Auris Nasus Larynx 51 (2024) 666–673

Contents lists available at ScienceDirect

Auris Nasus Larynx

journal homepage: www.elsevier.com/locate/anl

Machine learning application in otology

Hajime Koyama

Department of Otorhinolaryngology and Head and Neck Surgery, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

A R T I C L E I N F O

Keywords:
Machine learning
Otology
Surgical outcome prediction
Deep learning

A B S T R A C T

This review presents a comprehensive history of Artificial Intelligence (AI) in the context of the revolutionary application of machine learning (ML) to medical research and clinical utilization, particularly for the benefit of researchers interested in the application of ML in otology. To this end, we discuss the key components of ML—input, output, and algorithms. In particular, some representation algorithms commonly used in medical research are discussed. Subsequently, we review ML applications in otology research, including diagnosis, influential identification, and surgical outcome prediction. In the context of surgical outcome prediction, specific surgical treatments, including cochlear implantation, active middle ear implantation, tympanoplasty, and vestibular schwannoma resection, are considered. Finally, we highlight the obstacles and challenges that need to be overcome in future research.

1. Introduction

Throughout the history of medicine, new techniques have continually shaped medical research and clinical practice. Recent advances in computer science, a field that is rapidly impacting multiple fields, have led to progress in machine learning (ML), particularly within the realm of artificial intelligence (AI). Although ML has been utilized in otolaryngology to some extent, its application lags behind that in other medical fields in terms of research volume. This review provides a clear overview of the historical context and highlights the impact of computer science and the growing role of ML in medicine. Specifically, we examine ML applications in otology, highlighting its current status, challenges, and successes in the field. We also consider the future, envisioning the untapped potential and transformative opportunities offered by ML in the diagnosis and treatment of ear, nose, and throat diseases. This exploration is expected to serve as a guide for researchers navigating the intersection of technology and medicine, and help shape the future landscape of otolaryngology.

2. A brief history of artificial intelligence

ML is a type of AI framework, which aims to enable machines to mimic human intelligence. This concept was proposed by Turing [1] and the terminology was introduced in 1956 in “The Dartmouth Summer Research Project on Artificial Intelligence”. AI was initially regarded as mere imitation of the step-by-step human reasoning using simple series of “if, then” rules. It was applied to reasoning or problem-solving tasks owing to the simplicity of their formulation in terms that were comprehensible to computers. Subsequently, knowledge representation has become a major topic of research, which aims to replace expert judgment by setting predefined rules of judgment. For example, knowledge representation system was used for the bacterial infection diagnosis system, Mycin [2]. Although such methods performed very effectively, they suffered from certain shortcomings, including the necessity for sufficiently comprehensive rules or knowledge to perform decision-making in the real world and certain contraindications between rules. In addition, they required significant processing power, with impractical processing times. Therefore, between the 1970s and the 2000s, funding and research interest in AI dwindled, leading to the so-called “AI winter” [3].

With the introduction of new algorithms and improvement in the processing power of computers in the 2000s, the solution of complex tasks within an acceptable time frame became possible. Several epoch-defining events occurred at this time, e.g., Big Blue, an AI developed by IBM, beat the world chess champion in 1997; Watson developed by IBM won first place in the television quiz show Jeopardy! in 2011; and AlexNet, a deep-learning system requiring a large amount of processing, dominated the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), an image classification contest, with a top-5 error of 15.3 %, which was lower than those of other methods by more than 10.8 %, in 2012. Since then, AI research has progressed rapidly [4].

E-mail address: hakoyama-tky@umin.ac.jp.

https://doi.org/10.1016/j.anl.2024.04.003
Received 7 December 2023; Accepted 2 April 2024
Available online 4 May 2024

3. The concept of machine learning

ML aims to find patterns in data by learning or adapting based on algorithms or statistical models, without following predefined instructions. ML comprises three components—input, output, and algorithms—it predicts outputs based on inputs using algorithms.

ML is primarily classified into three groups depending on the mode of output generation—supervised, unsupervised, and reinforcement learning. Supervised learning algorithms construct models based on datasets containing both inputs as well as desired outputs. The models predict outputs based on the inputs, compares them with the desired outputs, and adjusts the model to reduce the errors between the model's output and the desired output. In contrast, unsupervised learning algorithms construct models based on datasets that do not contain the desired output. Such models attempt to identify patterns based on a large number of inputs. This often requires a large amount of data and is primarily used to solve classification and dimension reduction problems. A large amount of inconsistent data can be sorted at first glance, or key features can be extracted from high-dimensional data. Reinforcement learning proceeds via recurrent trial-and-error processes to derive an appropriate prediction [5]. Positive or negative feedback strengthens the connection between the environment and the model, and promotes desired outcome and appropriate actions [6].

