minimum in ZEC magnitude in the 2–6 kHz frequency range, which correlates with a sharp phase rise of up to 180° in the ZEC phase. The very high prominence compared to other characteristics indicates very weak damping. This characteristic can be identified as the quarter-wave antiresonance property of the ear canal, since plausible acoustic lengths from the probe to the eardrum are obtained for frequencies in

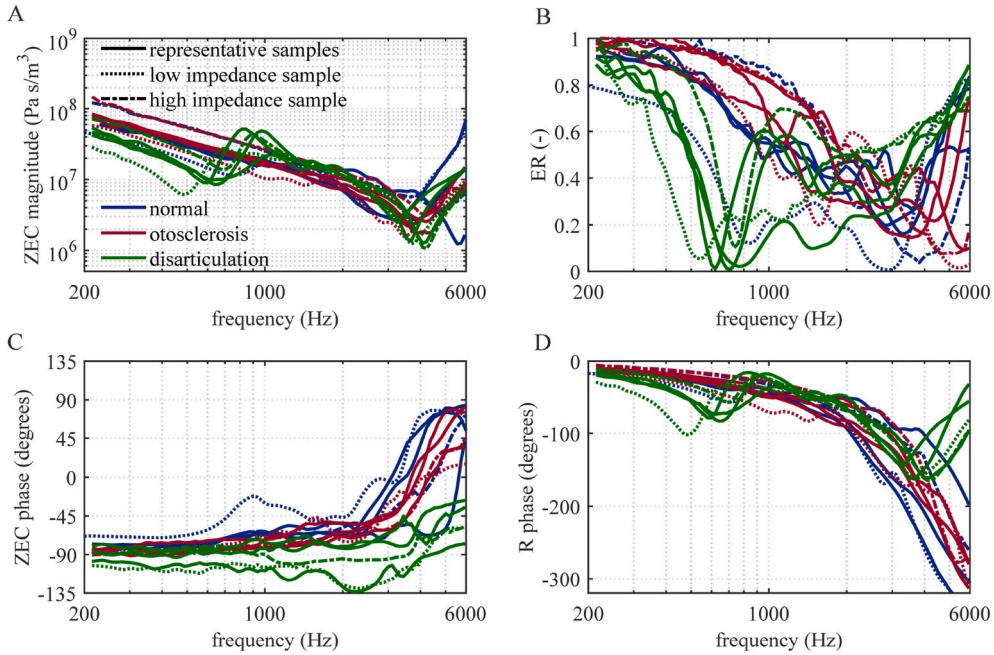



**Fig. 3.** Preprocessed representative ears from the WAI database [20] to show the variety of (A) ear canal impedance magnitude, (B) reflectance magnitude, (C) ear canal impedance phase, and (D) reflectance phase for the cases normal, otosclerosis, and disarticulation. For each case three representative ears that are closest to the total class median impedance and reflectance are plotted together with the samples with largest and smallest impedance and reflectance amplitude below 2000 Hz.

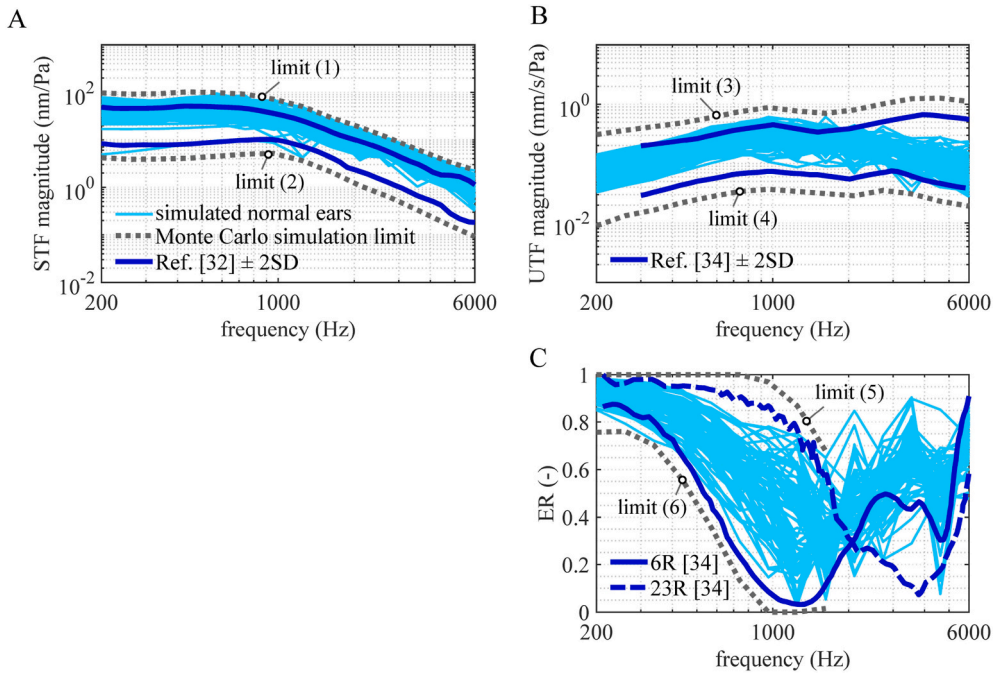



**Fig. 4.** Filtered population of simulated normal ears and references from literature for the measurands (A) stapes transfer function magnitude, (B) umbo transfer function magnitude, and (C) reflectance magnitude. In the ER plot (panel C), the two most extreme samples 23R and 6R with the largest and smallest ER at frequencies below f = 2000 Hz out of the 58 normal ears from Ref. [34] are plotted. The limits for the Monte-Carlo simulations are shown in dotted lines and labeled with numbers in round brackets according to the notation in section 2.3.

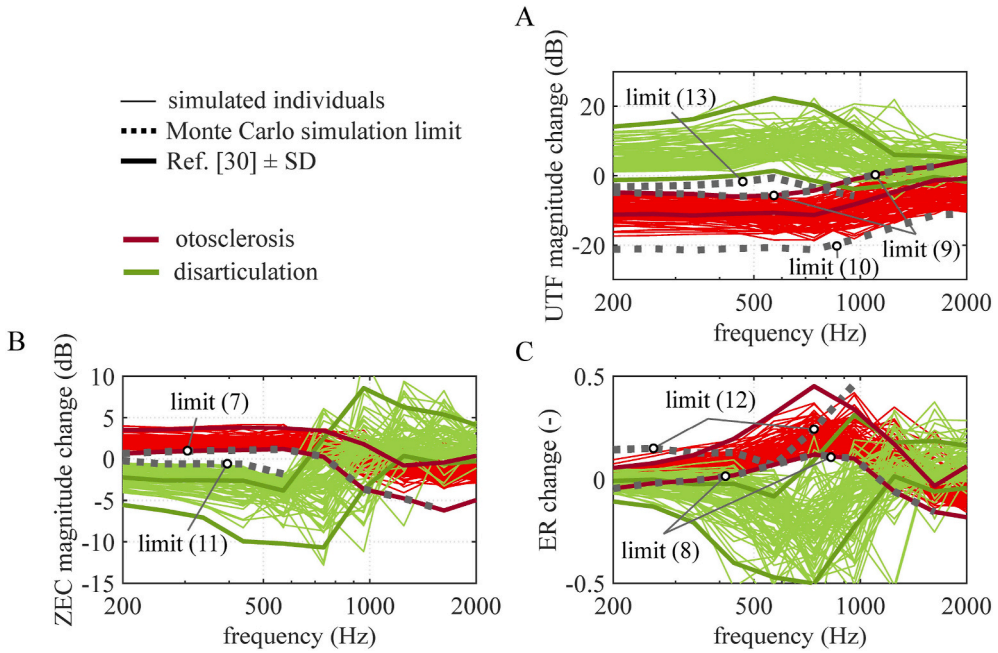



**Fig. 5.** Change from normal due to otosclerosis and disarticulation for the filtered simulated population compared with references from literature. (A) Change in umbo transfer function magnitude. (B) Change in ear canal impedance magnitude. (C) Change in reflectance magnitude. The limits for the Monte-Carlo simulations are shown in dotted lines and labeled with numbers in round brackets according to the notation in section 2.3.

this range.

The three representative normal ears show only weakly pronounced characteristics. The reason for this is that representative samples have been defined in such a way that the least squares error to the median is the smallest. Averaging leads to a smoothing of the curves [55] and the features lose prominence. In addition, the damping may be slightly higher in some people. These samples are then closer to the median in terms of the least squares measure and were therefore likely identified as representative samples. Searching for key features as in Ref. [12] would likely pull samples with more distinct features from the total set. Overall, there is considerable overlap in the WAI measures between the normal ear, otosclerosis, and disarticulation groups, and especially given the high and low impedance samples, purely visual discrimination is nearly impossible.

### 3.2. Generated data set from Monte Carlo simulations

Using Monte Carlo simulations, characteristic populations of normal and pathological ears could be generated, which are consistent with literature data. Application of the literature-based cut-off limits in the Monte Carlo simulations resulted in 23,509 remaining virtual individuals for the normal ear out of the 400,000 simulations. Additional application of the limits for relative changes due to otosclerosis and disarticulation, each with respect to the normal ear, resulted in 79 remaining simulated virtual subjects. Figs. 4 and 5 show the 79 individuals drawn in thin lines together with the filter limits drawn in thick dotted lines, which are numbered in correspondence to their definition in section 2.3. The simulated population is also compared to the mean ± 2SD range of [32] for the STF magnitude, the mean ± 2SD range of [34] for UTF magnitude and the two most extreme ER curves from Ref. [34], all drawn in thick blue lines. The virtual population of normal ears is well centered in the literature interval for ER. However, STF and UTF magnitudes for the normal ear are in general higher than the average of the referenced literature data.

