Toward Intelligent Intraoperative Guidance in Temporal Bone Surgery

Sudanthi Wijewickrema*
Australian Centre for Artificial Intelligence in Medical Innovation
La Trobe University
Melbourne, Australia

Phu Lai†
Australian Centre for Artificial Intelligence in Medical Innovation
La Trobe University
Melbourne, Australia

Lu Yu‡
Australian Centre for Artificial Intelligence in Medical Innovation
La Trobe University
Melbourne, Australia

Jean-Marc Gerard §
Royal Victorian Eye and Ear Hospital
Melbourne, Australia

Figure 1: The Future Operating Room: Intelligent Extended Reality–Based Intraoperative Guidance for Temporal Bone Surgery

Abstract

Temporal bone surgery (TBS) involves complex dissection for ear pathologies and hearing restoration and demands exceptional precision because of nearby critical structures such as the facial nerve, labyrinth, and ossicles. Despite this, intraoperative decision-making still relies heavily on preoperative planning, surgeon experience, and indirect cues. Several key steps, including cochlear electrode insertion, cannot be directly controlled visually or by touch. Recently, image-based guidance and electrophysiological monitoring have shown promise, reducing uncertainty in traditionally "blind" steps. At the same time, extended reality (XR) technologies have rapidly advanced in surgical education and are beginning to enter clinical practice. We argue that the next major advance in intraoperative guidance for TBS will come from integrated systems that combine imaging and electrophysiological data with XR-based visualization. This work reviews the current state of guidance technologies, evaluates existing XR applications, identifies remaining gaps, and proposes a path toward XR-enabled, intraoperative guidance systems for TBS.

*e-mail: s.wijewickrema@latrobe.edu.au
†e-mail: p.lai@latrobe.edu.au
‡e-mail: l.yu@latrobe.edu.au
§e-mail: jean-marc.gerard@eyeandear.org.au

Index Terms: Intraoperative Guidance, Extended Reality, Temporal Bone Surgery, Intelligent Surgical Guidance.

1 Intraoperative Guidance in Temporal Bone Surgery

Temporal bone surgery (TBS) involves procedures on the side and base of the skull to address conditions such as hearing loss, ear infections, tumors, or traumatic injuries. Due to the close proximity of structures essential for hearing, balance, and facial function, TBS demands exceptional skill and precision. This challenge is heightened by the fact that surgical trauma in this region is frequently irreversible. For example, injury to the facial nerve can result in facial paralysis.

Advances in surgical microscopes have contributed significantly to patient safety and outcomes in TBS by improving visualization, depth perception, and anatomical precision. However, the view offered by the surgical microscope is inherently limited, providing only a surface-level, line-of-sight perspective. Underlying or obscured structures remain hidden until exposed through drilling. As such, some risk of surgical trauma remains, particularly in cases of abnormal anatomy, underlying pathology, or when procedures are performed by less experienced surgeons. Recent advances, from electrophysiological monitoring to image-based guidance, have helped to mitigate this risk.

1.1 Image-based Guidance

Image-based guidance assists the surgeon in understanding the location of instruments relative to critical structures during the procedure. Registration of preoperative models to intraoperative images provides real-time feedback, helping surgeons correlate what they see in the surgical field with deeper, obscured, or partially exposed structures. In TBS, these capabilities are delivered through a spectrum of modalities that trade off field-of-view, tissue contrast, latency, and workflow disruption.

1.1.1 Computed Tomography

Intraoperative computed tomography (CT) and Cone-beam CT (CBCT) offer one of the most practical mechanisms to "re-anchor" the surgeon's spatial understanding to a current 3D dataset once drilling has altered the operative field [25]. In cochlear implantation, intraoperative or mobile CBCT is most commonly used to confirm electrode position and identify gross malposition that may warrant immediate correction. Automated interpretation is also advancing; for example, detecting electrode tip fold-over directly from intraoperative CT/CBCT data [20]. However, intraoperative scanning interrupts operative flow and introduces ionizing radiation. Furthermore, interpretability can degrade precisely when it is most needed due to implant-related artifacts and limited soft-tissue contrast.

1.1.2 Projection Imaging

Projection imaging (X-ray/fluoroscopy) remains attractive because it is fast, widely available, and can provide near-real-time reassurance or early warning during implant insertion [20]. Synchronized fluoroscopy with intracochlear pressure measurements has helped characterize the mechanics of tip fold-over and supports the view that projection imaging can reveal distinctive signatures of malposition during insertion [17]. However, the core technical ceiling is geometric: 2D projections are limited for continuous anatomy-aware guidance in complex 3D spaces and are generally better at detecting gross malposition than quantifying proximity to hazards [28]. A recent systematic review [2] found that studies on intraoperative X-ray imaging in cochlear implantation were heterogeneous and included no randomized controlled trials. This indicates that current practice is guided more by local workflow and risk tolerance than by strong outcome-based evidence.

