Original Investigation

Deep Learning-Enhanced Ultra-high-resolution CT Imaging for Superior Temporal Bone Visualization

Lavinia Brockstedt, MD, Nils F. Grauhan, MD, Andrea Kronfeld, PhD, Mario Alberto Abello Mercado, MD, Julia Döge, MD, Antoine Sanner, MSc, Marc A. Brockmann, MD, Ahmed E. Othman, MD

Rationale and Objectives: This study assesses the image quality of temporal bone ultra-high-resolution (UHR) Computed tomography (CT) scans in adults and children using hybrid iterative reconstruction (HIR) and a novel, vendor-specific deep learning-based reconstruction (DLR) algorithm called AiCE Inner Ear.

Material and Methods: In a retrospective, single-center study (February 1–July 30, 2023), UHR-CT scans of 57 temporal bones of 35 patients (5 children, 23 male) with at least one anatomical unremarkable temporal bone were included. There is an adult computed tomography dose index volume (CTDIvol 25.6 mGy) and a pediatric protocol (15.3 mGy). Images were reconstructed using HIR at normal resolution (0.5-mm slice thickness, 512^2 matrix) and UHR (0.25-mm, 1024^2 and 2048^2 matrix) as well as with a vendor-specific DLR advanced intelligent clear-IQ engine inner ear (AiCE Inner Ear) at UHR (0.25-mm, 1024^2 matrix). Three radiologists evaluated 18 anatomic structures using a 5-point Likert scale. Signal-to-noise (SNR) and contrast-to-noise ratio (CNR) were measured automatically.

Results: In the adult protocol subgroup ($n = 30$; median age: 51 [11–89]; 19 men) and the pediatric protocol subgroup ($n = 5$; median age: 2 [1–3]; 4 men), UHR-CT with DLR significantly improved subjective image quality ($p < 0.024$), reduced noise ($p < 0.001$), and increased CNR and SNR ($p < 0.001$). DLR also enhanced visualization of key structures, including the tendon of the stapedius muscle ($p < 0.001$), tympanic membrane ($p < 0.009$), and basal aspect of the osseous spiral lamina ($p < 0.018$).

Conclusion: Vendor-specific DLR-enhanced UHR-CT significantly improves temporal bone image quality and diagnostic performance.

Key Words: Computed tomography; Deep learning; Temporal bone; Image quality; Ultra-high resolution.

Abbreviations: AiCE (Advanced Intelligent Clear-IQ Engine), **AiCE Inner Ear** (Advanced Intelligent Clear-IQ Engine Inner Ear), **AIDR 3D** (Adaptive Iterative Dose Reconstruction), **CNR** (contrast to noise ratio), **ConvCT** (conventional multi-detector row computed tomography), **CT** (computed tomography), **DLP** (dose-length products), **DLR** (deep learning- based image reconstruction), **DCNN** (deep convolutional neural network), **CTDIvol** (computed tomography dose index volume), **ICC** (Intraclass Correlation Coefficient), **FBP** (Filtered back projection), **FoV** (field of view), **HIR** (hybrid iterative image reconstruction), **HR-CT** (high-resolution computed tomography), **IR** (iterative image reconstruction), **MBIR** (model-based iterative reconstruction), **MDCT** (multi-detector row computed tomography), **NR** (normal resolution), **SNR** (signal to noise ratio), **STARD** (Standards for Reporting of Diagnostic Accuracy), **UHR** (ultra-high resolution), **UHR 1024** (ultra-high resolution 1024×1024 matrix), **UHR 2048** (ultra-high resolution 2048×2048 matrix), **UHR-CT** (ultra-high resolution computed tomography)

INTRODUCTION

Computed tomography (CT) imaging is frequently used in patients with hearing disorders or vertigo, as it offers insights into the intricate anatomy of the middle and inner ear structures. The small size, three-dimensional configuration, and the complex interconnections of bony structures pose unique challenges for imaging modalities, necessitating advanced technology to achieve accurate visualization. Accurate imaging of these structures is critical for diagnosing pathologies such as otosclerosis, cholesteatoma, and congenital malformations, as well as for planning surgical interventions like cochlear

Acad Radiol 2025; 32:3618–3630

From the Department of Neuroradiology, University Medical Centre Mainz, Johannes Gutenberg University Mainz, Mainz, Germany (L.B., N.F.G., A.K., M.A.A.M., A.S., M.A.B., A.E.O.); Department of ear, nose, and throat medicine, University Medical Centre Mainz, Johannes Gutenberg University Mainz, Mainz, Germany (J.D.). Received November 28, 2024; revised January 23, 2025; accepted February 3, 2025. **Address correspondence to:** A.O. e-mail: ahmed.e.othman@googlemail.com

https://doi.org/10.1016/j.acra.2025.02.002

implantation or ossiculoplasty. Multi-detector row CT (MDCT) enabled the identification of subtle anatomical structures such as the auditory ossicles, which were challenging to detect in earlier single-detector row scans (1). Alongside hardware improvements, image reconstruction algorithms evolved rapidly. Filtered back projection (FBP) was one of the first methods employed to allow, fast image reconstruction but was dependent on high radiation doses to allow significant image quality. Iterative reconstruction (IR) has been used since the early 2010s as it demanded increased computational capacities (2). IR methods enabled radiologists to reduce radiation dose and improve image quality by lowering image noise in temporal bone imaging (3,4). In the last decade, machine learning has entered the field of health care and revolutionized medical imaging (5,6). In medical image reconstruction, deep learning has already been applied to improve image quality by increasing signal-to-noise (SNR) and contrast-to-noise ratio (CNR). DLR algorithms have been successfully employed on various body parts (10–20).

Not only has the reconstruction software improved over time, but also the hardware of the CT scanners. A distinction can be made between high-resolution (HR) and ultra-high-resolution (UHR) CT scanners. While UHR-CT has a precise definition, HR-CT lacks a universally accepted standard. HR-CT typically refers to thin slices, usually ranging from 0.625 to 1.25 mm, which are acquired and reconstructed using sharp algorithms (7). UHR-CT is defined by detector elements measuring 0.25 mm or smaller along the z-axis, resulting in a slice thickness of 0.25 mm or less (8). This allows for detailed visualization of micro-anatomical structures, making it particularly useful for imaging small and intricate anatomical details. (2–4).

Recent studies have shown that DLR improves the image quality of the temporal bone HR-CT (9,10). Imaging on UHR-CT is known to have higher noise and lower CNR and SNR (17).

This study therefore aimed to assess the image quality and visibility of anatomic microstructures in the temporal bone using UHR-CT with vendor-specific deep learning reconstruction (Advanced intelligent Clear-IQ Engine (AiCE) Inner Ear, Canon Medical Systems, Otawara, Tochigi, Japan) and compare it to hybrid iterative image reconstruction (AIDR 3D) in an adult and pediatric subgroup.

MATERIAL AND METHODS

The Ethics Committee of the Rhineland-Palatinate Chamber of Physicians gave its approval for this retrospective, single-center study (approval number: 2021–15948_1-retrospektv, approval number: 06.07.2021). The 2013 revision of the Helsinki Declaration was followed in conducting the study.

Study Sample

Between February 1st, 2023 and July 30th, 2023, a retrospective analysis included a total 52 patients (43 adults and 9 children; 20 female; resulting in 104 temporal bones) who underwent UHR-CT of the temporal bone. The primary indications for imaging were unclear hearing loss, progressive hearing impairment and/or preoperative evaluation prior to cochlear implant. As the aim of the underlying work was the evaluation of image quality by analyzing relatively small anatomical structures, only volume UHR-CT scans with at least one anatomical unremarkable temporal bone were included. This results in a final cohort of 35 patients. The STARD flowchart in Figure 1 illustrates the inclusion und exclusion criteria for this study.