The type of ML is usually selected based on the research question or the goal of a study. However, in medical tasks, supervised learning algorithms have been the widely used type owing to the nature of medical datasets. In medical applications, patient characteristics, such as age, preoperative condition, and type of surgical intervention, are often the targets of analysis, and patient diagnosis, examination results, and postoperative outcomes are the main prediction targets. Thus, patient attributes are considered as inputs and the diagnosis, examination result, and postoperative outcome is considered to be desired outputs. Supervised models are constructed based on datasets comprising both inputs and outputs, and postoperative outcomes are predicted based on different patient attributes (of new patients).

3.1. The inputs

A wide variety of data types can be handled by ML as inputs in healthcare applications. Patient demographics, such as age, sex, and medical history, and the results of clinical tests, such as pure-tone audiometry thresholds, radiological imaging, and gene expression, can all be treated as input data. Input data type is usually selected manually based on previous reports or clinical experience. However, some ML algorithms select important inputs automatically and ignore the rest based on learning. Some ML algorithms require pre-processing of the input data, in the form of normalization or standardization.

Input data can be divided into two types: structured data and unstructured data. Structured data, such as pure-tone audiometry thresholds, age, and clinical staging, is quantitative, formatted, and easy to manipulate or analyze using a computer. In contrast, unstructured data, such as medical records and video files, is stored in a variety of forms and doesn't have a fixed format, making it difficult to analyze. Although unstructured data are complex and difficult to handle for computer [7], previous reports clarified the importance of unstructured data [8–10], partly because unstructured data represent the complexity of patients' condition, which increases the reliability of analysis [11].

3.2. The outputs

In supervised learning, the outputs (sometimes called objective variables) comprise results to be predicted by the examiner. In medicine, outputs are primarily clinical outcomes, such as diagnoses, test results, and medical or surgical outcomes.

ML tasks can be categorized as classification or regression tasks depending on whether their outputs are categorized as discrete or continuous variables, respectively. However, regression tasks can be transformed into classification tasks via categorization of continuous output variables. Generally, classification tasks are easier than regression tasks. The evaluation of an ML system depends on the type of task involved.

3.3. Types of algorithms

Several algorithms have been documented in ML, each tailored to specific tasks [12]. No single algorithm universally outperforms others in all contexts; therefore, the algorithm type must be chosen judiciously depending on the nature of the input data and desired outputs [13]. Given the impracticality of exhaustively cataloging all algorithms in this review, we focus on a subset of representative algorithms that are commonly used in healthcare analysis [14].

The decision-making process regarding algorithm selection is of paramount importance in data analysis and predictive modeling in the healthcare domain. The algorithms under review encompass diverse methodologies, ranging from decision trees (DTs) and support vector machines (SVMs) to k-nearest neighbors (k-NN) and neural networks (NNs). Each algorithm exhibits unique characteristics that make it suitable for specific applications in healthcare [15]. The choice of algorithm influences the accuracy, interpretability, and generalizability of the resulting ML model fundamentally.

In the following sections, representative algorithms are discussed, highlighting their underlying mechanisms and practical applications in healthcare analytics. This academic endeavor is intended to contribute to the broader discourse on the intricate interplay between algorithmic selection and the nuances of healthcare data analysis.

3.3.1. Lasso regression

Similar to multiple regression analysis, LASSO regression is distinguished by the inclusion of a regularization term, which makes it a powerful tool for variable selection and regularization [16]. In traditional multiple regression, the loss function is defined as follows:

$$L = (y - Xw)^T (y - Xw)$$

where L represents the loss function, y represents the outcome or the objective variable, X represents explanatory variable, and w represents the weight.

In contrast, Lasso regression introduces a regularization term, resulting in the following loss function:

$$L = \frac{1}{2}(y - Xw)^T (y - Xw) + \lambda \| w \|_1$$

where λ denotes the regularization parameter and $\| w \|_1$ represents the L1 norm of the regression coefficients.

The addition of the penalty term, $\lambda \| w \|_1$, plays a crucial role in both variable selection and regularization. This regularization term imposes a constraint on the magnitude of the coefficients and promotes sparsity in the model by driving certain coefficients to zero. Consequently, LASSO regression facilitates the identification of the most influential variables, effectively performing automatic feature selection [17].