1.1.3 Optical Coherence Tomography

Optical coherence tomography (OCT) addresses a different gap by providing depth-resolved images of soft-tissue microstructure at the point of care. This raises the possibility of intraoperative tissue characterization rather than purely coordinate-based navigation. Recent work has aligned middle-ear OCT with temporal bone CT in both normal and pathological ears. These studies show that CT provides global bony anatomy, while OCT reveals local soft-tissue details that are often difficult to see on CT, supporting the feasibility of multimodal data fusion [50]. Contemporary clinic-facing OCT studies (including volumetric OCT and OCT angiography through the tympanic membrane) further indicate a broadening translational base for middle-ear pathology assessment [39]. However, OCT has a limited field of view and penetration depth, and remains sensitive to blood, fluid, and probe geometry.

1.1.4 Endoscopy

Endoscopy and image-based augmentation are increasingly used to overcome microscope line-of-sight constraints. Recent work has explored the feasibility of linking endoscopic video to preoperative CT through robust registration so that subsurface anatomy can be projected onto the surgeon's view. Taleb et al. [44] proposed a fully automatic 2D registration pipeline between preoperative temporal bone CT and otoendoscopic video, explicitly targeting elimination of fiducials and manual landmarking that slow adoption and introduce errors. Challenge of using endoscopy include loss of depth perception, limited views, sensitivity to viewpoint changes, and instrument crowding.

1.2 Electrophysiological Monitoring

Electrophysiological monitoring serves as the functional counterpart to image-based structural guidance in TBS. It fulfills three primary objectives: localization of neural structures obscured by complex anatomy or pathology [26], real-time detection of iatrogenic stress to prevent permanent injury [1], and prognostication of postoperative functional outcomes [12]. By translating physiological activity into audible or visual feedback, these systems allow surgeons to map the course of the facial nerve during dissection and assess cochlear integrity during implant insertion. However, current standards remain largely reactive rather than predictive and are frequently limited by signal artifacts, latency, and the requirement for expert interpretation.

1.2.1 Auditory Pathway Monitoring

Monitoring the functional integrity of the auditory system is critical during procedures involving the inner ear and cerebellopontine angle. While Auditory Brainstem Response (ABR) and Cochlear Nerve Action Potential (CNAP) monitoring remain the standard for detecting retrocochlear pathway compromise, they are often limited by averaging delays that prevent true real-time feedback [37]. Consequently, electrocochleography (ECochG) has emerged as the primary tool for guiding intracochlear steps, particularly during cochlear implant electrode insertion. Intraoperative drops in the amplitude of the cochlear microphonic, a potential generated by outer hair cells, are significantly correlated with intracochlear trauma and the subsequent loss of residual hearing [9]. To overcome the latency and noise inherent in raw electrophysiological signals, advanced model-based signal processing is being deployed. Recent research has demonstrated the ability to extract actionable signal trends related to trauma events without the delays associated with traditional averaging [5]. Furthermore, automated systems can now classify trauma patterns with high sensitivity, potentially creating an automated early warning system [30, 52].

1.2.2 Facial Nerve Monitoring

Facial nerve monitoring via electromyography (EMG) is the established standard of care for preventing iatrogenic injury to the facial nerve during surgery [26]. Artificial intelligence (AI) is increasingly employed to enhance the specificity and spatial context of facial nerve monitoring. AI models have demonstrated superior performance over traditional threshold-based methods in distinguishing true nerve responses from artifacts caused by electrocautery or anesthesia [54]. Emerging AI models are also enabling objective, automated grading of facial nerve function from surface EMG data, facilitating more consistent intraoperative assessments [34].

2 Current Role of XR in TBS

2.1 Training, Preoperative Planning and Rehearsal

In TBS, XR, particularly virtual reality (VR) simulators, has seen its greatest impact in training and education. The face, content, and construct validity of TBS simulators are well established [11, 38]. Their ability to improve objective performance metrics such as drilling economy, anatomical landmark identification, and injury avoidance, as well as reduction of operating time is also well substantiated [4, 14, 47]. The integration of automated feedback, procedural guidance, and objective assessment within simulation-based training has enabled the development of self-directed systems that can be used independently, without the need for continuous expert supervision [51, 53, 27, 42, 52].

The development of patient-specific models, derived from preoperative patient imaging, has enabled the use of these training systems for surgical planning and rehearsal. The utility of VR in planning TBS has been established [46, 13], while early feasibility studies have shown that case-specific rehearsal is technically feasible and is regarded highly by participants for its ability to improve surgical planning and confidence [6, 32].

Although these systems validate XR's capacity to represent complex anatomy and surgical sequences accurately, they remain largely disconnected from the realities of live surgery: real-time patient-specific challenges, dynamic tissue behavior, stresses of operating on real patients, and complexities of human interaction in the operating theatre.

2.2 Intraoperative Guidance

While VR is well suited for training, preoperative planning, and surgical rehearsal, it is not appropriate for intraoperative guidance because it occludes the surgeon's view of the operative field. In contrast, augmented reality (AR) and mixed reality (MR) integrate virtual content with the real world along the "virtuality continuum" [36], making them more suitable for intraoperative use.