Two distinct CT scanning protocols are employed based on the patient's age, resulting in a pediatric and an adult protocol subgroup. Table 1 provides an overview of patient demographics and the indications for CT scans within these subgroups. In to the absence of standardized diagnostic reference values for the radiation dose in temporal bone scans, it is the responsibility of individual clinical departments to establish protocols that balance sufficient diagnostic quality and radiation.

The scanning protocols are align with our standard clinical practices, where three age groups are differentiated: children under four years, children aged four to ten years, and children older than ten years, with the latter being scanned using the adult protocol.

Image Acquisition and Reconstruction

All enrolled patients underwent imaging of the temporal bone using a non-contrast volume UHR-CT scan on the Aquilion Precision CT scanner of Canon Medical Systems (Otawara/Tochigi/Japan). The scanner features a focal spot size of 0.4 mm × 0.5 mm, a detector element size of 0.25 mm × 0.25 mm, and beam collimation of 0.25 mm × 160 rows with 1792 channels. The scan range of a volume CT scan is determined by the beam collimation resulting in a scan width of 40 mm (160 × 0.25 mm). The scanning protocol includes a tube voltage of 120 kV, with tube currents set at 250 mA for the adult protocol and 220 mA for the pediatric protocol. Rotation times are 0.75 s for the adult protocol and 0.5 s for the pediatric protocol. A field of view (FoV) of 200 mm and a scan range of 4 cm are maintained to capture detailed volume images. The computed tomography dose index volume (CTDIvol) measures 25.6 mGy (DLP 102.4 mGy·cm) for the adult protocol and 15.30 mGy (DLP 61.20 mGy·cm) for the pediatric protocol. In the pediatric protocol group, all patients underwent short sedation (anesthesia) during the CT scan to minimize motion artifacts and ensure optimal image quality.

The two primary reconstruction methods utilized are HIR, using AIDR 3D (mild level) with an FC81 kernel, and a novel delicated DLR especially developed for temporal bone imaging AiCE Inner Ear (mild level), Canon Medical Systems, Otawara, Tochigi, Japan).

Deep Learning-Based Image Reconstruction for Temporal Bone Imaging: AiCE Inner Ear

Using high-quality, sophisticated MBIR target images as training material, the AiCE Inner Ear algorithm learns to convert low-quality input data into crisp, low-noise images.

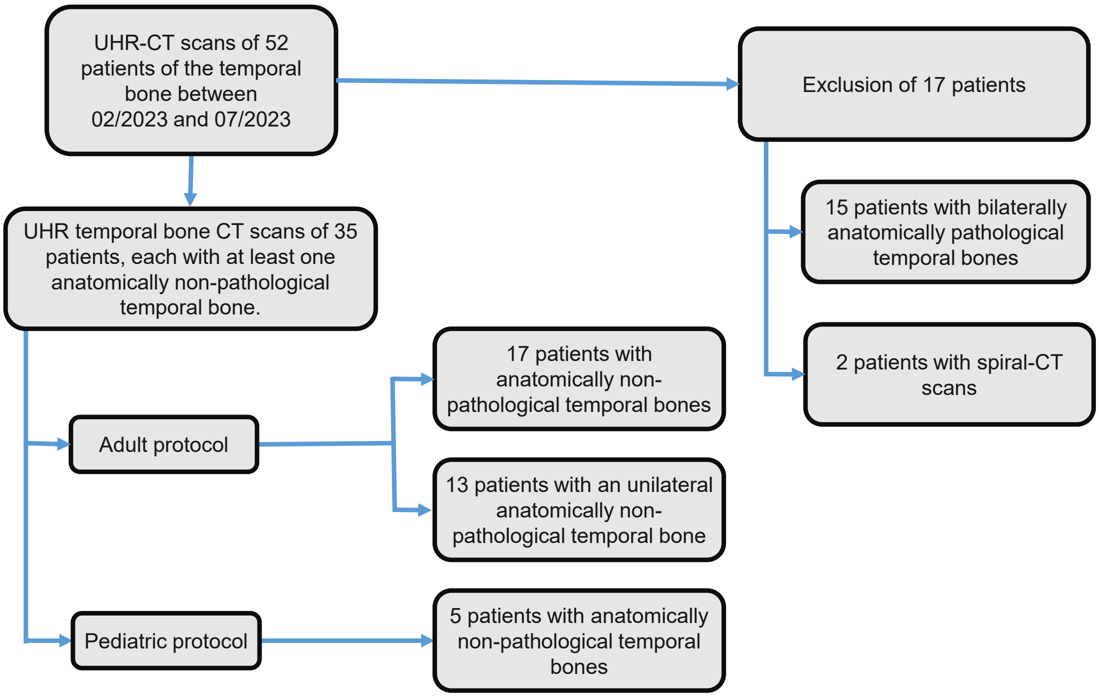

Figure 1. STARD Flowchart of the inclusion/exclusion process. UHR CT scans of anatomical, non-pathological temporal bones from 35 patients were included, resulting in a total of 57 temporal bones.

TABLE 1. Patients Characteristics

Parameter	Adult protocol subgroup	Pediatric protocol subgroup
Mean age in years (range)	51 (11–89)	2 (1–3)
Gender	11 female; 19 male	1 female; 4 male
Indication for CT		
preoperative assessment before cochlear implantation in cases of severe hearing loss bordering on deafness	20	5
Assessment of electrode and implant positions of cochlear implants	3	
diagnosis and management of cholesteatoma	2	
Evaluation of acute hearing loss	1	
hearing impairment and pulse-synchronous ear noise	1	
sound conduction disorders in auditory canal atresia	1	
Assessment of chronic otitis media	1	
suspected perilymphatic fistula following surgery	1	

Thousands of phantom and patient images were included in the development. In order to effectively reconstruct pictures, AiCE Inner Ear functions in both the raw and image domains in the clinic. The AiCE Inner Ear algorithm leverages high-quality reference images, reconstructed with advanced model-based iterative reconstruction (MBIR), as its training material. The training process involves comparing the output of the deep convolutional neural network (DCNN) to the reference images and adjusting the network's weights iteratively through backpropagation. This optimization ensures that the algorithm minimizes discrepancies between its output and the reference images. The training dataset includes millions of image pairs, ensuring robustness across various clinical scenarios. To enhance its utility in low-dose imaging, the algorithm was trained with low-quality data to enable it to generate high-quality images while maintaining signal integrity and spatial resolution.

The reconstruction parameters are categorized into normal resolution (NR), ultra-high resolution 1024^2 matrix (UHR 1024), ultra-high resolution 2048^2 matrix (UHR 2048), and DLR UHR. NR employs a 512^2 matrix and DLR UHR is reconstructed with a 1024^2 matrix. All data are reconstructed with a FoV of 60 mm. For NR, a slice thickness of 0.5 mm, a slice increment of 0.5 mm, and 80 reconstructed slices are used. For UHR 1024 and UHR 2048 a slice thickness of 0.25 mm, a slice increment of 0.125 mm and 320 reconstructed slices are employed.

Subjective Image Analyses

Subjective image analyses were conducted by three radiologists: one certified neuroradiologists with eight years of experience in head and neck imaging and two radiologist with five years of experience in head and neck imaging (M.A.A.M., N.F.G. and L.A.B.).

To minimize bias and ensure objective assessment, the four image reconstructions were randomized and the readers were blinded during the evaluation process. Additionally, a washout period of more than six weeks was implemented between the two readings to further reduce potential carryover effects.

The readers evaluate the acquired thin-slice axial images with a slice thickness of 0.25 mm (matrix size 1024^2 and 2048^2) or 0.5 mm (matrix size 512^2). They were able to reconstruct the images in the MPR (multi-planar reformation) mode as needed, enabling optimal visualization of the small anatomical structures of the temporal bone. To ensure standardized image reading, all three readers received a briefing on the evaluation criteria and the use of a 5-point Likert scale prior to image evaluation. The subjective parameters included image noise and sharpness, the extent of artifacts, and the diagnostic confidence of the reader during image analysis. Additionally, the assessability of 18 predefined micro-anatomical structures of the middle and inner ear was rated by each of the three readers using the 5-point Likert scale (Table 2).