Fig. 5 shows the change from the normal condition due to otosclerosis and disarticulation for the magnitudes of UTF, ZEC and ER. The filtered population of otosclerosis fits well in the mean ± SD range of [30] for ZEC and ER. Some otosclerotic ears have a lower UTF magnitude, which might be caused by the choice of one high annular ligament stiffening factor. Further, not displayed investigations showed that stricter limits for UTF magnitude would not change the sample distribution for ZEC and ER. The population of disarticulation ears fits well in the mean ± SD range of umbo velocity. Some disarticulation samples have less change in ZEC and ER than the mean ± SD range.

Fig. 6 shows the median, 5 % and 95 % percentiles of the initially simulated data from the model with nominal ear canal length and area labeled as before transformation, and the augmented virtual population data after transformations with the sampled ear canal length and area. Table 6 and 7 in the Appendix show the numeric values of the curves in Fig. 6 for the median impedance magnitude and phase and the differences to the 5 % and 95 % percentiles at the octave frequencies 0.25 kHz, 0.5 kHz, 1 kHz, 2 kHz, 4 kHz and at 6 kHz. The median ZEC magnitude of the augmented data at these frequencies decreases by 49 %–84 % up to 2 kHz compared to the initially simulated data. This agrees with the in average longer ear canal in case of augmented data. Also, the phase decreases by 9 %–114 % up to

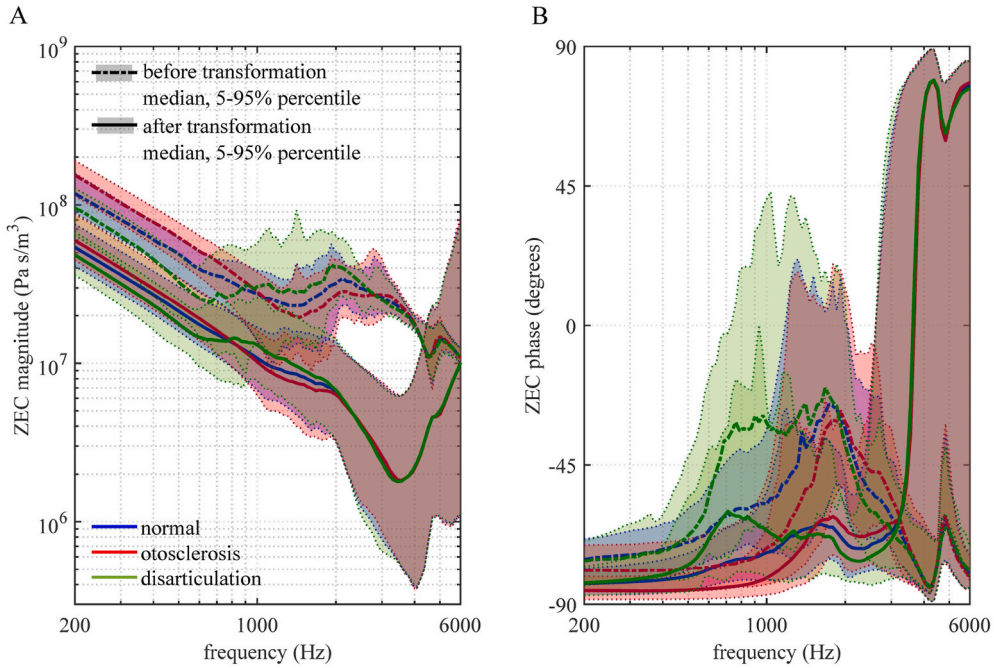



**Fig. 6.** Median, 5 % and 95 % percentiles of the initially simulated data from the model with nominal ear canal length and area (before transformation) and augmented virtual training population after transformations with the sampled ear canal length and area. (A) Ear canal impedance magnitude. (B) Ear canal impedance phase.

2 kHz. The changes in ZEC magnitude and phase are comparable among the three cases. The differences from the 5 % and 95 % percentiles decrease by 16 %–78 % up to 2 kHz for the amplitude and −4 % to −64 % for the phase. Case and frequency specific systematics in these changes of the variation are not visible. At 4 kHz and 6 kHz amplitude and phase variations increased in the data after transformation by up to a factor of 40 compared to before transformation.

### 3.3. Comparison of simulated and measured data

Fig. 7 compares the representative preprocessed test samples derived from the WAI database to the closest simulated individuals used for the CNN training. The representative samples plotted in dashed lines are the same as shown in Fig. 3 and are selected as described in Section 3.1. The simulated individuals plotted in solid lines are determined as the closest samples to the representative samples in terms of summed least squares difference. The differences are calculated for the measurands ER, R phase, ZEC phase, and logarithmic ZEC magnitude. Except ZEC magnitude all measurands were evaluated across the frequency range 226–2000 Hz. ZEC magnitude was evaluated in the range 226–6000 Hz, where the maximum-normalized summed difference below 2000 Hz was weighted by 0.7 and above by 0.3. All other quantities were normalized to the maximum of all samples and were weighted once in the overall difference to find the closest samples.

A detailed description and interpretation of the measured curves can be found in section 3.1. At this point, particular attention will be paid to the characteristics of the simulated data in comparison to the measured data. The simulated curves follow the essential features of the measured data mostly qualitatively well in the frequency ranges used for their assignment. As with the measured data, ER in the frequency range from 226 Hz to 1000 Hz is higher in otosclerosis and lower in disarticulation than in normal ears. The first prominent resonance of the simulations, apparent by a local minimum in the ER, is at lower frequencies (600–735 Hz) for disarticulations and tends to be at higher frequencies (1560–2060 Hz) for otosclerotic ears than for normal ears (935–1675 Hz). This is comparable to the measurements. Consistent with this, below the first resonance, ZEC magnitude is higher, ZEC phase is lower, and R phase tends to be higher in otosclerotic ears than in normal ears. For disarticulations it is the reverse. Above 2 kHz, the qualitative responses in the ZEC phase are comparable. However, the quarter-wave antiresonance is less damped in the simulated data than in the measured data, as evidenced by a more pronounced antiresonance in the simulated ZEC magnitude and a steeper increase in the ZEC phase. Larger deviations occur for the disarticulation ears in the ZEC phase over the entire frequency range and the R phase from about 4 kHz. However, as outlined in Section 3.1, the ZEC phase in the measured disarticulations is not consistent with other publications, while our simulated samples more plausibly reproduce these features. A sharp increase in R phase from 4 kHz as in the measured samples is also not apparent in the measurements in Ref. [30]. Furthermore, there are larger deviations between measured and simulated samples in the ER from about 3 kHz. This range is masked by a gray bar in Fig. 7B and D. As explained in Ref. [12], this frequency range could be influenced by anatomical variations of the tympanic cavity. Sackmann et al. [12] also suggest, based on the investigations in Ref. [30], that this frequency range in ER appears to be less relevant for distinguishing pathologies. Therefore, R phase and ER data in this frequency range were not included in the classification.

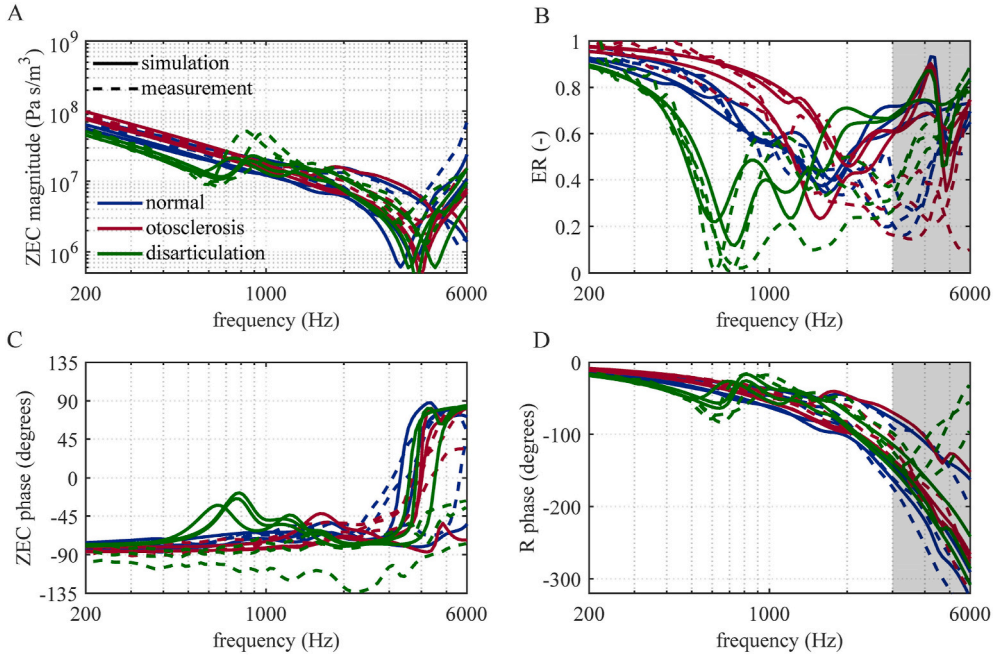



**Fig. 7.** Comparison of representative data from the WAI database [20] to closest simulated WAI data. (A) Ear canal impedance magnitude. (B) Reflectance magnitude. (C) Ear canal impedance phase. (D) Reflectance phase.