High-fidelity head-mounted displays (HMDs), particularly those that support AR, allow for the projection of internal anatomical structures directly onto the patient's body by anchoring digital preoperative models and plans within the physical operative space [40]. Alternatively, AR projections can be integrated into real-time microscopic views or exoscopic video.

XR guidance systems typically follow a pipeline comprising image acquisition, 3D reconstruction, registration, and tracking. AI plays a central role in several stages of this pipeline. Volumetric preoperative CT and/or magnetic resonance imaging (MRI) scans are processed with AI models to automatically segment critical anatomical structures and generate virtual representations. AI models can also predict safe surgical corridors and generate preoperative plans. During surgery, registration and tracking algorithms, often incorporating AI, align these virtual models with the patient's physical anatomy in real time. AI can further enhance tracking robustness by compensating for tissue deformation, instrument-induced motion, or partial occlusions.

2.3 Early Prototypes

Research into XR-guided intraoperative TBS remains in its early stages, with several notable examples reported.

Tsuchida et al. [48] investigated the use of AR in transcanal endoscopic ear surgery for treating middle ear cholesteatoma. AR was employed to project the lesion onto the intraoperative endoscopic view, helping surgeons estimate how much external auditory canal bone needed removal and choose the optimal surgical approach. Their findings suggested that AR guidance can enhance intraoperative assessment and technique selection in cholesteatoma surgery.

Hussain et al. [22] evaluated the performance of a vision-based AR system for middle ear surgery by integrating real-time video from an operating microscope with virtual images generated from preoperative CT scans. Results demonstrated high clinical precision while maintaining stability even when the field of view was partially obstructed by bleeding or surgical instruments.

In a randomised controlled trial, Hadida et al. [19] investigated whether AR guidance could improve mastoidectomy performance on 3D-printed models in novice surgeons. Those using AR achieved significantly higher proficiency scores, demonstrating better margin definition, exposure of key anatomy, and preservation of critical structures compared with the control group.

Ito et al. [23] evaluated a novel AR system that integrated 3D holographic models directly into live exoscopic surgical footage in cochlear implant (CI) and Bonebridge surgeries. They found that the 3D graphical guides significantly improved performance metrics for CI by reducing mental, physical, and temporal demands as well as lowering frustration levels. However, for Bonebridge implantation traditional 2D guides performed better. The system demonstrated high stability with an imperceptible video latency.

Guigou et al. [18] evaluated the feasibility of an AR system to guide transmodiolar auditory implantation. They integrated preoperative CT scan data with real-time video from the surgical field for 3D resin models. The study concluded that AR can provide the sub-millimetric accuracy necessary to identify the extremities and axis of the cochlear modiolus intraoperatively, potentially leading to improved frequency stimulation and reduced electrode interference compared to standard cochlear implantation.

Bautista et al. [8] developed a visualization system to support CI surgery that integrated a semi-autonomous handheld surgical tool with an optical tracking system and an operating microscope. The microscope's live view was augmented with AR-based guidance to assist surgeons in positioning the tool and maintaining alignment during insertion. Their approach demonstrated promising tool alignment performance, comparable to state-of-the-art methods.

Tian et al. [45] investigated the feasibility and accuracy of using an HMD to guide the placement of bone-anchored hearing aid implants. They projected 3D holographic models highlighting critical anatomical features derived from preoperative CT scans onto the surgical field of cadaveric heads to assist implant positioning. Results showed that the preoperative digital models corresponded well with the actual implant positions reconstructed postoperatively.

Ito et al. [24] introduced a MR platform for lateral temporal bone resection. Using patient-specific 3D models derived from preoperative contrast-enhanced CT scans, they visualized holographic renderings of critical anatomy directly superimposed on the surgical field via a HMD. The authors concluded that using 3D holograms with HMDs provides a revolutionary tool for assisting in the navigation of complex otologic and skull base surgeries where traditional 2D monitors lack sufficient intuitive depth perception.

McJunkin et al. [35] developed a MR platform using a HMD to visualize interactive 3D holograms of temporal bone anatomy derived from CT scans projected onto physical cadaveric models. The platform generated accurate, interactive representations of soft tissue, bony landmarks, and internal ear structures to assist in lateral skull base surgery. However, the registration error remained too high for precise intraoperative use.

Hoing et al. [21] utilized digital microscopes with Picture-in-Picture (PiP) displays to directly display intraoperative ECochG potentials in the surgeon's field of view. The study demonstrated a positive effect of intraoperative ECochG visualization on residual hearing preservation outcomes compared to those without PiP.

Eichler et al. [16] showed that digital visualization of intraoperative ECochG in the surgeon's field of view can facilitate real-time feedback during CI insertion. In addition, simplified semantic visualizations, such as color-coded arrows indicating signal health, have proven to be as effective as complex raw graphs. This suggests that XR overlays can effectively convey critical functional data without distracting from the surgical task.