The selection of the analyzed micro-anatomical structures was guided by their clinical relevance in diagnosing and managing various otologic conditions, as well as their importance in preoperative planning for interventions such as cochlear implantation and ossicular prosthesis placement. Furthermore, several structures were chosen specifically for their fine anatomical detail or small size, making them ideal for exploring the diagnostic boundaries of UHR-CT.

The tympanic membrane, for example, is critical for sound transmission and typically only becomes visible in conventional multi-detector row computed tomography (ConvCT) imaging when thickened or affected by pathology. Similarly, the ossicular chain structures, such as the stapes and the incudomalleolar and incudostapedial joints, are crucial for sound conduction but challenging to evaluate in standard CT due to their minute size. Other structures, such as the tendons of the tensor tympani and stapedius muscles, as well as the chorda tympani, present particular challenges for conventional imaging. Assessment of the bony coverage of the facial canal is paramount in preoperative planning for cochlear implants to avoid damaging the facial nerve or causing unintended stimulation. Additionally, the tegmen tympani is vital to evaluate for potential defects that could lead to cerebrospinal fluid leaks. Other structures, such as the singular canal, the osseous spiral lamina, and the modiolus within the cochlea, offer a unique challenge due to their exceptionally small and thin features, often partially or entirely invisible on ConvCT. Finally, structures such as the cochlear and vestibular aqueducts, as well as the petromastoid canal, hold significance due to their roles in inner ear physiology and their potential to present abnormalities that may impact hearing. Enlarged vestibular aqueducts, for instance, are associated with sensorineural hearing loss, emphasizing the importance of their accurate imaging. By including these challenging yet clinically significant structures, this study aims to thoroughly evaluate the capabilities and limitations of UHR-CT in visualizing critical temporal bone anatomy. For the purpose of image evaluation, the Sectra software (Sectra Workstation; Linköping, Sweden) was employed, harnessing its capabilities including the ability to manually adjust slice thickness and perform multi-planar reconstructions.

For the assessment of overall image quality, the adapted European guidelines on quality criteria for CT imaging (11,12) were employed. In order to rate the visibility of very small anatomical structures, a 5-point Likert scale originally developed for rating facial nerve canal dehiscence (13) was applied (Table 2). Consistent use of the 5-point Likert scale was maintained across all patients and categories.

Objective Image Analyses

For an observer-independent assessment of image quality, images were exported from Sectra (Sectra Workstation; Linköping, Sweden) in DICOM format to a workstation equipped with

TABLE 2. Adapted 5-point Likert Scale for Evaluation of the Subjective Overallimage Quality and the Visibility of Micro-anatomical Structures

Grading	Overall image quality, noise, sharpness, and diagnostic precision	General Artifacts	Visibility of micro-anatomical Structures
5	outstanding/excellent	minimal, noticeably little (well below average)	clearly continuous delimitable
4	very good/above average	low (somewhat below average)	generally delimitable with punctiform blur or discontinuity
3	good/average	moderate (average)	generally delimitable with linear blur or discontinuity
2	poor/below average	marked (somewhat above average)	partially delimitable
1	very poor	pronounced (unacceptable)	not delimitable

MATLAB tools (2022b, The MatWorks). The following objective image parameters were evaluated: noise, signal-to-noise ratio (SNR), contrast-to-noise ratio (CNR).

A radiologist with 5 years of experience in head and neck imaging selected an exemplary 0.25 mm (DLR UHR 1024, HIR UHR 2048, HIR UHR 1024) or 0.5 mm (HIR NR) image focused on the epitympanic recess, the incus and malleus. In the chosen slice, a region of interest was measured in the air of the meatus externus and in the partially depicted cerebellar hemisphere. The regions of interest (ROIs) were circular in shape with a diameter of 1 cm. For each patient, the same location was used for all reconstructions to ensure consistency in the analysis.

Statistical Analysis

The SPSS (SPSS IBM Statistics for Windows, Version 23.0 IBM Corp) statistical analysis program was used. If regularly distributed, the mean and standard deviation are used to express all continuous variables. Variables for categorization were presented in absolute frequencies and ratios. A Wilcoxon signed-rank test was used to compare subjective image quality graded for each anatomic structure on the temporal bone. Intrer-rater correlation (Fleiss' kappa and intraclass correlation coefficient) was evaluated. The level of agreement was defined as follows: >0.75 as excellent, 0.6–0.75 as very good, 0.5–0.6 as good, 0.4–0.5 as fair, and <0.40 as poor (14). Only significant Fleiss' kappa values ($p<0.001$) are reported.

RESULTS

Patient Cohort

The final study sample consisted of 35 patients, encompassing both infants and adult patients with a gender distribution of 12 female and 23 males. The STARD flowchart in Figure 1 illustrates the inclusion und exclusion criteria for this study. The pediatric protocol group included five patients (10 temporal bones) with a median age of 2 years (rage: 1–3 years). The adult protocol group incorporating 30 patients (47 temporal bones) with a median age of 51 years (rage: 11–89 years).

The indications for CT imaging within the patient cohort varied, and they are delineated alongside the patients' characteristics in Table 1.

Subjective Image Quality

The analysis of both adult and pediatric protocol subgroups demonstrated the clear superiority of DLR over HIR. DLR provided superior image quality, reduced noise, and enhanced visibility of numerous anatomical landmarks.

The following sections provide a detailed breakdown of the findings within the adult and pediatric subgroups.

Adult Protocol Subgroup

Within the adult protocol subgroup, the analysis of general subjective image parameters revealed that a DLR specifically trained for inner ear structures demonstrated unequivocal superiority over all HIR techniques ($p<0.001$) (Fig 2, Table 4). Specifically, DLR facilitated significant clearer identification of almost all selected anatomical landmarks (Fig 4, Fig 5), with the cochlear aqueduct being the sole exception, showing no significant difference.

Within the group of hybrid iterative reconstructions, HIR UHR 2048 exhibited significantly higher image noise and artifacts. Nonetheless, some anatomical landmarks, such as the stapes crus anterior ($p<0.010$), tendon of the stapedius muscle ($p<0.021$), chorda tympani in the tympanic segment ($p<0.010$) and posterior canaliculus ($p<0.018$) as well as tegmen tympani ($p<0.003$), were delineated better with HIR UHR 2048 compared to HIR UHR 1024. Furthermore, 12 of the 18 anatomical landmarks were clearer to identify with HIR UHR 1024 compared to HIR NR. However, no significant difference in identification was observed for the tympanic membrane, posterior limb of the stapes, incudomalleolar joint, the apical aspect of the osseous spiral lamina, modiolus of the cochlea, and the cochlear aqueduct.

Table 3 presents the inter-rater agreement assessed using Fleiss' Kappa. Inter-rater reliability between the three readers ranged from 0.5 to 0.8 with a mean κ value of 0.7. Notably, the lowest inter-rater agreement was observed concerning the identification of the incudostapedial joint with DLR. The Intraclass Correlation Coefficient (ICC) analysis is available as supplementary data (S1), offering deeper insights into the reliability of the ratings and providing a more comprehensive assessment of the agreement among the raters.

Pediatric Protocol Subgroup

In the pediatric protocol subgroup, the assessment of general subjective image parameters unveiled DLR's superior performance in comparison to HIR UHR 2048 (Fig 3). Notably, DLR yielded higher general image quality ($p<0.024$), reduced image noise ($p<0.006$), and fewer artifacts ($p<0.006$) (Table 5). However, no significant disparities were observed in terms of diagnostic confidence or general image sharpness. Moreover, DLR demonstrated superior performance across all general subjective image parameters when juxtaposed with both HIR UHR 1024 ($p<0.027$) and HIR NR ($p<0.013$) techniques.