In addition, the regularization term promotes model robustness by preventing overfitting, particularly when the number of predictors exceeds the number of observations. The dual functionality of LASSO regression, including variable selection and regularization, enhances prediction accuracy and generalizability, making it a valuable asset in predictive modeling and feature engineering efforts. The nuanced integration of the regularization parameter allows practitioners to fine-tune the trade-off between accuracy and sparsity by adapting the model to the specific needs of the data at hand [18–20].

3.3.2. K nearest neighbor (Fig. 1(A))

The k-NN algorithm is a simple and intuitive approach applicable to classification and regression tasks. It operates based on the principle of proximity, in which each new data point is judged based on its similarity to its k nearest neighbors [21,22]. Thus, in classification tasks, new data are classified based on the most common class among their k nearest neighbors, whereas in regression tasks, the prediction is often determined based on the average of the values corresponding to these neighbors.

The concept of "nearest" is defined in terms of distance, typically measured using metrics such as Euclidean distance or Manhattan distance [23]. However, in medical contexts, where variables often represent different types of data, normalization is critical to ensure fair comparison as well as improve accuracy and reliability. Normalization ensures that each variable contributes proportionally to distance calculation, preventing features with larger scales from dominating the algorithm [24].

It is worth considering the implications of choosing an appropriate value for k because smaller values can make the model sensitive to noise, whereas larger values can lead to oversimplification [25]. Further, although k-NN is known for its simplicity and ease of implementation, it may suffer from sensitivity to outliers and high computational cost, especially in scenarios involving large datasets [26].

3.3.3. *Support vector machine (Fig. 1(B))*

The SVM algorithm is used to define boundaries between data points. The SVM treats data as p-dimensional vectors, where p denotes the number of features. The algorithm aims to construct a (p-1)-dimensional hyperplane with the maximum margin from the nearest data vectors [27]. This hyperplane serves as the decision boundary, effectively separating different classes in the dataset. It is important to note that the SVM is not limited to classification tasks; it can also be applied to regression tasks [28], making it a versatile tool in various domains.

SVM strives to identify the global optimal solution, which contributes to its effectiveness in capturing complex relationships within the data [29]. The implementation of an SVM is straightforward, and its versatility has led to its widespread use in medical research. For example, SVMs have been successfully applied to cancer diagnosis [30], biomarker selection for Alzheimer's disease [31], and novel drug discovery [32]. Its popularity in the healthcare domain is evident, with SVMs being the most widely used algorithm, constituting 42 % of all ML applications in healthcare [33].

3.3.4. *Decision tree (Fig. 1(C))*

The DT algorithm is a hierarchical model that exhibits a tree-like structure, with branching points involving classification conditions. This structure is arranged hierarchically, similar to a flowchart, to provide a systematic and interpretable representation of DT processes [34]. In ML, algorithms construct DTs based on given inputs, make assumptions regarding the output, and iteratively adjust the classification conditions [35].

The structure of a DT is defined by its nodes, which represent decision points, and branches, which represent possible outcomes. Each branching point in the tree corresponds to a classification condition, and these conditions are organized hierarchically based on their relevance to the decision process. The algorithm constructs a DT by evaluating the input features and making decisions at each node, ultimately leading to prediction of the target output.

A notable feature of DT is its interpretability [36]. As a tree grows, it becomes a visual representation of decision logic, enhancing its comprehensibility and interpretability. DTs are versatile, and can be applied to both classification and regression tasks.

The DT algorithm is used to predict clinical parameters but it often serves more as the basis for ensemble learning methods, such as Random Forest (RF) and extreme Gradient Boosting (XGBoost) algorithms, which are explained below.

3.3.5. *Random forest (Fig. 1(D))*

The RF algorithm represents the convergence of DTs and ensemble learning, and provides a robust and versatile approach in various domains, particularly in clinical applications [37]. Consisting of a collection of DT algorithms arranged in parallel, the defining characteristic of RF lies in the random determination of the input data and features for each tree [38].

In clinical contexts, the ensemble-learning nature of RFs, formed by merging multiple algorithms, contributes significantly to increased model robustness [39]. In addition, the algorithm's use of random inputs and features results in a low correlation between individual trees, which enhances its ability to capture diverse patterns in clinical data. In particular, RF eliminates the need for variable normalization, which accommodates the inherent variability in clinical data and often includes a mix of categorical and numerical variables. This flexibility allows the processing of multiple data types without the need for complex pre-processing steps [40].