Shah et al. [43] and Saadya et al. [41] demonstrated how AR is addressing the spatial uncertainty of the facial nerve's course. Holographic overlays of the facial nerve, segmented from preoperative imaging, were successfully projected onto the patient's anatomy in pilot studies, improving surgical confidence by providing a form of "X-ray vision" and providing an effective tool for surgical planning and education. Complementing these digital overlays, novel molecular contrast agents like bevonescein are enabling fluorescence-guided surgery, illuminating the nerve structure to provide visual confirmation independent of electrical function [29].

3 Future of XR Enabled Intraoperative Guidance

3.1 Vision

While important research has been conducted on intraoperative guidance in TBS, a major limitation of current systems is that critical data streams remain siloed, requiring surgeons to mentally integrate information from electrophysiological monitoring, imaging, and preoperative planning. Although AI models already exist that process individual modalities (for example, for electrocochleography analysis, anatomical segmentation, and preoperative-to-real-time image alignment) these approaches largely operate in isolation. AI-based fusion models that integrate multimodal data into unified, intuitive AR overlays, combined with real-time microscope views, could enable faster interpretation and more effective intraoperative decision-making. Moreover, predictive models leveraging multiple modalities could provide early warnings and introduce non-intrusive redundancy by cross-validating spatial, imaging, and physiological signals, thereby enhancing surgical safety. Intelligent, real-time path planning could further assist surgeons in adapting to unforeseen challenges encountered in the operating room.

3.2 Challenges and Potential Solutions

Despite its conceptual promise, intelligent XR-based intraoperative guidance has not yet achieved widespread clinical adoption in TBS. This gap reflects not a lack of technical feasibility, but the convergence of unresolved challenges, both general to surgical guidance and specific to the high-stakes nature of TBS, which demands exceptional levels of validation, reliability, and trustworthiness. Addressing these challenges is essential to enable safe and effective intraoperative use.

3.2.1 Challenges Common to Multiple Surgical Domains

Intraoperative guidance systems must operate under strict latency constraints to ensure that visual overlays, warnings, and navigation cues remain perceptually aligned with the surgeon's actions. However, many state-of-the-art AI models are computationally intensive and ill-suited for low-latency inference in resource-constrained environments. Achieving real-time deployment necessitates the use of techniques such as model optimization and pruning, knowledge distillation, and compact architectures, all while maintaining a balance between performance, efficiency, and robustness.

Furthermore, real-time guidance pipelines should be structured hierarchically, with lightweight models providing immediate alerts, coarse tracking, and safety warnings, and more computationally demanding models running asynchronously for higher-level reasoning, uncertainty estimation, or post hoc analysis. Additionally, edge and embedded inference performed directly on devices integrated into the surgical microscope or HMD, supported by dedicated onboard parallel processing, can mitigate latency and reliability issues and ensure consistent performance under high computational load.

Long-duration surgical procedures impose significant physical and cognitive demands on surgeons, and current HMDs are often perceived as bulky and heavy, leading to neck strain and fatigue [15, 10]. Consequently, hardware design should prioritize lightweight, ergonomic HMDs with balanced weight distribution and adjustable support systems to minimize physical strain during extended use. At the same time, intraoperative guidance must be presented in an intuitive and contextually relevant manner to avoid information overload, distraction, and errors. Interfaces should be designed using principles from human–computer interaction and human–computer integration such as hierarchical information layering, context-aware alerts, adaptive display prioritization, and gaze- or gesture-based interaction.

Beyond physical and interface considerations, psychological barriers, particularly trust, remain a major challenge for clinical adoption. While AI assurance mechanisms addressing explainability, transparency, accuracy, and reliability [7] are essential, education, training, and familiarity also play a critical role in enabling acceptance and effective use [3].

Effective integration of XR-based guidance into the surgical workflow is critical to avoid disrupting sterile technique or restricting surgeon movement. HMDs, microscope adapters, and additional cabling can interfere with instrument handling and the overall ergonomics of the operating room. To address these challenges, AR displays can be incorporated directly into existing surgical microscopes or eyepieces, minimizing additional hardware in the field. Lightweight, ergonomically designed HMDs and sterile draping solutions will reduce physical and sterility-related constraints. Wireless, low-latency data links will eliminate cumbersome cables, while gesture- and eye-tracking interfaces, already available in some HMDs, enable intuitive, hands-free control of AR content, allowing surgeons to interact with guidance systems without breaking focus or compromising sterility.

A recent systematic review highlighted several ethical challenges associated with the intraoperative use of AR in surgery, including the "datafication" of patients and potential erosion of the patient–physician relationship [49]. These concerns can be mitigated through robust data governance frameworks, strict ethical standards, and clear communication with patients. The review also emphasized the risk of "negative training," whereby surgeons may inadvertently acquire inappropriate behaviors through over-reliance on digital overlays. This can be addressed by ensuring AR complements rather than replaces core surgical skills and by reinforcing training through structured curricula and simulation without AR [49]. In addition, the high cost of AR hardware, software licenses, and tracking infrastructure may limit adoption and exacerbate inequities in surgical care [33], motivating the development of cost-effective, modular systems, shared institutional resources, open-source platforms, and cloud-based or federated approaches.

Finally, regulatory approval requires clear evidence of patient safety and clinical benefit, necessitating structured and robust validation frameworks that demonstrate reliability, transparency, and explainability before routine clinical adoption.