DLR enhances the visibility of the selected anatomic landmarks in comparison to HIR UHR 2048. These included the tympanic membrane ($p<0.021$), tendon of the stapedius muscle ($p<0.009$), tendon of the tensor tympanic muscle ($p<0.024$) and basal aspect of the osseous spiral lamina ($p<0.018$). Additionally, compared to HIR UHR 1024, DLR improved the visibility of the crus anterior of the stapes ($p<0035$), the vestibular aqueduct ($p<0.036$) and the posterior canaliculus of the chorda tympani ($p<0.036$). Furthermore, in comparison to HIR NR, DLR improved the visibility of the crus posterior of the stapes ($p<0.024$), the tympanic segment of the chorda tympani ($p<0.028$), the

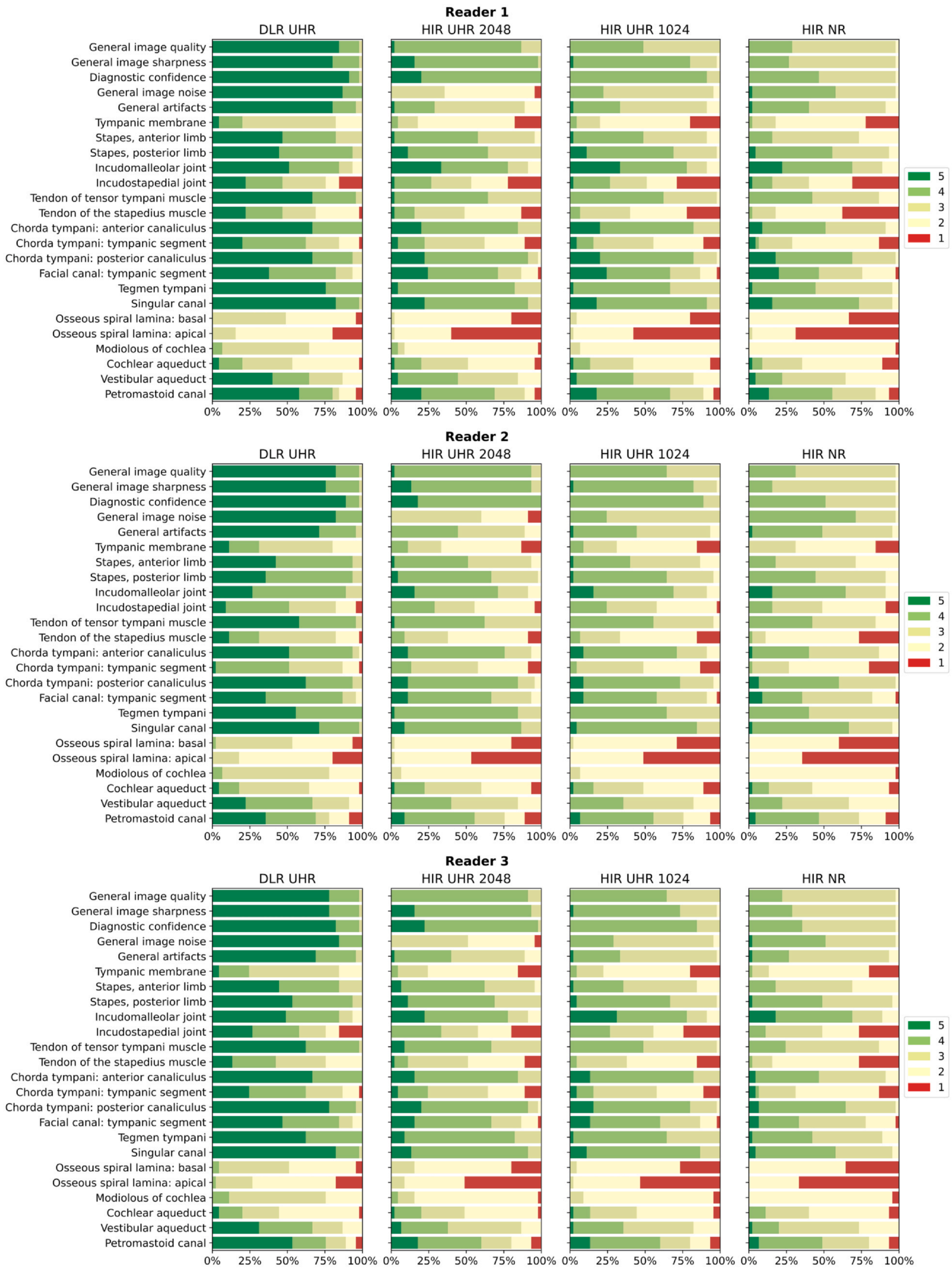

Figure 2. Subjective image quality in the adult protocol subgroup. Three radiologists assessed image quality of 18 anatomic structures using a 5-point Likert scale (1 = very poor, 2 = poor, 3 = good/average, 4 = very good/above average, 5 = outstanding/excellent; further Likert scale details are provided in Table 4).