To quantitatively evaluate the discrepancies between each simulation curve and the measured curves in the frequency range from 226 Hz to 2000 Hz, the absolute differences for ER, ZEC phase, R phase, and the logarithmic relationship for ZEC magnitude between the three representative measured curves shown and the nearest identified simulated samples were evaluated. The maximum of the absolute difference for each class and the average of the absolute difference over the frequency range from 226 Hz to 2000 Hz for each pair of representative sample and corresponding identified simulation sample were used as the evaluation measure. To put these measures in the context of interindividual variations, the absolute frequency-dependent range was further calculated for the representative measurement samples for all classes for ER, R phase, ZEC phase. For the ZEC magnitude, the absolute logarithmic ratio between the smallest and the largest value per frequency was calculated. From the span or logarithmic ratio, the maximum over all samples per class and the mean per sample over the frequency range 226 Hz–2000 Hz were evaluated. The order of magnitude of these values can already be estimated from Fig. 7. Detailed values are summarized in Table 8 in the Appendix. From the values, it can be summarized that the variations between the representative literature individuals and the nearest identified simulated samples are mostly smaller or in the same order of magnitude as the interindividual variations between the representative literature samples within a class. Both discrepancies and interindividual variations are greater in ears with disarticulation than in otosclerotic and normal ears. This is because the measurement curves of the disarticulation show very prominent, weakly damped resonances. Small deviations in the resonant frequencies lead to relatively large deviations in the magnitudes and phases. However, the magnitude deviations for the first prominent resonance at ER are only in the range of 0.10–0.21, which is comparable to the maximum deviations between simulation and measured samples of the other classes.

### 3.4. Training and classification performance of the CNN

The training and cross-validation of the network took approximately 105 s. The training and validation loss decreased substantially during the first three epochs and further slightly decreased until the end of the training at epoch 20 without evidence for overfitting. At the end of the training a validation accuracy of 92 % was reached. The inference of all 452 test samples took only approximately 0.2 s using a 3.90 GHz Intel CPU and was significantly faster than the training. Table 3 summarizes the number of training, validation, and test samples for each class and in total.

Classification of the data from the WAI database led to the confusion matrix shown in Fig. 8. The percentages of correctly and incorrectly classified cases of the number of cases in each true class are shown as a row-summary. The percentages of correct classifications are the class-wise recalls or sensitivities. The percentage of correctly and incorrectly classified cases of the number of cases in each predicted class are shown as a column-summary. The percentages of correct classification are the class-wise precisions.

Fig. 9 shows the confusion matrix updated by additional application of stapes reflex with a conservative and optimistic scenario of reflex prevalence for normal ears and ears with otosclerosis and disarticulation. Further performance criteria are listed in Table 4. Using ASRT as a confirmation tool of the WAI diagnosis improves specificity of otosclerosis from 85.2 % to 97.8 and 99.1 % depending on the reflex prevalence scenario. Similarly, the specificity of disarticulation is improved from 93.1 % to 98.9 % and 99.6 %, respectively. At the same time sensitivity is not changed. But precision significantly increases as shown in Table 4.

Table 4 compares further the results from the CNN trained with simulation data to some other classification approaches for conductive middle-ear pathologies. Shahnaz et al. [10] used the 90th percentile of the ER from 211 to 6000 Hz of normal ears as a cutoff point to identify otosclerosis cases and separated them from the normal ears in their study. Their study contained 115 ears from 62 normal patients and 28 ears from otosclerotic ears. They reached a sensitivity of 82 % and a specificity of 83 % with this approach. Using third-octaves bands of ER they reached an AUC of 86 % at 500 Hz.

Nakajima et al. [8] plotted absorbance level, which is directly related to ER, versus air-bone gap for 14 patients with otosclerosis, 6 disarticulation patients and 11 semicircular canal dehiscence (SCD) cases. In this 2D plot representation they could build groups which allowed to separate most of these pathologies. Otosclerosis was detected with a sensitivity of 86 % and a specificity of 100 %. Disarticulation was separated with a sensitivity of 83 % and a specificity of 96 %. In their study they did not include normal ears, justified thereby that in clinical settings normal patient could already be separated from patients with conductive hearing loss by audiometry. Keefe et al. [56] used moment analyses of WBT and ambient-pressure absorbance. Their predictors based on moment deviations from the 10th to 90th percentile of normal ears in their study. Their study contained 42 normal ears and 18 ears with conductive hearing loss. They did not distinguish by type of conductive pathology. For a fixed specificity of 90 % they reached up to a sensitivity of 72 % for ambient-pressure absorbance and 94 % for WBT. Their AUC for ambient-pressure absorbance reached 90 % and for WBT 95 %. Nie et al. [13] trained a CNN with a comparable architecture to the CNN of our study with WBT measurement data to separate otosclerotic ears from a control group of normal ears. One CNN in their study was trained directly with the measurement data. Further CNNs were trained with synthetic data, which were gathered with standard approaches of data augmentation from the real data and tuned with transfer learning strategies. They reached an accuracy of 89.7 % and AUC of 94.5 % for the CNN trained with real data and an accuracy of 94.9 % and AUC of 96.8 % for the CNN with the best F1-score trained using data augmentation and transfer

**Table 3**
Number of virtual training and validation samples and real test samples used for the CNN in this study.

| | Normal | Otosclerosis | Disarticulation | Total |
|---|---|---|---|---|
| Virtual training samples | 2520 | 2520 | 2520 | 7560 |
| Virtual validation samples | 640 | 640 | 640 | 1920 |
| Real test samples | 440 | 7 | 5 | 452 |

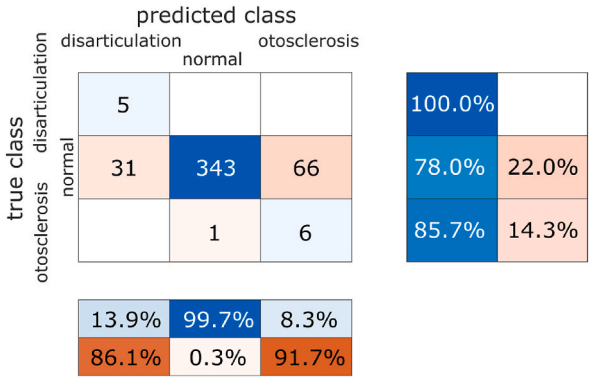



**Fig. 8.** Confusion matrix for the classification of WAI data for the neural network trained with virtual data.

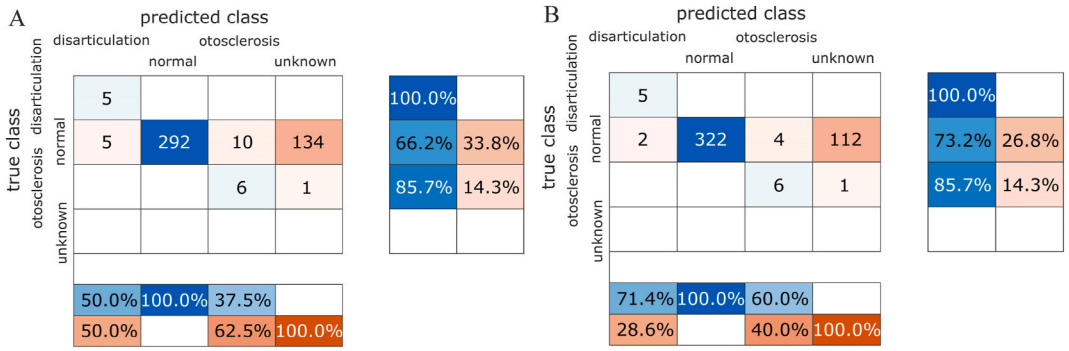



**Fig. 9.** Confusion matrix for the classification of WAI data for the neural network trained with virtual data combined with subsequent confirmation of WAI diagnosis by acoustic stapes reflex. Displayed are the estimated classifications for the (A) conservative acoustic reflex prevalences and (B) optimistic prevalences. The calculated proportions are rounded to natural numbers.

learning. Their study population consisted of 80 otosclerosis patients and 55 normal patients. Further performance parameters are given in Table 4.

## 4. Discussion

### 4.1. Generation of virtually simulated samples

Due to the difficulty in measuring the stiffness and damping parameters of the ligaments and joints in the middle ear and tympanic membrane, there is considerable variation in the parameter ranges published in the literature. When these uncertainties are propagated directly by the model, the variation in simulated middle-ear measurements is much greater than in published normative data. Furthermore, as shown in Ref. [12], it is possible to have parameter combinations within these parameter ranges that lack sufficient sensitivity of individual measurement curves to pathological parameter changes. Therefore, as part of our model validation and training data generation, we apply filters to ensure that the individual virtual simulated responses of the samples fall within the normal and pathological ear ranges from the literature. By combining filters from different measurement metrics of normal and pathological ears, we ensure that only plausible parameter combinations are allowed. Through this calibration, we were able to generate individual simulated training data that both matched the literature normative data and accurately represented the curve characteristics of real individuals (see Fig. 7). These data are well suited for training the CNN. The methods and simulation model have the following limitations.