In addition, RF serves as a valuable tool for exploring feature importance, which is a critical step in clinical environments [41]. This feature allows practitioners to identify influential factors that contribute to predictions and provides critical insights to benefit decision-making

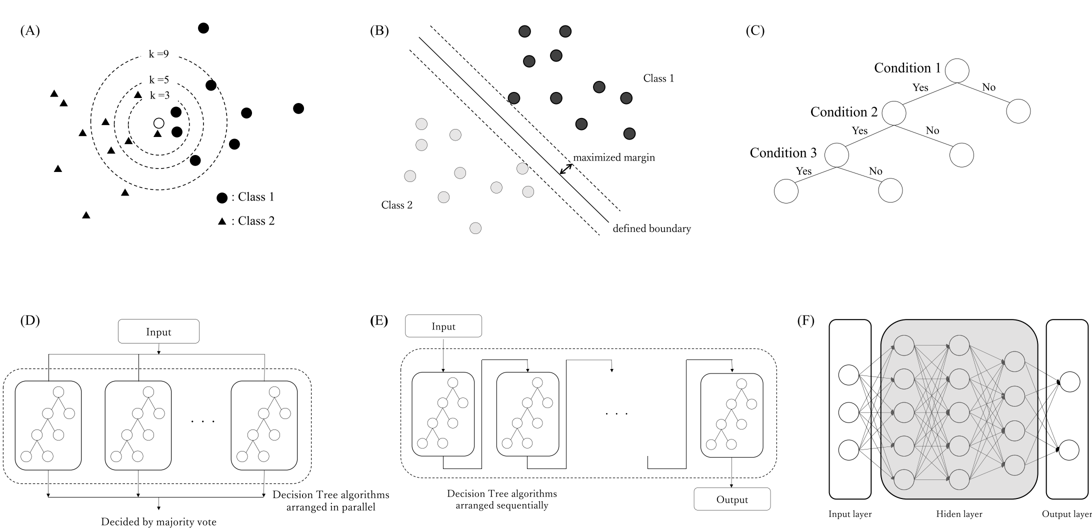

Fig. 1. Schemas of representative machine learning algorithms, (A): K nearest neighbor, (B): Support vector machine, (C): Decision tree, (D): Random forest, (E): eXtreme gradient boosting, (F): Neural network.

[42]. In the clinical landscape, where understanding influential factors is paramount, the ability of RF to assess and prioritize these factors positions it prominently among ML algorithms. Its ability to handle diverse data types, resist overfitting, and provide interpretable insights makes RF a powerful and preferred choice in clinical research and predictive modeling [43].

3.3.6. eXtreme gradient boosting (Fig. 1(E))

XGBoost is an advanced ML algorithm based on the DT framework. It differs from traditional DTs as it constructs a direct connection between them and creates an ensemble learning model. Ensemble learning involves combining multiple weak learners, such as DTs, to form a robust and accurate predictive model [44,45].

The XGBoost algorithm adds DTs sequentially, each of which corrects mistakes made by its predecessors. This iterative process results in a highly adaptive and powerful model. A notable advantage of XGBoost is its ability to handle complex relationships within data, making it particularly effective at capturing intricate patterns [46].

However, XGBoost is sensitive to the selection of explanatory variables that can significantly affect the performance. Sensitivity to variable selection is a characteristic shared by other ensemble learning methods [47]. Despite this drawback, when explanatory variables are selected judiciously, XGBoost exhibits superior predictive accuracy compared to other ensemble learning algorithms, such as RF [46]. Its predictive accuracy, especially when properly configured, positions XGBoost as a powerful tool in the ML arsenal, contributing to its widespread adoption in various domains, including finance [48–50], healthcare [51–53], and data science [54,55].

3.3.7. Neural networks (Fig. 1(F))

Inspired by the intricate structure and function of the human brain, this algorithm operates as a neural network, a concept fundamental for mimicking human cognitive processes [56]. The NN consists of layers of interconnected nodes, where the nodes are analogous to neurons and the connections between them reflect neural activity. In computer science, these nodes encapsulate information and the connections carry weights that influence the flow of information.

Within this network, nodes are categorized into three main types—input, hidden, and output [57]. Input nodes serve as conduits for real-world data and are particularly relevant in medical contexts, where they represent explanatory variables [58]. Attributes, such as age, hearing threshold, and history of otorrhea, are encapsulated within the input nodes. Hidden nodes, aptly named for their lack of a direct connection to the outside world, play a central role in processing and transforming input data. They are essential for capturing complex patterns and relationships within information [59].