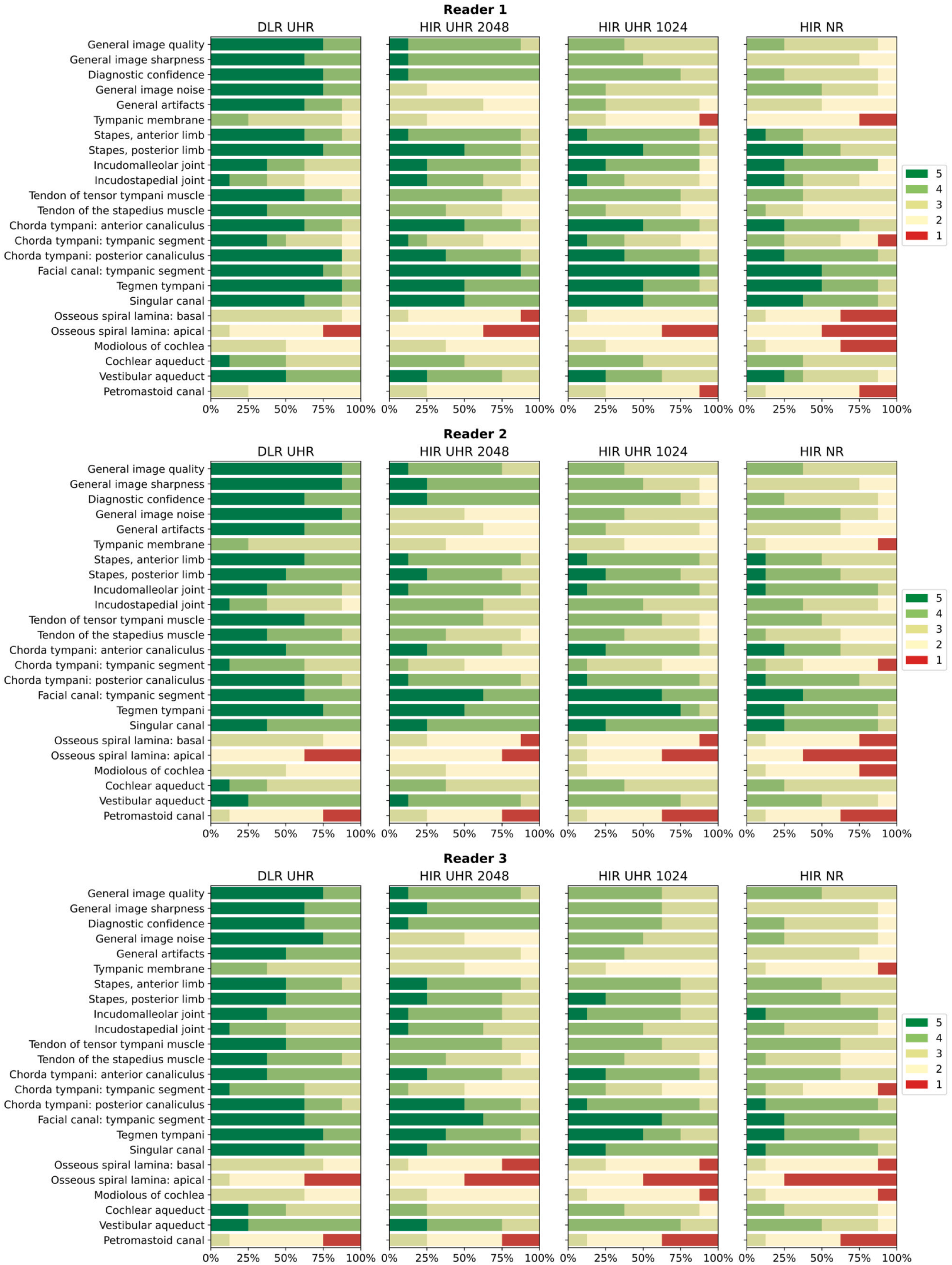

Figure 3. Subjective image quality in the pediatric protocol subgroup. Three radiologists assessed image quality of 18 anatomic structures using a 5-point Likert scale (1 = very poor, 2 = poor, 3 = good/average, 4 = very good/above average, 5 = outstanding/excellent; further Likert scale details are provided in Table 4).

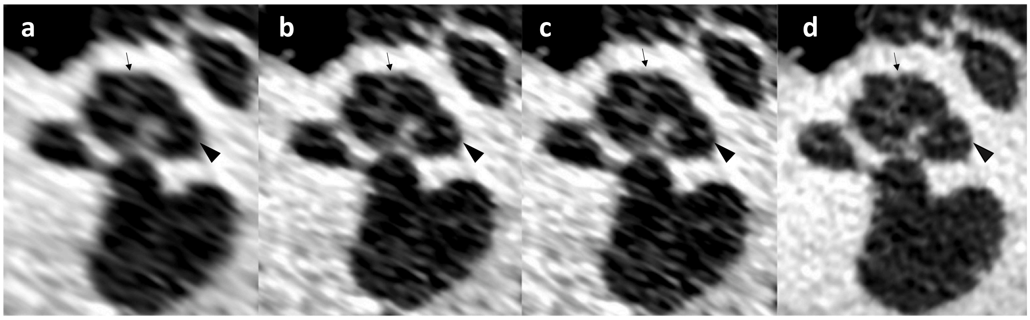

Figure 4. UHR-CT volume scan. Modiolus of the cochlea (arrow). Osseous spiral lamina (arrowhead). **(a)** HIR NR (AIDR 3D, 512² matrix). **(b)** HIR UHR 1024 (AIDR 3D, 1024² matrix). **(c)** HIR UHR 2048 (AIDR 3D, 2048² matrix). **(d)** DLR UHR 1024 (AiCE Inner Ear, 1024² matrix).

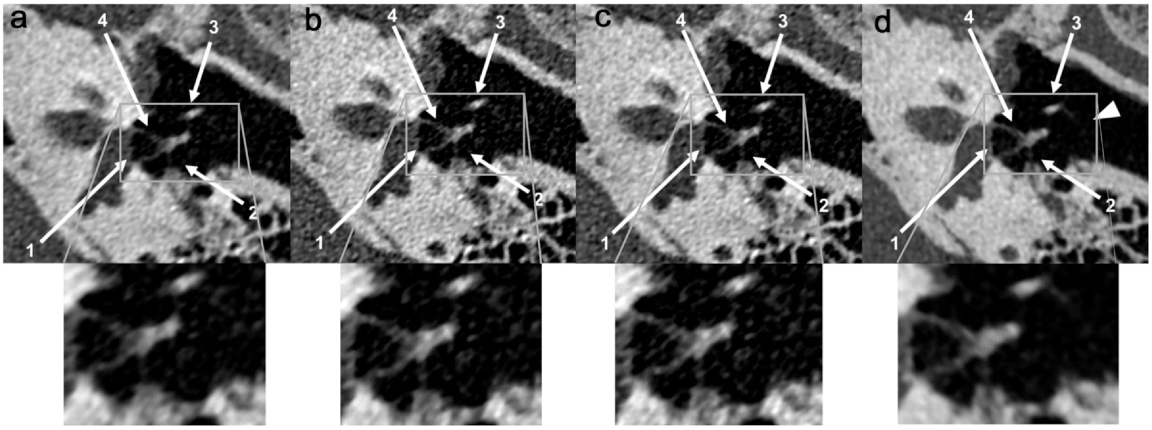

Figure 5. UHR-CT volume scan. Oval window (arrow 1). Pyramidal eminence and tendon of the stapedius muscle (arrow 2). Handle of the malleus (arrow 3). Stapes with anterior and posterior crus (arrow 4). Tympanic membrane (arrowhead). **(a)** HIR NR (AIDR 3D, 512² matrix). **(b)** HIR UHR 1024 (AIDR 3D, 1024² matrix). **(c)** HIR UHR 2048 (AIDR 3D, 2048² matrix). **(d)** DLR UHR 1024 (AiCE, 1024² matrix).

tegmen tympani ($p < 0.036$) and the modiolus of the cochlea ($p < 0.044$).

Table 3 presents the inter-rater agreement assessed using Fleiss' Kappa. Inter-rater reliability between the three readers ranged from fair to excellent, with κ values ranging from 0.4 to 0.8 with a mean κ value of 0.6. The lowest inter-rater agreement in the pediatric protocol subgroup was observed regarding the singular canal with DLR. The ICC analysis is available as supplementary data (S1) to provide additional insights into the reliability of the ratings and offers a more comprehensive evaluation of the agreement among the raters.

Objective Image Quality

DRL demonstrates significantly lower noise levels with a mean of 96 HU ± 17 HU ($p < 0.001$), while HIR UHR 2048 exhibits the highest noise at 236 HU ± 21 HU ($p < 0.001$), as shown in Figure 6. The contrast-to-noise ratio (CNR), calculated from bone, air, and noise signal intensities, is notably higher for DLR with a mean of 25 ± 7 ($p < 0.001$) and significantly lower for HIR UHR 2048 with a mean of 12 ± 3 ($p < 0.001$). Similarly, the signal-to-noise ratio (SNR), derived from bone signal intensity and noise, is significantly higher for DLR (mean: 16 ± 4; $p < 0.001$) and markedly lower for HIR UHR 2048 (mean: 7 ± 2; $p < 0.001$).

DISCUSSION

This study aimed to evaluate the subjective and objective image quality of temporal bone imaging using a UHR-CT scanner equipped with a vendor-specific DLR (AiCE Inner Ear, matrix: 1024^2). The performance of DLR was compared to the widely used HIR-based algorithm (AIDR 3D; Canon Medical Systems; FC81 kernel) across varying resolutions: UHR with a matrix size of 2048^2, UHR with a matrix size of 1024^2, and NR with a matrix size of 512^2). Depending on the patient's age, imaging was conducted using either an adult protocol (CTDIvol: 25 mGy) or a low-dose pediatric protocol (CTDIvol: 15 mGy). Notably, this is the first study to compare a vendor-specific DLR to HIR UHR 2048 Matrix while including pediatric patients.