While qualitative characteristics for the normal, otosclerosis, and disarticulation populations from the Monte-Carlo simulations fit quite well, there are a few quantitative discrepancies to the referenced literature data that are further discussed here. For instance, the STF and UTF magnitudes of the simulated data are generally in the upper magnitude range of the magnitudes allowed by the Monte-Carlo filters. The limits for disarticulation were further chosen less restrictively than for otosclerosis.

In this study, the pathologies were modeled in a very simplified way to keep the already high model complexity moderate and to still be able to use Monte Carlo methods, which require high numbers of model evaluations. Furthermore, the mechanically relevant effects of the pathologies have not yet been investigated in sufficient detail in experimental measurements. For instance, otosclerosis was represented in the model by a stiffening factor of 10,000. However, in reality there are also partial degrees of stiffening that can

**Table 4**
Performance of the simulation-based convolutional neural network (CNN) for wideband acoustic immittance (WAI) data classification with and without confirmation by acoustic stapes reflex threshold (ASRT) from this study in comparison to other classification approaches for conductive middle-ear pathologies. Shown performance parameters are accuracy (acc.), precision (prec.), recall or sensitivity (sens.), F1-score and the area under the curve (AUC) of the receiver operating characteristic. Further abbreviations are ABG: Air-bone gap; SCD: semicircular canal dehiscence; ER: energy reflectance; WBT: wideband tympanometry.

| Study design, cases (samples) | case | acc. | prec. | recall/ sens. | spec. | F1-score | AUC |
|---|---|---|---|---|---|---|---|
| **This study**. WAI-measurements CNN classification, diagnose otosclerosis (7) and disarticulation (5) with separation from normal ears (440) | otosclerosis<br>disarticulation | 78.3 %<br>– | 8.3 %<br>13.9 % | 85.7 %<br>100 % | 85.2 %<br>93.1 % | 15.2 %<br>24.4 % | 91.1 %<br>99.5 % |
| **This study**. WAI-measurements CNN classification + ASRT confirmation, diagnose otosclerosis (7) and disarticulation (5) with separation from normal ears (440), *conservative scenario* | otosclerosis<br>disarticulation | 66.9 %<br>– | 37.5 %<br>50.0 % | 85.7 %<br>100 % | 97.8 %<br>98.9 % | 52.2 %<br>66.7 % | – |
| **This study**. WAI-measurements CNN classification + ASRT confirmation, diagnose otosclerosis (7) and disarticulation (5) with separation from normal ears (440), *optimistic scenario* | otosclerosis<br>disarticulation | 73.7 %<br>– | 60.0 %<br>71.4 % | 85.7 %<br>100 % | 99.1 %<br>99.6 % | 70.6 %<br>83.3 % | – |
| [10]: ER measurement, separate otosclerotic ears (28) from normal ears (115) | otosclerosis | – | – | 82 %[1] | 83 %[1] | – | (86 %)[2] |
| [8]: ER + ABG separate otosclerosis (14), disarticulation (6), SCD (11) | otosclerosis<br>disarticulation<br>SCD | – | – | 86 %<br>83 %<br>100 % | 100 %<br>96 %<br>95 % | – | – |
| [56]: WAI, separate conductive pathology (18) from normal (42) | conductive pathology | – | – | 72 % | 90 % | – | 90 % |
| [56]: WBT, separate conductive pathology (18) from normal (42) | conductive pathology | – | – | 94 % | 90 % | – | 95 % |
| [13]: CNN trained with measured WBT data, separate normal ears (55) from otosclerosis (80) | otosclerosis | 89.7 % | 89.4 % | 89.2 % | – | 89.3 % | 94.5 % |
| [13]: CNN trained with measured WBT and usage of data augmentation and transfer learning techniques, separate normal ears (55) from otosclerosis (80) | otosclerosis | 94.9 % | – | – | – | 94.7 % | 96.8 % |

[1] Performance of 'ER below the 90th percentile of the normal ears from 211 to 6000 Hz' [10] as cutoff point.
[2] Performance of 'ER averaged in 1/3 octave bands centered at 500 Hz' [10].

occur with significantly lower changes in the stiffness of the annular ligament. Additionally, translational and rotational stiffnesses are not likely to be equally affected by otosclerosis in the patient, as otosclerosis usually begins with local stiffening at one side of the annular ligament. In the case of disarticulation, the exact mechanically relevant changes have hardly been investigated by measurement. Some modelers have assumed that separation of the incudostapedial joint is accompanied by a pretension loss throughout the ossicular chain aside from decoupling the stapes and inner ear from the ossicular chain [11,12]. However, this has not yet been proven, but indicated by measurements of tympanic membrane shape changes during separation of the incudostapedial joint published by Ref. [57].

Since both the normal ear and pathological ear criteria are required to be satisfied in this study, such assumptions in one of the cases may influence the populations of other cases. On the other hand, this linkage also ensures that generally implausible parameter combinations are avoided, as explained in Ref. [12], and it allows the limited available measurement data to be utilized more extensively. Another reason for the offset of the virtual data could be that the number of Monte Carlo simulations (400,000) was still very small considering the large number of uncertain model parameters (N = 66). Therefore, statistical statements about the mean of the simulated virtual population and comparisons with populations from measurements with sometimes also very small sample sizes can only be made to a limited extent. With iterative Monte-Carlo algorithms, convergence for the defined filter criteria could be targeted, but this would only make sense if more measurement data were available, especially for the pathologies.

The chosen methodology of first drawing all samples and then filtering them gives insight into systematic deviations between simulations and different measurement data sources. Applying the filters in varying combinations, makes it possible to identify causes for deviations. For example, the superimposed application of the filter criteria for the normal ear and the pathological ears leads to the exclusion of many samples near the lower limit of the STF and UTF magnitudes of the normal ear. This is not the case if only the normal ear filter criteria are applied. The exclusion of these samples can mainly be attributed to the filter criteria for the lower limit of the change in ZEC and ER due to otosclerosis (limits (7) and (8)). Ears with stiffer ligaments are predominantly excluded. These ears exhibit limited sensitivity to additional stiffening of the ossicular chain caused by otosclerosis.

One limitation of the model in this study is, that except for variations in the tympanic membrane thickness and modifications in the ear canal length and cross-section, we do not account for further geometry variations like tympanic membrane shape, and ossicles' positions and orientations. A fully parameterized, stochastic middle ear model would be required to account for such geometric variations. Such a model would be very helpful and an important research goal, but it does not yet exist.

It has been observed that effects like evanescence or the oblique insertion of the probe affects measurement data only at higher

frequencies [58,59]. We did not take these two influences into account in the simulation, and therefore made sure in the post-processing by the filters that the limits of the filters for the reflectance already end at 1600 Hz. Age-related changes in middle ear material properties were also not considered.

### 4.2. Classification performance and clinical relevance

This study cannot definitively determine whether training with virtual data leads to better or worse performance than training with measurement data alone or hybrid training with both measurement and virtual data. An insufficient number of individual measurements for pathologies to train a CNN prevents such an assessment. However, when compared to other studies that have trained classifiers using measured data, the CNN trained with virtual data shows at least similar performance. The CNN trained with virtual WAI simulation data in this study achieved sensitivities, specificities and AUC values for the diagnosis of otosclerosis and disarticulations that are comparable to corresponding values of other approaches in the literature for classifying WAI data. For otosclerosis, a sensitivity of 85.7 %, specificity of 85.2 %, and AUC of 91.1 % were achieved. Shahnaz et al. [10] achieved a sensitivity of 82 %, specificity of 83 %, and AUC of 86 % for ER data of patients with otosclerosis. Keefe et al. [56] achieved with ER data a specificity of 72 %, a sensitivity of 90 %, and an AUC of 90 % for conductive pathologies in general.

In this study, mostly normal ears were misclassified. If measurement of the acoustic stapes reflex threshold is performed as a second test to confirm the WAI diagnosis, this study estimates that higher overall specificities for the pathologies can be achieved. For otosclerosis, a specificity of up to 99.1 % and a sensitivity of 85.7 % could be achieved in this study. These values are comparable to the specificity of 100 % with sensitivity of 86 % of [8], who suggest a combined evaluation of ER and air-bone gap. However, unlike classical audiometry, stapes reflex threshold measurement can be performed without active patient participation, i.e., even in neonates. Using the procedures of this study, comparable values to Ref. [8] were also obtained for disarticulation. While the number of test data from measurement data of normal ears was sufficiently high, only 7 otosclerosis and 5 disarticulation measurements in total could be used as test data for the CNN in this study. This in turn leads to the fact that the sensitivity and specificity of the classification are very susceptible to single misclassifications. In case of otosclerosis and disarticulation one measurement represents 14.3 % and 20 % of the total number of each class. Hence, to evaluate the diagnostic reliability and scalability of the obtained results, more measurement data of the pathological cases is needed. Databases like the one from Ref. [20], which was used in this study and contains publicly available WAI measurement data, offer a great contribution to research and should be fed with further data.