3.2.2 Specific Challenges for TBS

While data requirements and validation are challenges for intraoperative guidance systems in general, the specific technical and clinical demands of TBS, (such as the small operating space, limited maneuverability, and proximity to critical anatomy and microstructures, where even minor errors can be catastrophic) introduce additional constraints that must be addressed.

Data quality remains a critical challenge, particularly given the sub-millimeter precision required in TBS. Many current intraoperative modalities lack sufficient spatial resolution for accurate segmentation, registration, and tracking. Although super-resolution techniques may partially compensate, they are not substitutes for true high-fidelity sensing, underscoring the need for parallel advances in acquisition hardware. Noise further complicates intraoperative data use, for example, in electrophysiological modalities, where signal quality varies across systems and conditions. While preprocessing and denoising can offer partial mitigation, long-term solutions will require hardware-level improvements and standardized acquisition pipelines.

Most AI approaches require large, well-annotated datasets, yet surgical data remain scarce, expensive to label, and highly heterogeneous. For example, standard practice varies across countries and centers, with some requiring both CT and MRI imaging for TBS, while others rely solely on CT. Both the spatial resolution of these image datasets and the quality of electrophysiological measurements vary considerably. Multi-center data consortia with standardized acquisition protocols, shared open benchmarks, and high-fidelity simulated datasets could enable more robust training and evaluation. Federated learning further supports collaborative model development without direct data sharing, while annotation burdens can be reduced through semi-supervised and active learning strategies, as well as automated detection of noisy or inconsistent labels.

Robust validation is essential for the safe deployment of AI- and XR-based intraoperative guidance systems, yet it remains largely lacking in TBS. For instance, a recent review by Liu et al. [31], which evaluated the quality of AI models in head and neck surgery using the AI assurance framework introduced by Batarseh et al. [7], found that most intraoperative AI models in TBS fail to meet these standards, often lacking external validation, prospective testing, and clinically meaningful performance metrics.

While multimodal data fusion can enhance robustness and precision through cross-validation and redundancy, integrating different modalities for TBS (for example, electrophysiological measurements with preoperative imaging and real-time microscopic data) introduces additional complexities that necessitate rigorous testing. Future validation efforts should adopt standardized, task-specific evaluation protocols that reflect real-world surgical constraints. This includes benchmarking against expert performance, assessing robustness to noise and anatomical variability, and performing external validation across institutions and devices. Continuous post-deployment monitoring and model-updating pipelines are also critical for detecting performance drift and maintaining safety over time.

XR-based guidance introduces additional validation requirements beyond algorithmic performance. They should be evaluated not only for spatial accuracy and registration fidelity, but also for perceptual reliability, depth consistency, occlusion handling, and temporal stability. Human-in-the-loop validation should assess whether the displayed information is interpretable, actionable, and non-distracting under realistic surgical conditions.

Finally, validation should follow a staged, risk-aware deployment pathway tailored to TBS. Initial testing on synthetic data and cadaveric specimens can establish baseline performance under controlled conditions. This should be followed by restricted studies in the operating room, where the system operates in an observational or advisory capacity without influencing surgical decisions. Only after demonstrating safety, robustness, and clinical utility should randomized controlled trials be conducted.

4 Conclusion

The future of intraoperative guidance in TBS lies in the convergence of XR, AI, low-latency networks, advanced data acquisition, and improved XR hardware. As AI models become more precise and capable of fusing multimodal data, intelligent XR platforms will provide unified, intuitive AR overlays that integrate microscope views, imaging, and physiological signals. Predictive models will deliver early warnings and real-time surgical planning, simultaneously cross-validating multiple modalities to introduce non-intrusive redundancy. Coupled with advances in high-resolution, low-noise data acquisition and lightweight, ergonomic HMDs, XR may evolve into an almost invisible yet indispensable layer of the surgical workflow. This has the potential to redefine the boundaries of human precision and patient safety in TBS.

References

[1] J. Abari, M. Matulic, P. Galeazzi, M. Z. Assadi, P. Van de Heyning, and V. Topsakal. Retrospective evaluation of facial nerve monitoring to prevent nerve damage during robotic drilling in the largest series of patients undergoing the hearo-procedure. *PloS one*, 20(6):e0326614, 2025. 2

[2] Z. Abbasi, S. Khoshsirat, and M. Khajavi. A systematic review of intraoperative x-ray images in cochlear implant. *Ear, Nose & Throat Journal*, p. 01455613251366053, 2025. 2

[3] M. Alberto, J. Xu, O. Patel, D. Bolton, and J. Ischia. Barriers to introducing new transformative surgical technology in australian healthcare: A comprehensive review and guide. *Société Internationale d'Urologie Journal*, 6(4):49, 2025. 4

[4] S. A. W. Andersen, S. Foghsgaard, L. Konge, P. Cayé-Thomasen, and M. S. Sørensen. The effect of self-directed virtual reality simulation on dissection training performance in mastoidectomy. *The Laryngoscope*, 126(8):1883–1888, 2016. 2