TABLE 3. Fleiss Kappa for Image Quality for the Pediatric and Adult Protocol Subgroup

Assessed parameters	Fleiss' Kappa: adult protocol subgroup				Fleiss' Kappa: pediatric protocol subgroup			
	HIR NR	HIR UHR 1024	HIR UHR 2048	DLR UHR 1024	HIR NR	HIR UHR 1024	HIR UHR 2048	DLR UHR 1024
General image quality	0.620	0.517	0.517	0.517	0.517	0.664	0.634	0.747
General image sharpness	0.700	0.747	0.747	0.747	0.747	0.529	0.747	0.597
Diagnostic confidence	0.628	0.686	0.686	0.686	0.686	0.616	0.700	0.813
General image noise	0.597	0.584	0.584	0.584	0.584	0.644	0.486	0.747
General artifacts	0.632	0.644	0.644	0.644	0.644	0.678	0.597	0.678
Tympanic membrane	0.633	0.586	0.586	0.586	0.586	0.644	0.644	0.644
Stapes. anterior limb	0.714	0.702	0.702	0.702	0.702	0.586	0.817	0.538
Stapes. posterior limb	0.719	0.598	0.598	0.598	0.598	0.607	0.607	0.657
Incudomalleolar joint	0.578	0.639	0.639	0.639	0.639	0.592	0.667	0.598
Incudostapedial joint	0.624	0.553	0.553	0.553	0.553	0.564	0.598	0.522
Tendon of tensor tympani muscle	0.656	0.667	0.667	0.667	0.667	0.733	0.597	0.678
Tendon of the stapedius muscle	0.724	0.723	0.723	0.723	0.723	0.727	0.598	0.553
Chorda tympani: anterior canaliculus	0.694	0.591	0.591	0.591	0.591	0.713	0.607	0.535
Chorda tympani: tympanic segment	0.708	0.642	0.642	0.642	0.642	0.503	0.604	0.505
Chorda tympani: posterior canaliculus	0.595	0.667	0.667	0.667	0.667	0.664	0.569	0.634
Facial canal: tympanic segment	0.618	0.644	0.644	0.644	0.644	0.597	0.597	0.644
Tegmen tympani	0.706	0.591	0.591	0.591	0.591	0.561	0.690	0.747
Singular canal	0.618	0.686	0.686	0.686	0.686	0.625	0.625	0.373
Osseous spiral lamina: basal	0.776	0.686	0.686	0.686	0.686	0.586	0.667	0.747
Osseous spiral lamina: apical	0.769	0.644	0.644	0.644	0.644	0.686	0.644	0.538
Modiolous of cochlea	0.742	0.686	0.686	0.686	0.686	0.515	0.813	0.832
Cochlear aqueduct	0.717	0.644	0.644	0.644	0.644	0.686	0.644	0.719
Vestibular aqueduct	0.654	0.623	0.623	0.623	0.623	0.517	0.709	0.625
Petromastoid canal	0.582	0.569	0.569	0.569	0.569	0.579	0.707	0.500

Our findings demonstrate that temporal bone UHR-CT with DLR is superior to HIR UHR 2048 matrix in both, subjective and objective image quality for the adult and pediatric protocol group. Notably, in the adult protocol group, DLR enhanced the clarity of 17 out of 18 selected anatomical landmarks compared to both HIR UHR with superior SNR and CNR. Our findings indicate that a DLR explicitly trained for reconstruction of inner ear structures provides significantly improved image quality compared to HIR UHR 2048 in adults as well as in a pediatric population, underscoring its potential for widespread clinical application. DLR UHR 1024 allows excellent visualization of the anatomical structures of the temporal bone in the submillimeter range. The enhanced visualization of anatomical microstructures is particularly relevant for clinical diagnostics, as accurate imaging of these structures plays a crucial role in diagnosing pathologies such as otosclerosis, cholesteatoma, and congenital malformations. Furthermore, improved imaging is essential for planning complex surgical interventions, including cochlear implantation and ossiculoplasty.

Our findings are in line with the results of a technical report, showing that UHR-CT with DLR provides a better visibility of the chorda tympani nerve and the stapes in c-omparison to HIR UHR 1024 Matrix and a model-based reconstruction algorithm in different dose levels (15). Nagayama et al. highlight the potential of DLR to improve image quality at lower radiation doses for pediatric temporal bone imaging (16).

In HR-CT imaging, low-dose DLR (CTDI vol: 6.7 mGy, matrix size: 512^2, slice thickness 0.5 mm) also improves the image quality in comparison to filtered back projection (CTDI vol: 32.5 mGy, matrix size: 512^2, slice thickness 0.5 mm) with satisfactory clinical diagnosis making in patients with otitis media and mastoiditis (10). Another recently published study showed that DLR provides superior image quality compared to HIR and filtered back projection using HR-CT (9).

Our analysis of hybrid iterative reconstruction methods revealed a significant benefit of HIR UHR 2048 (matrix size: 2048^2; slice thickness: 0.25 mm) compared to the HIR UHR 1024 (matrix size: 1024^2; slice thickness: 0.25 mm) and HIR NR (matrix size: 512^2; slice thickness: 0.5 mm) in adult protocol group, even though HIR UHR 2048 suffers from higher image noise and artifacts. In the adult protocol subgroup, the subjective analysis indicates that HIR UHR 2048 provided a better definition for five out of 18 anatomic landmarks compared to HIR UHR 1024. In the pediatric

TABLE 4. Adult Protocol Subgroup: p-values of Subjective Image Analysis

Assessed parameters	DLR UHR 1024 vs. HIR UHR 2048			HIR UHR 2048 vs. HIR UHR 1024			HIR UHR 1024 vs. HIR NR		
Reader	1	2	3	1	2	3	1	2	3
General image quality	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	0.001	0.013	< 0.001	< 0.001
General image sharpness	< 0.001	< 0.001	< 0.001	0.001	0.001	< 0.001	< 0.001	< 0.001	< 0.001
Diagnostic security	< 0.001	< 0.001	< 0.001	0.001	0.001	0.001	< 0.001	< 0.001	< 0.001
General image noise	< 0.001	< 0.001	< 0.001	1.000	1.000	1.000	1.000	1.000	0.989
General artifacts	< 0.001	< 0.001	< 0.001	0.978	0.857	0.655	0.978	0.932	0.055
Tympanic membrane	< 0.001	< 0.001	< 0.001	0.681	0.074	0.117	0.074	0.036	0.018
Stapes. anterior limb	< 0.001	< 0.001	< 0.001	0.010	0.003	< 0.001	< 0.001	< 0.001	< 0.001
Stapes. posterior limb	< 0.001	< 0.001	< 0.001	0.725	0.117	0.073	0.001	< 0.001	0.002
Incudomalleolar joint	< 0.001	< 0.001	< 0.001	1.000	0.500	0.960	0.010	0.212	0.003
Incudostapedial joint	< 0.001	< 0.001	< 0.001	0.091	0.579	0.021	0.007	0.005	0.001
Tendon of tensor tympani muscle	< 0.001	< 0.001	< 0.001	0.074	0.010	< 0.001	< 0.001	0.001	< 0.001
Tendon of the stapedius muscle	< 0.001	< 0.001	< 0.001	0.003	0.021	0.001	< 0.001	< 0.001	< 0.001
Chorda tympani: anterior canaliculus	< 0.001	< 0.001	< 0.001	0.500	0.036	0.212	< 0.001	< 0.001	< 0.001
Chorda tympani: tympanic segment	< 0.001	< 0.001	< 0.001	0.010	0.001	0.005	< 0.001	< 0.001	< 0.001
Chorda tympani: posterior canaliculus	< 0.001	< 0.001	< 0.001	0.018	0.010	0.005	0.005	0.021	0.001
Facial canal: tympanic segment	< 0.001	< 0.001	< 0.001	0.173	0.013	0.036	< 0.001	< 0.001	< 0.001
Tegmen tympani	< 0.001	< 0.001	< 0.001	0.003	0.001	0.001	0.001	0.002	0.001
Singular canal	< 0.001	< 0.001	< 0.001	0.173	0.074	0.074	0.001	0.001	< 0.001
Osseous spiral lamina: basal	< 0.001	< 0.001	< 0.001	0.977	0.065	0.007	0.007	0.021	0.033
Osseous spiral lamina: apical	< 0.001	< 0.001	< 0.001	0.807	0.149	0.091	0.055	0.010	0.020
Modiolous of cochlea	< 0.001	< 0.001	< 0.001	0.286	1.000	0.033	0.036	0.036	0.091
Cochlear aqueduct	0.149	0.303	0.574	0.007	0.005	0.060	0.074	0.370	0.106
Vestibular aqueduct	< 0.001	< 0.001	< 0.001	0.173	0.074	0.076	< 0.001	< 0.001	0.002
Petromastoid canal	< 0.001	< 0.001	< 0.001	0.173	0.807	0.173	0.001	0.005	0.003

TABLE 5. Pediatric Protocol Subgroup: p-values of Subjective Image Analysis.