The small number of pathologic measurement data used as test data with enough available normal ear data ($n = 440$) also results in very inhomogeneous class sizes. This in turn leads to a much lower precision of 8.3–60.0 % for otosclerosis in our study compared to other studies (89.4 %). Considering the number of pathologies compared to normal ears in this study, a prevalence of pathologies of 1.1–1.6 % can be calculated. This value is comparable to the 0.2–1.3 % prevalence [60–62] of clinical otosclerosis with hearing impairment in the total adult population. Thus, our results provide performance values for a test applied to the entire population. However, other prevalences might be more realistic for a clinical population. When the rows of the confusion matrix for classification of WAI data are multiplied to give the same class sizes for all classes (otosclerosis, disarticulation, normal), a precision of 85.1 % and 93.4 % is calculated for otosclerosis and disarticulation, respectively. The corresponding F1-scores also increase to 85.4 % and 96.6 %, respectively, and the accuracy increases to 87.9 %. These values are comparable to those reported in the study by Ref. [13] for the identification of otosclerotic ears using WBT.

By varying the static ear canal pressure, the WBT measurement, unlike the WAI measurement, also contains information about the nonlinear behavior of the soft tissues of the middle ear, such as the annular ligament. Pathologies which locally alter this nonlinear stiffening behavior, might be more distinctive in WBT than in WAI only, as pointed out in Ref. [15]. Therefore, WBT is expected to contain additional information relevant to diagnosis compared with the WAI measures included in this study. This can be argued, for example, based on the study of [56]. Using WBT measurement data for otosclerosis, they achieved a sensitivity of 94 % with a specificity of 90 % and an AUC of 95 %, whereas WAI measurement data only achieved a sensitivity of 72 % with a specificity of 90 % and AUC of 90 %. Also, [43] were able to achieve an improvement of the AUC from 88 % to 95 % for discrimination of otosclerosis and normal ears with WBT compared to WAI with a multivariate predictor with three reflectance variables. Currently, only the [43] study in the WAI database from Ref. [20] includes pressurized WAI (WBT) measurements from a few otosclerotic individuals with limited frequency and pressure resolution. Unfortunately, there are no WBT measurements available for individuals with disarticulation in this database. Due to the scarce data, we cannot quantitatively calibrate a model based on WBT data at this time. The classification of WAI data in this study is also affected by the uncertainty of unbalanced static pressure at the tympanic membrane. By using WBT measurements, this uncertainty can be avoided, which should reduce the overall variance of each class and thus improve the discriminability of the classes. In addition, the variation of the measurements can be further reduced by more reproducible measurement methods. The measurement data from the database we used for this study came from measuring devices with different levels of technical development. For some datasets, the data was also not physically consistent. The model for generating training data was fitted only to the limits given in the literature using the Monte Carlo method. A true statistical fit of the model to statistical curves would increase accuracy, and additional data augmentation with noise could make the method more robust to new data. Transfer learning of a network pre-trained on virtual data to real measured data could probably also achieve higher accuracies. However, these methods require more pathological measurement data.

The phase of WAI data was included in this study. However, the measuring instruments typically used in clinical routine only output the reflectance and the impedance is not available as a tool for diagnosis. It is expected that the impedance phase provides characteristic information relevant for the differentiation of pathologies in frequency ranges that are not as sensitive for magnitude changes. This is because resonance shifts that occur on the frequency axis are mapped more clearly in the phase information than on

the magnitude axis. After resonances, the magnitude difference between normal and pathological cases is often lost in interindividual variations while the phase still retains a delayed offset, as seen in the averaged results from Ref. [30]. Due to the limited amount of available individual test data and data to calibrate the simulated phases in the model, a final assessment of the added value of phase in diagnosis cannot be made in this study. However, the consistency check during post-processing of the data already shows how important the phase is. If consistency were not checked, the sensitivity for otosclerosis of the CNN in this study would drop from 85.7 % to only 61.5 % without significantly changing the specificities for the investigated pathologies.

Furthermore, the phase information in the impedance can be useful to achieve sufficient quality and reproducibility of the measurement data, e.g. also to detect leakage [21] based on the criteria in Ref. [45] already directly during the measurement. In this case, the measurement could simply be repeated or extended. Since providing phase information from WAI measurements does not require any hardware changes, we strongly recommend not to dismiss this information in diagnostic tests but make it accessible at least for future evaluations.

Furthermore, many measurement devices currently used in clinical practice assume fixed values for the ear canal area at the probe position and do not account for the measurement position when calculating the WAI measurement quantities. However, as obvious from the equations for the calculation of reflectance from impedance, differences in the assumption of the ear canal area in the calculation of the characteristic impedance, which is needed to calculate ER from ZEC, to the ear canal area inherently accounted for in the measurements of impedance, introduce an error to ER. Ear canal area and measurement position further significantly influence impedance. In this study, we attempted to include the area of the ear canal in the classification. However, assumptions had to be made about the cross-sectional area of the ear canal based on the age of the subjects, because this information was often missing from the data or was obtained using procedures that are not very reproducible. The length of the ear canal is taken into account by the anti-resonance frequency, as described in section 2.2. It would be helpful if these variables could be reliably measured in future diagnostic devices.

Despite these limitations, the preliminary results of this study for the detection of otosclerosis and disarticulation using a CNN trained solely with simulated WAI data are quite promising. Without the use of virtually generated data, a considerable amount of measurement data of the various pathologies would be required to train a network that can classify the pathologies with similar accuracy. Conducting clinical studies to obtain such measurement data is connected to a high cost and time effort. Additionally, pathological ears are usually only included in such studies if their respective pathology has been confirmed by a subsequent surgery. Thus, only pathologies that are in an advanced stage are examined. However, the generation of WAI data based on a FE model also allows the modeling of pathologies that are in early stages and not as pronounced yet. This could enable earlier detection of the pathologies and thus lead to better treatment opportunities. The method also has the potential to diagnose problems in reconstructed ears like a dislocated prothesis. For such cases, measurement data from well-defined situations, where the exact issue is known, can hardly be obtained. The presented approach could be used to simulate such cases and thus may contribute to their detection.

## 5. Conclusion

The main finding of the present study is that simulated diagnostic data are similarly suitable for training a convolutional neural network as measured data. We focussed on the discrimination of normal, otosclerotic and disarticulated ears. The FEM simulation model employed is highly complex. We used the Monte Carlo method to generate sufficient data to represent both qualitatively and in most cases quantitatively the variability of measurements in normal, otosclerotic and disarticulated ears. The convolutional neural network trained with the simulated WAI data achieved an AUC of the receiver operating characteristic of 91.1 %, a sensitivity of 85.7 %, and a specificity of 85.2 % for otosclerosis. For disarticulation, an AUC of 99.5 %, sensitivity of 100 %, and specificity of 93.1 % were achieved. Furthermore, it was estimated that specificity could potentially be significantly increased to about 99 % if stapes reflex threshold measurements were used to confirm the diagnosis based on wideband acoustic immittance measurements. The overall classification performance is thus comparable to other classification approaches in the literature that rely only on measured data as training data. In this context, simulated data offer the advantage that pathological cases and early-stage cases can also be trained for which very little measurement data is available.

To evaluate the clinical potential of this preliminary study, the approach would need to be tested on more real measurement data and additional pathologies would need to be considered.

## Data availability

Data associated with our study has not been deposited into a publicly available repository. Data will be made available on request.

## Ethics declarations

Review and/or approval by an ethics committee was not needed for this study because we did not perform any measurements. The measurement data used were obtained from a freely available database that contains data only from published and peer-reviewed papers, which are cited in this paper.

## CRediT authorship contribution statement

**Michael Lauxmann:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology,

Funding acquisition, Formal analysis, Conceptualization. **Felix Viehl:** Writing – original draft, Software. **Barbara Priwitzer:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Benjamin Sackmann:** Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

This work was funded by the Vector foundation and the German Federal Ministry for Economic Affairs and Climate Action within the scope of a Central Innovation Programme for small and medium-sized enterprises (SMEs), KK5206106BA1.

## APPENDIX A

**Table 5**
Overview of studies from the WAI database [20] that contain data that meet the requirements described in Section 2.2.

| Reference | No. of subjects* | No. of samples* | Age range | Ear condition | Inclusion criterion | Measurement device |
|---|---|---|---|---|---|---|
| Abur et al. [44] | 7 | 86 | 19–22 | normal | Not known | Not known |
| Groon et al. [45] | 21 | 21 | 19–30 | normal | air-conduction behavioral thresholds, 226-Hz tympanometry, normal otoscopy, no middle-ear surgery or disease during the preceding 2 years | Custom |
| Lewis and Neely [22] | 15 | 28 | 21–45 | normal | pure-tone audiometry, 226 Hz tympanometry, and normal otoscopic inspection | Custom |
| Lewis [46] | 30 | 60 | 22–33 | normal | pure-tune audiometry, present ipsilateral broadband-noise acoustic reflex, and 226-Hz tympanometry | custom |
| Merchant et al. [47] | 5 | 5 | 19–31 | normal | normal hearing screening, otoscopy, and 226 Hz tympanometry, no history of balance disorder, dizziness, neurologic pathologies, or middle-ear surgery except tympanostomy tube placement | Titan probe, custom calibration and evaluation |
| Rosowski et al. [34] | 29 | 58 | 22–64 | normal | no history of significant middle-ear disease, no recent otologic surgery, and no abnormalities in their auricles or eardrums. In addition, audiometric measurements, air-bone gap, and 226-Hz tympanometry values had to be within the normal range. | Mimosa Acoustics HearID |
| Shaver and Sun [48] | 48 | 48 | 18–35 | normal | no history of chronic otitis media, active flu, abnormal otoscopy and with normal pure-tone audiometry, 226 Hz tympanometry and tympanometric peak pressure | Interacoustics research prototype with custom ER probe and Interacoustics AT235 tympanometer |
| Sun [49] | 84 | 84 | 18–35 | normal | otoscopy, air conduction pure-tone audiometry, air-bone gap and normal tympanometric peak pressure, no history of abnormal middle ear condition, hearing loss, family otosclerosis or recent flu. | Interacoustics research prototype with custom ER probe and Interacoustics AT235 tympanometer |
| Voss and Allen [4] | 10 | 10 | 18–24 | normal | normal hearing thresholds with standard clinical methods | custom |
| Voss et al. [50] | 12 | 60 | 19–42 | normal | negative history of middle-ear problems, normal hearing thresholds and tympanograms | Mimosa Acoustics HearID |
| Nakajima et al. [8] | 20 | 20 | 22–72 | 14 otosclerosis 6 disarticulation | diagnosis of the respective pathology based on the air-bone gap had to be present and confirmed during a subsequent surgery | Mimosa Acoustics HearID |

* Number of normal, otosclerosis, or disarticulation subjects. Some studies contain also subjects of other cases irrelevant for this study.