[5] R. R. Andonie, W. Wimmer, R. A. Wildhaber, M. Caversaccio, and S. Weder. Real-time feature extraction from electrocochleography with impedance measurements during cochlear implantation using linear state-space models. *IEEE Transactions on Biomedical Engineering*, 70(11):3137–3146, 2023. 2

[6] A. Arora, C. Swords, S. Khemani, Z. Awad, A. Darzi, A. Singh, and N. Tolley. Virtual reality case-specific rehearsal in temporal bone surgery: a preliminary evaluation. *International Journal of Surgery*, 12(2):141–145, 2014. 3

[7] F. A. Batarseh, L. Freeman, and C.-H. Huang. A survey on artificial intelligence assurance. *Journal of Big Data*, 8(1):60, 2021. 4, 5

[8] D. Bautista-Salinas, D. Kundrat, A. Kogkas, M. E. Abdelaziz, S. Giannarou, and F. R. y Baena. Integrated augmented reality feedback for cochlear implant surgery instruments. *IEEE Transactions on Medical Robotics and Bionics*, 3(1):261–264, 2020. 3

[9] C. W. Bester, L. Campbell, A. Dragovic, A. Collins, and S. J. O'Leary. Characterizing electrocochleography in cochlear implant recipients with residual low-frequency hearing. *Frontiers in neuroscience*, 11:141, 2017. 2

[10] E. J. Brown, K. Fujimoto, B. Blumenkopf, A. S. Kim, K. L. Kontson, and H. L. Benz. Usability assessments for augmented reality head-mounted displays in open surgery and interventional procedures: A systematic review. *Multimodal Technologies and Interaction*, 7(5):49, 2023. 4

[11] E. C. Compton, S. K. Agrawal, H. M. Ladak, S. Chan, M. Hoy, S. C. Nakoneshny, L. Siegel, J. C. Dort, and J. T. Lui. Assessment of a virtual reality temporal bone surgical simulator: a national face and content validity study. *Journal of Otolaryngology-Head & Neck Surgery*, 49(1):17, 2020. 2

[12] J. Cooper, J. Mittal, M. Zalta, N. DiStefano, D. L. Klassen, K. McKenna, D. A. Godur, A. Monterrubio, M. Moosa, R. Mittal, et al. Bridging the gap: A systematic review of intraoperative electrocochleography during cochlear implantation and preservation of residual hearing. *PloS one*, 20(5):e0323493, 2025. 2

[13] B. Copson, S. Wijewickrema, X. Ma, Y. Zhou, J.-M. Gerard, and S. O'Leary. Surgical approach to the facial recess influences the acceptable trajectory of cochlear implantation electrodes. *European Archives of Oto-Rhino-Laryngology*, 279(1):137–147, 2022. 2

[14] B. Copson, S. Wijewickrema, Y. Zhou, P. Piromchai, R. Briggs, J. Bailey, G. Kennedy, and S. O'Leary. Supporting skill acquisition in cochlear implant surgery through virtual reality simulation. *Cochlear Implants International*, 18(2):89–96, 2017. 2

[15] R. D'Amato, F. Cutolo, G. Badiali, M. Carbone, H. Lu, H. Hogenbirk, and V. Ferrari. Key ergonomics requirements and possible mechanical solutions for augmented reality head-mounted displays in surgery. *Multimodal Technologies and Interaction*, 6(2):15, 2022. 4

[16] T. Eichler, A. Lakomek, L. Waschkies, M. Meyer, N. Sadok, S. Lang, and D. Arweiler-Harbeck. Two different methods to digitally visualize continuous electrocochleography potentials during cochlear implantation: a first description of feasibility. *European Archives of Oto-Rhino-Laryngology*, 281(6):2913–2920, 2024. 3

[17] J. R. Gonzalez, N. D. Cass, R. M. B. Hartl, J. Peacock, S. P. Cass, and N. T. Greene. Characterizing insertion pressure profiles during cochlear implantation: simultaneous fluoroscopy and intracochlear pressure measurements. *Otology & Neurotology*, 41(1):e46–e54, 2020. 2

[18] C. Guigou, R. Hussain, A. Lalande, and A. B. Grayeli. Augmented reality based transmodiolar cochlear implantation. *Otology & Neurotology*, 43(2):190–198, 2022. 3

[19] D. Hadida Barzilai, S. Tejman-Yarden, D. Yogev, O. Vazhgovsky, N. Nagar, L. Sasson, R. Sion-Sarid, Y. Parmet, A. Goldfarb, and O. Ilan. Augmented reality-guided mastoidectomy simulation: A randomized controlled trial assessing surgical proficiency. *The Laryngoscope*, 135(2):894–900, 2025. 3

[20] C. Högerle, A. Englhard, F. Simon, I. Grüninger, R. Mlynski, J.-M. Hempel, and J. Müller. Cochlear implant electrode tip fold-over: our experience with long and flexible electrode. *Otology & Neurotology*, 43(1):64–71, 2022. 2