Assessed parameters	DLR UHR 1024 vs. HIR UHR 2048			HIR UHR 2048 vs. HIR UHR 1024			HIR UHR 2048 vs. HIR NR		
Reader	1	2	3	1	2	3	1	2	3
General image quality	0.024	0.015	0.010	0.044	0.102	0.172	0.293	0.579	0.383
General image sharpness	0.036	0.018	0.074	0.018	0.027	0.044	0.024	0.060	0.010
Diagnostic security	0.018	0.149	0.065	0.074	0.044	0.036	0.018	0.065	0.065
General image noise	0.006	0.006	0.006	0.996	0.993	0.991	0.715	0.715	0.172
General artifacts	0.006	0.005	0.006	0.986	0.986	0.972	0.018	0.065	0.044
Tympanic membrane	0.021	0.013	0.013	0.386	1.000	0.173	0.172	0.186	0.500
Stapes. anterior limb	0.036	0.018	0.242	1.000	1.000	0.074	0.036	0.074	0.212
Stapes. posterior limb	0.074	0.036	0.036	1.000	1.000	1.000	0.074	0.173	0.074
Incudomalleolar joint	0.807	0.212	0.036	0.500	1.000	0.681	1.000	1.000	0.977
Incudostapedial joint	0.962	0.862	0.725	0.172	0.500	0.173	0.574	0.173	0.074
Tendon of tensor tympani muscle	0.024	0.007	0.010	0.681	0.425	0.383	0.074	0.607	0.579
Tendon of the stapedius muscle	0.009	0.003	0.007	0.500	1.000	0.681	0.036	0.036	0.036
Chorda tympani: anterior canaliculus	0.445	0.102	0.117	1.000	0.977	0.977	0.074	0.173	0.036
Chorda tympani: tympanic segment	0.052	0.004	0.004	0.970	0.977	0.970	0.036	0.074	0.036
Chorda tympani: posterior canaliculus	0.036	0.036	0.500	1.000	1.000	0.074	0.500	0.500	1.000
Facial canal: tympanic segment	0.970	0.681	0.681	1.000	1.000	1.000	0.117	0.212	0.117
Tegmen tympani	0.074	0.173	0.036	0.500	0.807	0.681	0.681	0.102	0.212
Singular canal	0.607	0.500	0.074	1.000	1.000	1.000	0.500	0.500	0.173
Osseous spiral lamina: basal	0.013	0.018	0.013	0.977	0.500	0.970	0.074	0.500	0.500
Osseous spiral lamina: apical	0.173	0.977	0.173	1.000	0.681	1.000	0.386	0.172	0.173
Modiolous of cochlea	0.386	0.386	0.117	0.386	0.173	0.212	0.036	0.173	0.681
Cochlear aqueduct	0.500	0.500	0.087	1.000	1.000	0.681	0.500	0.500	0.386
Vestibular aqueduct	0.036	0.212	0.212	0.500	0.173	0.212	0.074	0.074	0.074
Petromastoid canal	1.000	0.977	0.977	0.500	0.173	0.173	0.173	1.000	1.000

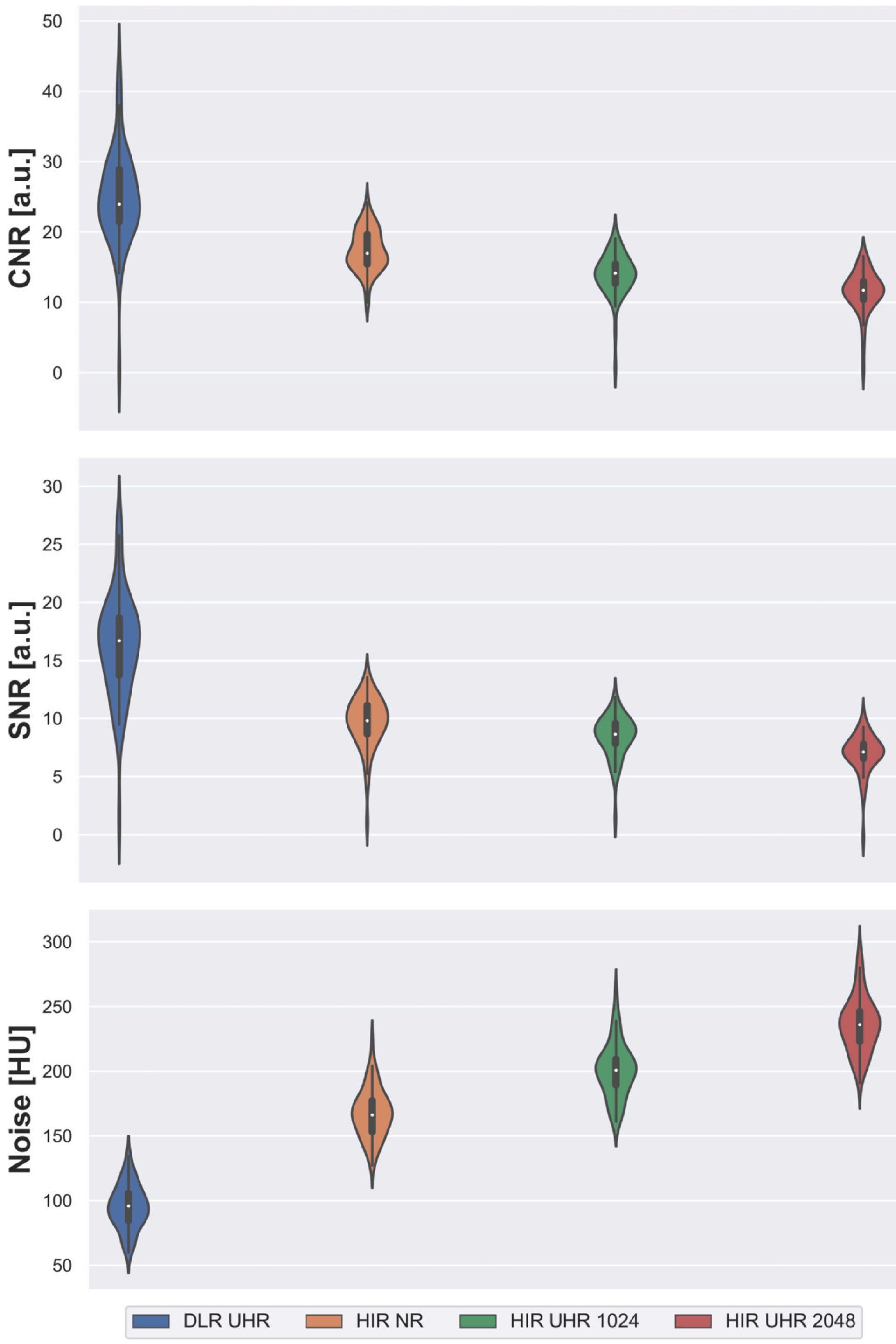

Figure 6. Image noise, signal to noise ratio and contrast to noise ratio: noise measured in the air of the meatus externus and the signal measured in the cerebellar hemisphere.

protocol subgroup, there was no significant difference between HIR UHR 2048 and HIR UHR 1024, most probably due to the smaller number of cases ($n = 10$). Only the tendon of the stapedius muscle showed a better definition in HIR UHR 2048 compared to HIR NR.