**Table 6**
Median of ZEC magnitude, and differences of 5 % and 95 % percentiles of ZEC magnitude to median ZEC magnitude of the (a) augmented virtual population data after transformations with the sampled ear canal length and area and (b) in round brackets initially simulated data from the model with nominal ear canal length and area.

| Freq. (Hz) | | 250 | 500 | 1000 | 2000 | 4000 | 6000 |
|---|---|---|---|---|---|---|---|
| Normal ZEC Mag. (Pa s/m$^3$) ·10$^6$ | $P_{95}$–$P_{50}$ | 18.0 (30.8) | 9.9 (13.0) | 6.6 (10.9) | 6.8 (19.8) | 11.4 (2.3) | 64.1 (1.5) |
| | $P_{50}$ | 41.2 (95.8) | 20.1 (49.0) | 10.2 (27.7) | 6.3 (32.1) | 2.5 (17.6) | 10.6 (10.8) |
| | $P_{50}$–$P_5$ | 8.6 (24.3) | 5.0 (14.2) | 3.8 (13.8) | 3.4 (12.5) | 1.9 (1.3) | 9.5 (0.8) |
| Otosclerosis ZEC Mag. (Pa s/m$^3$) ·10$^6$ | $P_{95}$–$P_{50}$ | 23.6 (30.0) | 12.5 (14.8) | 6.6 (9.3) | 7.4 (16.8) | 11.4 (2.4) | 69.4 (1.6) |
| | $P_{50}$ | 46.0 (123.6) | 22.3 (62.2) | 9.6 (28.3) | 6 (26.9) | 2.5 (17.8) | 10.6 (10.8) |
| | $P_{50}$–$P_5$ | 11.9 (32.2) | 6.0 (18.5) | 3.8 (14.5) | 3.2 (11.4) | 2.0 (1.2) | 9.5 (0.8) |
| Disarticulation ZEC Mag. (Pa s/m$^3$) ·10$^6$ | $P_{95}$–$P_{50}$ | 17.3 (27.4) | 9.4 (16.1) | 11.8 (26.7) | 7.9 (22.9) | 11.7 (1.5) | 62.4 (1.5) |
| | $P_{50}$ | 37.3 (76.4) | 16.8 (32.8) | 12.8 (30.4) | 6.5 (40.9) | 2.5 (17.4) | 10.6 (10.9) |
| | $P_{50}$–$P_5$ | 9.6 (25.1) | 4.6 (13.1) | 6.4 (15.2) | 3.6 (16.7) | 1.9 (1.6) | 9.5 (0.8) |

**Table 7**
Median of ZEC phase, and differences of 5 % and 95 % percentiles of ZEC phase to median ZEC amplitude of the (a) augmented virtual population data after transformations with the sampled ear canal length and area and (b) in round brackets initially simulated data from the model with nominal ear canal length and area.

| Freq. (Hz) | | 250 | 500 | 1000 | 2000 | 4000 | 6000 |
|---|---|---|---|---|---|---|---|
| Normal ZEC phase (degrees) | $P_{95}$–$P_{50}$ | 4.8 (6.8) | 6.6 (10.8) | 22.4 (24) | 29.4 (36.6) | 12.9 (16.9) | 15.7 (6.8) |
| | $P_{50}$ | −83.2 (−74.5) | −80.8 (−69.9) | −74.7 (−57.5) | −70.2 (−34.1) | 74.0 (−82.2) | 69.7 (−80.3) |
| | $P_{50}$–$P_5$ | 3.9 (8.4) | 5.3 (9.1) | 7.8 (14.4) | 10.0 (20.5) | 158.8 (4.3) | 151.2 (4.6) |
| Otosclerosis phase (degrees) | $P_{95}$–$P_{50}$ | 4.8 (8.3) | 5.4 (9.5) | 16.7 (22) | 37.4 (46.5) | 13 (19.8) | 13.5 (7.0) |
| | $P_{50}$ | −85.9 (−79.0) | −85.4 (−78.4) | −81.7 (−71.8) | −66.0 (−31.7) | 74.1 (−82.5) | 72.1 (−81.1) |
| | $P_{50}$–$P_5$ | 3.0 (6.9) | 3.1 (7.4) | 5.6 (11.8) | 12.9 (19.9) | 159.2 (4.0) | 153.8 (4.1) |
| Disarticulation phase (degrees) | $P_{95}$–$P_{50}$ | 5.5 (8.6) | 24.0 (25.0) | 44.1 (72.1) | 20.4 (42.7) | 13.9 (15.7) | 16.0 (7.8) |
| | $P_{50}$ | −82.4 (−75.3) | −77.5 (−67.7) | −67.2 (−31.4) | −74.7 (−38.2) | 72.7 (−81.8) | 69.3 (−79.1) |
| | $P_{50}$–$P_5$ | 5.1 (7.8) | 8.6 (13.9) | 12.9 (33.6) | 7.3 (20.0) | 156.8 (4.9) | 150.7 (5.8) |

**Table 8**
Discrepancies between each measurement curve and the nearest identified simulation curve in the frequency range from 226 Hz to 2000 Hz and interindividual variations of the characteristic measurement samples per class.

| Measurand | Case | Max. variation of literature samples[1] | Max. abs. difference between literature and simulated samples[2] | mean of variation of literature samples[3] | mean abs. difference between literature and simulated samples[4] |
|---|---|---|---|---|---|
| ZEC mag. (dB) | Normal | 5.5 | 3.4 | 1.8 | 0.5–2.3 |
| | Otosclerosis | 3.3 | 3.7 | 1.3 | 0.9–1.2 |
| | Disarticulation | 12 | 9 | 3.3 | 1.5–3.1 |
| ZEC phase (degrees) | Normal | 14 | 13 | 6 | 3–4 |
| | Otosclerosis | 25 | 12 | 13 | 1–4 |
| | Disarticulation | 47 | 85 | 27 | 17–43 |
| ER (−) | Normal | 0.14 | 0.16 | 0.06 | 0.03–0.04 |
| | Otosclerosis | 0.36 | 0.20 | 0.16 | 0.02–0.05 |
| | Disarticulation | 0.51 | 0.42 | 0.23 | 0.09–0.16 |
| ER phase (degrees) | Normal | 40 | 24 | 9 | 2–12 |
| | Otosclerosis | 32 | 21 | 10 | 3–6 |
| | Disarticulation | 45 | 42 | 13 | 9–13 |

[1] Maximum over all frequencies in the range 226 Hz–2000 Hz of the (a) absolute range in ER, R phase, and ZEC phase of all measured samples per class, (b) absolute logarithmic ratio between the smallest and the largest value per frequency.

[2] Maximum over all frequencies in the range 226 Hz–2000 Hz and over all samples per class of the (a) absolute differences between each measurement curve and the nearest identified simulation curve for ER, R phase, and ZEC phase, (b) absolute logarithmic ratio between the smallest and largest value per frequency for ZEC magnitude between each measurement curve and the nearest identified simulation.

[3] Same definitions as (1) but with mean instead of maximum.

[4] Same definitions as (2) but with mean instead of maximum, and mean was calculated per pair of corresponding simulation and measurement samples.

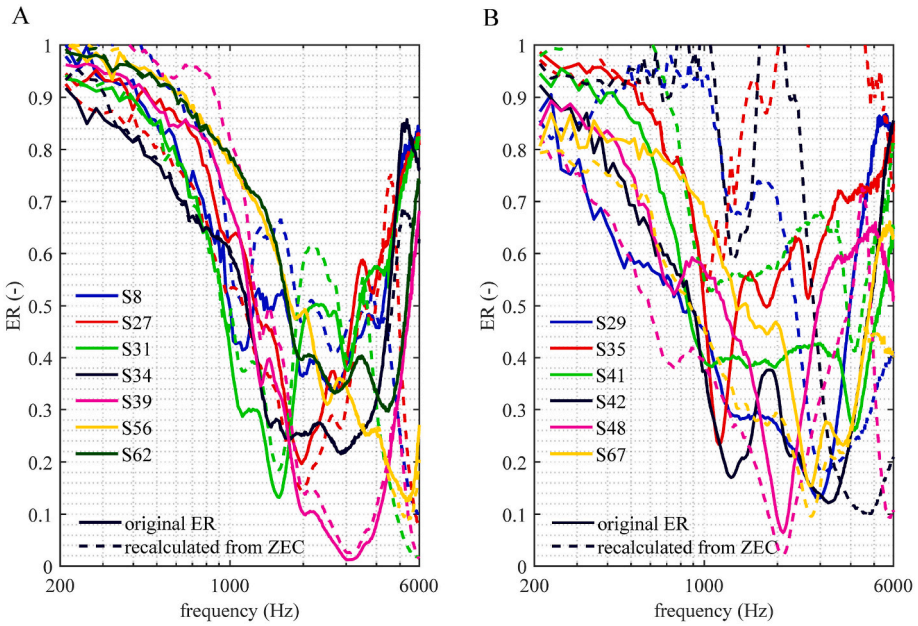



**Fig. 10.** Comparison of original ER and ER recalculated from original complex ZEC from the WAI-database [20] of the otosclerosis data from Ref. [8]. The comparison is split into the samples which where (A) consistent in ER and ZEC according to the criteria defined in this study and (B) inconsistent samples in ER and ZEC. The legend indicates the subject numbers as defined in the data from Ref. [8].