[21] B. Höing, T. Eichler, V. Juelly, M. Meyer, L. Jung, L. Waschkies, S. Lang, and D. Arweiler-Harbeck. Digital live imaging of in-

traoperative electrocochleography during cochlear implantation: the first 50 patients. *European Archives of Oto-Rhino-Laryngology*, 281(3):1175–1183, 2024. 3

[22] R. Hussain, C. Guigou, A. Lalande, and A. B. Grayeli. Vision-based augmented reality system for middle ear surgery: evaluation in operating room environment. *Otology & Neurotology*, 43(3):385–394, 2022. 3

[23] T. Ito, T. Fujikawa, T. Takeda, Y. Mizoguchi, K. Okubo, S. Onogi, Y. Nakajima, and T. Tsutsumi. Integration of augmented reality in temporal bone and skull base surgeries. *Sensors (Basel, Switzerland)*, 24(21):7063, 2024. 3

[24] T. Ito, Y. Kawashima, A. Yamazaki, and T. Tsutsumi. Application of a virtual and mixed reality-navigation system using commercially available devices to the lateral temporal bone resection. *Annals of Medicine and Surgery*, 72:103063, 2021. 3

[25] H. Jia, R. Torres, Y. Nguyen, D. De Seta, E. Ferrary, H. Wu, O. Sterkers, D. Bernardeschi, and I. Mosnier. Intraoperative conebeam ct for assessment of intracochlear positioning of electrode arrays in adult recipients of cochlear implants. *AJNR Am J Neuroradiol*, 39(4):768–774, Apr. 2018. doi: 10.3174/ajnr.A5567 2

[26] J. M. Kartush, K. S. Rice, R. E. Minahan, G. K. Balzer, C. D. Yingling, and C. N. Seubert. Best practices in facial nerve monitoring. *The Laryngoscope*, 131:S1–S42, 2021. 2

[27] T. Kerwin, G. Wiet, D. Stredney, and H.-W. Shen. Automatic scoring of virtual mastoidectomies using expert examples. *International journal of computer assisted radiology and surgery*, 7(1):1–11, 2012. 2

[28] S.-Y. Lee, J. H. Han, M. Carandang, Y. J. Bae, and B. Y. Choi. Simpler and effective radiological evaluations for modiolar proximity of a slim modiolar cochlear implant electrode. *Scientific Reports*, 10(1):17714, 2020. 2

[29] Y.-J. Lee, R. K. Orosco, M. Bouvet, J. D. Richmon, B. J. Berman, K. L. Crawford, M. Hom, Q. T. Nguyen, and E. L. Rosenthal. Intraoperative nerve-specific fluorescence visualization in head and neck surgery: a phase 1 trial. *Nature Communications*, 16(1):6060, 2025. 3

[30] Y. Liu, S. Wijewickrema, C. Bester, S. J. O'Leary, and J. Bailey. Time series representation learning with supervised contrastive temporal transformer. In *2024 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8. IEEE, 2024. 2

[31] Y. Liu, S. Wijewickrema, B. Copson, J.-M. Gerard, and S. Antani. Artificial intelligence assurance in head and neck surgery: Now and next. In *2025 IEEE 38th International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 977–982. IEEE, 2025. 4

[32] G. D. Locketz, J. T. Lui, S. Chan, K. Salisbury, J. C. Dort, P. Youngblood, and N. H. Blevins. Anatomy-specific virtual reality simulation in temporal bone dissection: perceived utility and impact on surgeon confidence. *Otolaryngology–Head and Neck Surgery*, 156(6):1142–1149, 2017. 3

[33] R. Magalhães, A. C. Lima, A. Marques, J. Pereira, and L. L. Santos. Usefulness of mixed reality in surgical treatment: Delphi study. *Journal of Medical Internet Research*, 27:e69964, 2025. 4

[34] I. Manzoor, A. Popescu, S. Ricchizzi, A. Spolaore, M. Gorbachuk, M. Tatagiba, G. Naros, and K. Machetanz. Automated neuromuscular assessment: Machine-learning-based facial palsy classification using surface electromyography. *Sensors*, 26(1):173, 2025. 2

[35] J. L. McJunkin, P. Jiramongkolchai, W. Chung, M. Southworth, N. Durakovic, C. A. Buchman, and J. R. Silva. Development of a mixed reality platform for lateral skull base anatomy. *Otology & Neurotology*, 39(10):e1137–e1142, 2018. 3

[36] P. Milgram and F. Kishino. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12):1321–1329, 1994. 3

[37] K. Niemczyk, I. Pobożny, R. Bartoszewicz, and K. Morawski. Intraoperative hearing monitoring using abr and tt-ecochg and hearing preservation during vestibular schwannoma resection. *Journal of Clinical Medicine*, 13(14):4230, 2024. 2

[38] S. J. O'Leary, M. A. Hutchins, D. R. Stevenson, C. Gunn, A. Krumpholz, G. Kennedy, M. Tykocinski, M. Dahm, and B. Pyman. Validation of a networked virtual reality simulation of temporal bone surgery. *The Laryngoscope*, 118(6):1040–1046, 2008. 2