Our study is partially consistent with another study using a UHR-CT scanner with HIR across different matrix sizes and slice thicknesses (NR (0.5-mm slice thickness, 512^2 matrix), HR (0.5-mm slice thickness, 1024^2 matrix), UHR 1024 (0.25-mm slice thickness, 1024^2 matrix) and UHR

2048 (0.25-mm slice thickness, 2048^2 matrix). Their findings indicated significantly higher image quality for UHR 2048 and UHR 1024 compared to NR and HR. In contrast to our findings, they did not observe a difference between UHR 2048 and UHR 1024 (17).

Furthermore, previous studies have consistently demonstrated the superior delineation of fine anatomical structures with UHR-CT compared to ConvCT (18,19). Another study specifically highlighted the improved visibility of the tendon of the stapedius muscle with UHR-CT in comparison to ConvCT (20). Additionally, Heutink et al. were able to detect new bone formation after cochlear implant placement using UHR-CT highlighting the impact on long-term residual hearing preservation in patients (21).

Limitations of the underlying study are (i) the relatively small sample size, particularly in the pediatric protocol subgroup, (ii) the single-center study design and (iii) the fact that only patients without pathological findings were included in order to standardize analysis of multiple small anatomical structures. Future research should aim to evaluate the effect of AiCE Inner Ear on specific diseases of the temporal bone in larger prospective, muticenter studies. While this study is the first to include a pediatric protocol subgroup for evaluating DLR algorithms in temporal bone imaging, we acknowledge that the small number of children limits the generalizability of the findings. Future research should include larger pediatric cohorts to validate these results and assess the broader applicability of the findings. As noted by Kurokawa et al., certain microanatomical structures are better visualized using HIR compared to MBIR (22), while Hempel et al. emphasize that MBIR outperforms FBP in terms of image quality (23). Therefore further research should focus on comparing DLR with both MBIR and FBP.

CONCLUSION

Our findings suggest that deep learning-based image reconstruction of the temporal bone with ultra-high-resolution CT significantly improves image quality and enhances diagnostic performance, hereby potentially affecting treatment decisions.

DECLARATION OF COMPETING INTEREST

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Ahmed Othman reports financial support was provided by Canon Medical Systems Corporation. Ahmed Othman reports a relationship with Canon Medical Systems Corporation that includes: funding grants and speaking and lecture fees. The other 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.

APPENDIX A. SUPPORTING INFORMATION

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.acra.2025.02.002.

REFERENCES

1. Jäger L, Bonell H, Liebl M, et al. CT of the normal temporal bone: comparison of multi- and single-detector row CT. Radiology 2005; 235(1):133–141.
2. Beister M, Kolditz D, Kalender WA. Iterative reconstruction methods in X-ray CT. Physica Medica 2012; 28(2):94–108.
3. Niu YT, Mehta D, Zhang ZR, et al. Radiation dose reduction in temporal bone CT with iterative reconstruction technique. AJNR Am J Neuroradiol 2012; 33(6):1020–1026.
4. Hempel J-M, Niklas Bongers M, Braun K, et al. Noise reduction and image quality in ultra-high resolution computed tomography of the temporal bone using advanced modeled iterative reconstruction. Acta Radiologica 2019; 60(9):1135–1143.
5. Nichols JA, Herbert Chan HW, Baker MAB. Machine learning: applications of artificial intelligence to imaging and diagnosis. Biophys Rev 2019; 11(1):111–118.
6. Koetzier LR, Mastrodicasa D, Szczykutowicz TP, et al. Deep learning image reconstruction for CT: technical principles and clinical prospects. Radiology 2023; 306(3):e221257.
7. Kazerooni EA. High-resolution CT of the lungs. AJR Am J Roentgenol 2001; 177(3):501–519.
8. Koweek L, Achenbach S, Berman DS, et al. Standardized medical terminology for cardiac computed tomography 2023 update: an expert consensus document of the Society of Cardiovascular Computed Tomography (SCCT), American Association of Physicists in Medicine (AAPM), American College of Radiology (ACR), North American Society for Cardiovascular Imaging (NASCI) and Radiological Society of North America (RSNA) with endorsement by the Asian Society of Cardiovascular Imaging (ASCI), the European Association of Cardiovascular Imaging (EACI), and the European Society of Cardiovascular Radiology (ESCR). J Cardiovasc Comput Tomogr 2023; 17(5):345–354.
9. Fujita N, Yasaka K, Hatano S, Sakamoto N, Kurokawa R, Abe O. Deep learning reconstruction for high-resolution computed tomography images of the temporal bone: comparison with hybrid iterative reconstruction. Neuroradiology 2024; 66(7):1105–1112. https://doi.org/10.1007/s00234-024-03330-1
10. Wang TJ, Wang Y, Zhang ZH, et al. Deep learning reconstruction improves the image quality of low-dose temporal bone CT with otitis media and mastoiditis patients. Heliyon 2024; 10(1):e22810.
11. Geleijns JJ, K.A.; Panzer, W.; Shrimpton, P.C.; Tosi, G. 2000;Pages. Accessed at ⟨https://op.europa.eu/de/publication-detail/-/publication/d229c9e1-a967–49de-b169–59ee68605f1a⟩. Accessed 19.05.2023.
12. Altmann S, Abello Mercado MA, Ucar FA, et al. Ultra-high-resolution CT of the head and neck with deep learning reconstruction-assessment of image quality and radiation exposure and intraindividual comparison with normal-resolution CT. Diagnostics (Basel) 2023; 13(9):1534. https://doi.org/10.3390/diagnostics13091534
13. Tang R, Zhang Z, Zhao P, et al. A novel imaging scoring method for identifying facial canal dehiscence: an ultra-high-resolution CT study. Eur Radiol 2023; 33(4):2830–2839.
14. Fleiss JLL, B, Paik MC. The measurement of interrater agreement. Statistical Methods for Rates and Proportions. 3rd ed., Hoboken: Wiley; 2003.
15. Beysang A, Villani N, Boubaker F, et al. Ultra-high-resolution CT of the temporal bone: comparison between deep learning reconstruction and hybrid and model-based iterative reconstruction. Diagn Interv Imaging 2024; 105(6):233–242.
16. Nagayama Y, Sakabe D, Goto M, et al. Deep learning-based reconstruction for lower-dose pediatric CT: technical principles, image characteristics, and clinical implementations. Radiographics 2021; 41(7):1936–1953.
17. Pham N, Raslan O, Strong EB, et al. High-resolution CT imaging of the temporal bone: a cadaveric specimen study. J Neurol Surg B Skull Base 2022; 83(5):470–475.
18. Hiraumi H, Obara M, Yoshioka K, et al. Detectability of minute temporal bone structures with ultra-high resolution CT. Auris Nasus Larynx 2019; 46(6):830–835.

19. Ohara A, Machida H, Shiga H, et al. Improved image quality of temporal bone CT with an ultrahigh-resolution CT scanner: clinical pilot studies. Jpn J Radiol 2020; 38(9):878–883.
20. Yamashita K, Hiwatashi A, Togao O, et al. Ultrahigh-resolution CT scan of the temporal bone. Eur Arch Otorhinolaryngol 2018; 275(11):2797–2803.
21. Heutink F, Klabbers TM, Huinck WJ, et al. Ultra-high-resolution CT to detect intracochlear new bone formation after cochlear implantation. Radiology 2022; 302(3):605–612.
22. Kurokawa R, Maeda E, Mori H, et al. Evaluation of the depiction ability of the microanatomy of the temporal bone in quarter-detector CT: model-based iterative reconstruction vs hybrid iterative reconstruction. Medicine 2019; 98(24):e15991.
23. Hempel JM, Niklas Bongers M, Braun K, et al. Noise reduction and image quality in ultra-high resolution computed tomography of the temporal bone using advanced modeled iterative reconstruction. Acta Radiol 2019; 60(9):1135–1143.