## References


[1] L.L. Hunter, M.P. Feeney, J.A. Lapsley Miller, P.S. Jeng, S. Bohning, Wideband reflectance in newborns: normative regions and relationship to hearing-screening results, Ear Hear. 31 (2010) 599–610, https://doi.org/10.1097/AUD.0b013e3181e40ca7.

[2] U. Fisch, G.O. Acar, A.M. Huber, Malleostapedotomy in revision surgery for otosclerosis, Otol. Neurotol. 22 (2001) 776–785, https://doi.org/10.1097/00129492-200111000-00011.

[3] S.N. Merchant, J.J. Rosowski, Acoustics and mechanics of the middle ear, in: A.J. Gulya, L.B. Minor, D.S. Poe, J.A. Gulya (Eds.), Glasscock-Shambaugh's Surgery of the Ear//Glasscock-Shambaugh Surgery of the Ear, sixth ed.),, People's Medical Publishing House; PMPH USA Ltd, Shelton, 2010, pp. 49–72.

[4] S.E. Voss, J.B. Allen, Measurement of acoustic impedance and reflectance in the human ear canal, J. Acoust. Soc. Am. 95 (1994) 372–384, https://doi.org/10.1121/1.408329.

[5] M.P. Feeney, I.L. Grant, L.P. Marryott, Wideband energy reflectance measurements in adults with middle-ear disorders, J. Speech Lang. Hear. Res. 46 (2003) 901–911, https://doi.org/10.1044/1092-4388(2003/070.

[6] N. Shahnaz, K. Bork, Wideband reflectance norms for Caucasian and Chinese young adults, Ear Hear. 27 (2006) 774–788, https://doi.org/10.1097/01.aud.0000240568.00816.4a.

[7] N. Shahnaz, N. Longridge, D. Bell, Wideband energy reflectance patterns in preoperative and post-operative otosclerotic ears, Int. J. Audiol. 48 (2009) 240–247, https://doi.org/10.1080/14992020802635317.

[8] H.H. Nakajima, D.V. Pisano, C. Roosli, M.A. Hamade, G.R. Merchant, L. Mahfoud, C.F. Halpin, J.J. Rosowski, S.N. Merchant, Comparison of ear-canal reflectance and umbo velocity in patients with conductive hearing loss: a preliminary study, Ear Hear. 33 (2012) 35–43, https://doi.org/10.1097/AUD.0b013e31822ccba0.

[9] G.R. Merchant, Functional Measurements of Ear Pathology in Patients and Cadaveric Preparations, 2014.

[10] N. Shahnaz, K. Bork, L. Polka, N. Longridge, D. Bell, B.D. Westerberg, Energy reflectance and tympanometry in normal and otosclerotic ears, Ear Hear. 30 (2009) 219–233, https://doi.org/10.1097/AUD.0b013e3181976a14.

[11] X. Zhang, R.Z. Gan, Finite element modeling of energy absorbance in normal and disordered human ears, Hear. Res. 301 (2013) 146–155, https://doi.org/10.1016/j.heares.2012.12.005.

[12] B. Sackmann, P. Eberhard, M. Lauxmann, Parameter identification from normal and pathological middle ears using a tailored parameter identification algorithm, J. Biomech. Eng. (2021), https://doi.org/10.1115/1.4052371.

[13] L. Nie, C. Li, F. Marzani, H. Wang, F. Thibouw, A.B. Grayeli, Classification of wideband tympanometry by deep transfer learning with data augmentation for automatic diagnosis of otosclerosis, IEEE J Biomed Health Inform 26 (2022) 888–897, https://doi.org/10.1109/JBHI.2021.3093007.

[14] J.V. Sundgaard, P. Bray, S. Laugesen, J. Harte, Y. Kamide, C. Tanaka, A.N. Christensen, R.R. Paulsen, A deep learning approach for detecting otitis media from wideband tympanometry measurements, IEEE J Biomed Health Inform 26 (2022) 2974–2982, https://doi.org/10.1109/JBHI.2022.3159263.

[15] B. Sackmann, E. Dalhoff, M. Lauxmann, Model-based hearing diagnostics based on wideband tympanometry measurements utilizing fuzzy arithmetic, Hear. Res. 378 (2019) 126–138, https://doi.org/10.1016/j.heares.2019.02.011.

[16] B. Sackmann, M. Lauxmann, D. Burovikhin, Model-based hearing diagnosis based on Monte-Carlo parameter estimation and artificial neural networks, in: 9th International Symposium on Middle Ear Mechanics in Research and Otology, 2022, p. 38.

[17] H. Motallebzadeh, M. Deistler, F. Schönleitner, J.H. Macke, S. Puria, Training a machine-learning differential diagnostic tool for conductive hearing loss using mechanistic models, in: 9th International Symposium on Middle Ear Mechanics in Research and Otology, 2022, p. 60.

[18] M. Golabbakhsh, W.R.J. Funnell, Use of simulated data to explore the application of optical coherence tomography for classifying middle-ear pathologies, J. Acoust. Soc. Am. 154 (2023) 2790–2799, https://doi.org/10.1121/10.0022051.

[19] A. Ebrahimian, H. Mohammadi, J.J. Rosowski, J.T. Cheng, N. Maftoon, Inaccuracies of deterministic finite-element models of human middle ear revealed by stochastic modelling, Sci. Rep. 13 (2023) 7329, https://doi.org/10.1038/s41598-023-34018-w.

[20] S.E. Voss, An online wideband acoustic immittance (WAI) database and corresponding website, Ear Hear. 40 (2019) 1481, https://doi.org/10.1097/AUD.0000000000000790.

[21] J.J. Rosowski, S. Stenfelt, D. Lilly, An overview of wideband immittance measurements techniques and terminology: you say absorbance, I say reflectance, Ear Hear. 34 (Suppl 1) (2013) 9S–16S, https://doi.org/10.1097/AUD.0b013e31829d5a14.

[22] J.D. Lewis, S.T. Neely, Non-invasive estimation of middle-ear input impedance and efficiency, J. Acoust. Soc. Am. 138 (2015) 977–993, https://doi.org/10.1121/1.4927408.

[23] S.E. Voss, N.J. Horton, K.E. Fairbank, L. Xia, L.R.K. Tinglin, K.D. Girardin, Measurements of ear-canal cross-sectional areas from live human ears with implications for wideband acoustic immittance measurements, J. Acoust. Soc. Am. 148 (2020) 3042, https://doi.org/10.1121/10.0002358.

[24] H. Motallebzadeh, S. Puria, Stimulus-frequency otoacoustic emissions and middle-ear pressure gains in a finite-element mouse model, J. Acoust. Soc. Am. 152 (2022) 2769–2780, https://doi.org/10.1121/10.0014901.

[25] M.R. Stinson, B.W. Lawton, Specification of the geometry of the human ear canal for the prediction of sound-pressure level distribution, J. Acoust. Soc. Am. 85 (1989) 2492–2503, https://doi.org/10.1121/1.397744.

[26] S.E. Voss, N.J. Horton, R.R. Woodbury, K.N. Sheffield, Sources of variability in reflectance measurements on normal cadaver ears, Ear Hear. 29 (2008) 651–665, https://doi.org/10.1097/AUD.0b013e318174f07c.

[27] M.R. Stinson, E.A. Shaw, B.W. Lawton, Estimation of acoustical energy reflectance at the eardrum from measurements of pressure distribution in the human ear canal, J. Acoust. Soc. Am. 72 (1982) 766–773, https://doi.org/10.1121/1.388257.

[28] S. Ihrle, M. Lauxmann, A. Eiber, P. Eberhard, Nonlinear modelling of the middle ear as an elastic multibody system — applying model order reduction to acousto-structural coupled systems, JCAM 246 (2013) 18–26, https://doi.org/10.1016/j.cam.2012.07.010.

[29] B. Sackmann, B. Warnholtz, J.H. Sim, D. Burovikhin, E. Dalhoff, P. Eberhard, M. Lauxmann, Investigation of tympanic membrane influences on middle-ear impedance measurements and simulations, in: A. Kecskeméthy, F. Geu Flores (Eds.), Multibody Dynamics 2019, Springer International Publishing, Cham, 2020, pp. 3–10.