[39] D. W. Pan, M. A. Morán, W. Kim, Z. Yang, B. E. Applegate, and J. S. Oghalai. Optical coherence tomography imaging and angiography of skull base tumors presenting as a middle ear mass in clinic. *Diagnostics*, 15(6):732, 2025. 2

[40] Z. Qi, M. H. Bopp, C. Nimsky, X. Chen, X. Xu, Q. Wang, Z. Gan, S. Zhang, J. Wang, H. Jin, et al. A novel registration method for a mixed reality navigation system based on a laser crosshair simulator: A technical note. *Bioengineering*, 10(11):1290, 2023. 3

[41] A. Saadya, S. Chegini, S. Morley, and M. McGurk. Augmented reality presentation of the extracranial facial nerve: an innovation in parotid surgery. *British Journal of Oral and Maxillofacial Surgery*, 61(6):428–436, 2023. 3

[42] C. Sewell, D. Morris, N. H. Blevins, S. Dutta, S. Agrawal, F. Barbagli, and K. Salisbury. Providing metrics and performance feedback in a surgical simulator. *Computer Aided Surgery*, 13(2):63–81, 2008. 2

[43] B. Shah, W. Xu, R. A. Bartholomew, N. Ben-Shlomo, A. Zhang, G. Fan, P. Unadkat, A. Ziaei, H. Zhou, J. P. Guenette, et al. Augmented reality for identification of temporal bone anatomy and comparison to conventional imaging. *Annals of Otology, Rhinology & Laryngology*, 134(11):828–836, 2025. 3

[44] A. Taleb, S. Leclerc, R. Hussein, A. Lalande, and A. Bozorg-Grayeli. Registration of preoperative temporal bone ct-scan to otoendoscopic video for augmented-reality based on convolutional neural networks. *European Archives of Oto-Rhino-Laryngology*, 281(6):2921–2930, 2024. 2

[45] X. Tian, Z.-q. Gao, Z.-h. Zhang, Y. Chen, Y. Zhao, and G.-d. Feng. Validation and precision of mixed reality technology in baha attract implant surgery. *Otology & Neurotology*, 41(9):1280–1287, 2020. 3

[46] T. Timonen, M. Iso-Mustajärvi, P. Linder, A. Lehtimäki, H. Löppönen, A.-P. Elomaa, and A. Dietz. Virtual reality improves the accuracy of simulated preoperative planning in temporal bones: a feasibility and validation study. *European Archives of Oto-Rhino-Laryngology*, 278(8):2795–2806, 2021. 2

[47] T. Timonen, M. Iso-Mustajärvi, P. Linder, H. Vrzakova, S. T. Sinkkonen, V. Luukkainen, J. Laitakari, A.-P. Elomaa, and A. Dietz. The feasibility of virtual reality for anatomic training during temporal bone dissection course. *Frontiers in Virtual Reality*, 3:957230, 2022. 2

[48] K. Tsuchida, M. Takahashi, T. Nakazawa, S. Kurihara, K. Yamamoto, Y. Yamamoto, and H. Kojima. Augmented reality-assisted transcanal endoscopic ear surgery for middle ear cholesteatoma. *Journal of Clinical Medicine*, 13(6):1780, 2024. 3

[49] F. Ursin, C. Timmermann, L. Benzinger, S. Salloch, and F.-A. Tietze. Intraoperative application of mixed and augmented reality for digital surgery: a systematic review of ethical issues. *Frontiers in surgery*, 11:1287218, 2024. 4

[50] J. Wang, F. Couvreur, J. D. Farrell, R. Ghedia, N. Shoman, D. P. Morris, and R. B. Adamson. Fusion of middle ear optical coherence tomography and computed tomography. *JAMA Otolaryngology–Head & Neck Surgery*, 151(5):476–484, 2025. 2

[51] S. Wijewickrema, X. Ma, P. Piromchai, R. Briggs, J. Bailey, G. Kennedy, and S. O'Leary. Providing automated real-time technical feedback for virtual reality based surgical training: is the simpler the better? In *International Conference on Artificial Intelligence in Education*, pp. 584–598. Springer, 2018. 2

[52] S. Wijewickrema, B. J. Talks, J. Lamtara, J.-M. Gerard, and S. O'Leary. Automated assessment of cortical mastoidectomy performance in virtual reality. *Clinical Otolaryngology*, 46(5):961–968, 2021. 2

[53] S. Wijewickrema, Y. Zhou, I. Ioannou, B. Copson, P. Piromchai, C. Yu, R. Briggs, J. Bailey, G. Kennedy, and S. O'Leary. Presentation of automated procedural guidance in surgical simulation: results of two randomised controlled trials. *The Journal of Laryngology & Otology*, 132(3):257–263, 2018. 2

[54] X. Zha, L. Wehbe, R. J. Sclabassi, Z. Mace, Y. V. Liang, A. Yu, J. Leonardo, B. C. Cheng, T. A. Hillman, D. A. Chen, et al. A deep learning model for automated classification of intraoperative continuous emg. *IEEE transactions on medical robotics and bionics*, 3(1):44–52, 2020. 2