[30] G.R. Merchant, S.N. Merchant, J.J. Rosowski, H.H. Nakajima, Controlled exploration of the effects of conductive hearing loss on wideband acoustic immittance in human cadaveric preparations, Hear. Res. 341 (2016) 19–30, https://doi.org/10.1016/j.heares.2016.07.018.

[31] S. Marelli, B. Sudret, UQLab: A framework for uncertainty quantification in MATLAB, Proc. 2nd Int. Conf. on Vulnerability, Risk Analysis and Management (ICVRAM2014), July 13-16, (2014,). Liverpool, United Kingdom, pp. 2554-2563, DOI: 10.1061/9780784413609.257.

[32] J.J. Rosowski, W. Chien, M.E. Ravicz, S.N. Merchant, Testing a method for quantifying the output of implantable middle ear hearing devices, AUD 12 (2007) 265–276, https://doi.org/10.1159/000101474.

[33] T.M. Essinger, M. Koch, H. Maier, J.H. Sim, N. Greene, L.J. Ren, M. Bornitz, M. Neudert, Unified database and methods for METF validation in temporal bone experiments, in: 9th International Symposium on Middle Ear Mechanics in Research and Otology, 2022, p. 17.

[34] J.J. Rosowski, H.H. Nakajima, M.A. Hamade, L. Mahfoud, G.R. Merchant, C.F. Halpin, S.N. Merchant, Ear-canal reflectance, umbo velocity, and tympanometry in normal-hearing adults, Ear Hear. 33 (2012) 19–34, https://doi.org/10.1097/AUD.0b013e31822ccb76.

[35] K.M. Nørgaard, E. Fernandez-Grande, C. Schmuck, S. Laugesen, Reproducing ear-canal reflectance using two measurement techniques in adult ears, J. Acoust. Soc. Am. 147 (2020) 2334–2344, https://doi.org/10.1121/10.0001094.

[36] Y. Lecun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proc. IEEE 86 (1998) 2278–2324, https://doi.org/10.1109/5.726791.

[37] M. Raghu, C. Zhang, J. Kleinberg, S. Bengio, Transfusion: Understanding Transfer Learning for Medical Imaging, thirty-second ed., 2019.

[38] C.M. Bishop, Pattern Recognition and Machine Learning, Reprint of the Original Firstst Ed. Twentieth0sixth, Springer, New York, 2016.

[39] A.A. Taha, A. Hanbury, Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool, BMC Med. Imaging 15 (2015) 29, https://doi.org/10.1186/s12880-015-0068-x.

[40] M.P. Feeney, D.H. Keefe, L.L. Hunter, D.F. Fitzpatrick, A.C. Garinis, D.B. Putterman, G.P. McMillan, Normative wideband reflectance, equivalent admittance at the tympanic membrane, and acoustic stapedius reflex threshold in adults, Ear Hear. 38 (2017) e142–e160, https://doi.org/10.1097/AUD.0000000000000399.

[41] R.S. Hong, C.M. Metz, D.I. Bojrab, S.C. Babu, J. Zappia, E.W. Sargent, E.Y. Chan, I.C. Naumann, M.J. LaRouere, Acoustic reflex screening of conductive hearing loss for third window disorders, Otolaryngol. Head Neck Surg. 154 (2016) 343–348, https://doi.org/10.1177/0194599815620162.

[42] T. Kan, H. Ueda, M. Kishimoto, Y. Tsuchiya, T. Ogawa, Y. Uchida, Availability of audiological evaluation for the differential diagnosis of clinical otosclerosis, Auris Nasus Larynx 47 (2020) 343–347, https://doi.org/10.1016/j.anl.2020.03.009.

[43] D.H. Keefe, K.L. Archer, K.K. Schmid, D.F. Fitzpatrick, M.P. Feeney, L.L. Hunter, Identifying otosclerosis with aural acoustical tests of absorbance, group delay, acoustic reflex threshold, and otoacoustic emissions, J. Am. Acad. Audiol. 28 (2017) 838–860, https://doi.org/10.3766/jaaa.16172.

[44] D. Abur, N.J. Horton, S.E. Voss, Intrasubject variability in power reflectance, J. Am. Acad. Audiol. 25 (2014) 441–448, https://doi.org/10.3766/jaaa.25.5.3.

[45] K.A. Groon, D.M. Rasetshwane, J.G. Kopun, M.P. Gorga, S.T. Neely, Air-leak effects on ear-canal acoustic absorbance, Ear Hear. 36 (2015) 155–163, https://doi.org/10.1097/AUD.0000000000000077.

[46] J.D. Lewis, The area discontinuity between probe and ear canal as a source of power-reflectance measurement-location variability, J. Acoust. Soc. Am. 143 (2018) 1106, https://doi.org/10.1121/1.5024360.

[47] G.R. Merchant, K.M. Schulz, J.N. Patterson, D. Fitzpatrick, K.L. Janky, Effect of cochlear implantation on vestibular evoked myogenic potentials and wideband acoustic immittance, Ear Hear. 41 (2020) 1111–1124, https://doi.org/10.1097/AUD.0000000000000831.

[48] M.D. Shaver, X.-M. Sun, Wideband energy reflectance measurements: effects of negative middle ear pressure and application of a pressure compensation procedure, J. Acoust. Soc. Am. 134 (2013) 332–341, https://doi.org/10.1121/1.4807509.

[49] X.-M. Sun, Wideband acoustic immittance: normative study and test-retest reliability of tympanometric measurements in adults, J. Speech Lang. Hear. Res. 59 (2016) 819–834, https://doi.org/10.1044/2016_JSLHR-H-14-0322.

[50] S.E. Voss, M.F. Adegoke, N.J. Horton, K.N. Sheth, J. Rosand, C.A. Shera, Posture systematically alters ear-canal reflectance and DPOAE properties, Hear. Res. 263 (2010) 43–51, https://doi.org/10.1016/j.heares.2010.03.003.

[51] Y.-W. Liu, C.A. Sanford, J.C. Ellison, D.F. Fitzpatrick, M.P. Gorga, D.H. Keefe, Wideband absorbance tympanometry using pressure sweeps: system development and results on adults with normal hearing, J. Acoust. Soc. Am. 124 (2008) 3708–3719, https://doi.org/10.1121/1.3001712.

[52] L.A. Werner, E.C. Levi, D.H. Keefe, Ear-canal wideband acoustic transfer functions of adults and two- to nine-month-old infants, Ear Hear. 31 (2010) 587–598, https://doi.org/10.1097/AUD.0b013e3181e0381d.

[53] G.R. Merchant, J.H. Siegel, S.T. Neely, J.J. Rosowski, H.H. Nakajima, Effect of middle-ear pathology on high-frequency ear canal reflectance measurements in the frequency and time domains, JARO 20 (2019) 529–552, https://doi.org/10.1007/s10162-019-00735-1.

[54] M.P. Feeney, I.L. Grant, D.M. Mills, Wideband energy reflectance measurements of ossicular chain discontinuity and repair in human temporal bone, Ear Hear. 30 (2009) 391–400, https://doi.org/10.1097/AUD.0b013e3181a283ed.

[55] K. Gladiné, J.J.J. Dirckx, Average middle ear frequency response curves with preservation of curve morphology characteristics, Hear. Res. 363 (2018) 39–48, https://doi.org/10.1016/j.heares.2018.02.005.

[56] D.H. Keefe, J.L. Simmons, Energy transmittance predicts conductive hearing loss in older children and adults, J. Anthropol. Soc. Oxf. 114 (2003) 3217–3238, https://doi.org/10.1121/1.1625931.

[57] H. Tang, P. Razavi, K. Pooladvand, P. Psota, N. Maftoon, J.J. Rosowski, C. Furlong, J.T. Cheng, High-speed holographic shape and full-field displacement measurements of the tympanic membrane in normal and experimentally simulated pathological ears, Appl. Sci. 9 (2019) 2809, https://doi.org/10.3390/app9142809.

[58] K.R. Nørgaard, E. Fernandez-Grande, S. Laugesen, Compensating for evanescent modes and estimating characteristic impedance in waveguide acoustic impedance measurements, J. Acoust. Soc. Am. 142 (2017) 3497, https://doi.org/10.1121/1.5016808.

[59] K.R. Nørgaard, E. Fernandez-Grande, S. Laugesen, Compensating for oblique ear-probe insertions in ear-canal reflectance measurements, J. Acoust. Soc. Am. 145 (2019) 3499, https://doi.org/10.1121/1.5111340.

[60] W. Chen, C.A. Campbell, G.E. Green, K. van den Bogaert, C. Komodikis, L.S. Manolidis, E. Aconomou, Y. Kyamides, K. Christodoulou, C. Faghel, C.M. Giguére, R. L. Alford, S. Manolidis, G. van Camp, R.J.H. Smith, Linkage of otosclerosis to a third locus (OTSC3) on human chromosome 6p21.3-22.3, J. Med. Genet. 39 (2002) 473–477, https://doi.org/10.1136/jmg.39.7.473.

[61] M. Thys, G. van Camp, Genetics of Otosclerosis, vol. 30, Otology & neurotology official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology, 2009, pp. 1021–1032, https://doi.org/10.1097/MAO.0b013e3181a86509.

[62] S. Hannula, R. Bloigu, K. Majamaa, M. Sorri, E. Mäki-Torkko, Ear diseases and other risk factors for hearing impairment among adults: an epidemiological study, Int. J. Audiol. 51 (2012) 833–840, https://doi.org/10.3109/14992027.2012.707334.