The output nodes are intricately linked to real-world data, specifically representing objective variables in medical environments [60]. In the context of a predictive model, these objective variables represent the outcomes to be predicted by researchers. Thus, an NN orchestrates a sophisticated combination of nodes and connections that translate input data into meaningful predictions, making it a powerful tool in the fields of predictive modeling and ML.

3.3.8. Deep learning

Deep learning (DL), a specific type of NN, is characterized by the incorporation of a large number of "deep" hidden layers into the learning process, rendering functions that are unattainable by classical NNs [61]. The term "deep" refers to the substantial depth of these NNs, which allows them to handle complex, non-linear data with varying structures effectively. This characteristic makes DL particularly adept at processing complex datasets that classical NNs find difficult to handle [62]. DL has garnered considerable attention in the field of image analysis primarily because image data inherently possess complex, voluminous, and non-linear characteristics. A notable example of deep learning application in image analysis is the use of convolutional neural networks (CNNs) [63]. CNNs recognize intricate patterns in images and create specific filters [64]. These filters, in turn, combine all identified features into connected layers, allowing for comprehensive analysis of visual data [65].

Several CNN algorithms have been developed to improve image analysis, each offering unique advantages. Examples include AlexNet [66], VGG [67], GoogLeNet [68], ResNet [69], and U-Net [70]. These algorithms contribute to the versatility and effectiveness of deep learning in handling image data, providing a range of tools for researchers and practitioners in various fields that require sophisticated image analysis techniques [71,72].

4. Machine learning applications in otology

4.1. Diagnosis

Diagnostic applications of ML have been the main targets of research since the early era of AI. The DL algorithm exhibits significant advantages, especially in image analysis; therefore, with improvements in the DL technique, its application has skyrocketed. In otology, the diagnosis of the tympanic membrane is the most common application of ML. However, it has also been applied to other tasks, such as the classification of hearing disorders.

Lee et al. used 1338 tympanic membrane photographs, predicted side judgments, and identified tympanic perforations using a CNN [73]. They reported the accuracy of side detection to be 97.9 %, and that of perforation detection to be 91.0 %. Another study analyzed a higher number of data points—10,544 tympanic photos—and created a program to diagnose six common ear diseases (normal, attic retraction, tympanic perforation, otitis externa±myringitis, tumor) [74] using CNN algorithms. The authors reported an average accuracy of 93.67 %. Wu et al. collected 12,203 pediatric otoendoscopic images and designed a program to diagnose normal, acute, and otitis media with effusion with 97.45 % accuracy, which achieved 90.66 % accuracy even on tympanic photos captured using a smartphone [75]. A novel platform (Teachable machine®) was introduced, which automatically creates a classification program without any programming knowledge on the part of the user. Analysis of 3024 tympanic membrane images using this platform revealed that the accuracy of detecting normal or abnormal cases was 90.1 %; the accuracy of detecting normal, OME, or COM + cholesteatoma was 89.0 %; and the accuracy of discriminating between all of these cases for pathology was 86.2 % [76]. Zeng et al. attempted to predict hearing levels based on otoscopic images using a DL model [77]. They used 2790 otoscopic images and achieved 81 % accuracy in predicting conductive hearing loss using the DL model.

Diagnosis based on audiometry data, which contain a large amount of information, has been a long-standing target for ML applications. Carey et al. reported a diagnostic algorithm using 767 vestibular schwannoma audiometry data points and 2000 normal audiometry data points [78]. They reported that ML algorithms performed similarly in identifying patients with vestibular schwannomas.

Computed tomography (CT) images are often used as targets for ML applications. Wang et al. analyzed 1147 temporal CT images (562 chronic otitis media (COM) images and 672 normal ear images) using a CNN model, and achieved 76.2 % accuracy in the differential diagnosis between these two diseases [79]. Eroğlu et al. applied DL to identify the existence of cholesteatoma in COM patients [80]. They reported 95.4 % accuracy in detecting cholesteatoma in patients with COM. Ayral et al. reported a similar result (93.33 % accuracy of the CNN model) in the identification of cholesteatoma in COM, and concluded that DL is a useful tool for improving diagnostic accuracy [81]. Su et al. tackled similar tasks using a 3-dimensional CNN instead of a 2-dimensional CNN and reported a 96.5 % accuracy in classifying COM, cholesteatoma, and normal ears [82]. Takahashi et al. focused on the presence of cholesteatomas in the mastoid area and developed a program to predict mastoid extension in pars flaccida cholesteatomas using DL [83]. They reported a mean accuracy of 81.14 % and concluded that DL was effective for automatic diagnosis and surgical approach selection.

Fujima et al. analyzed 198 temporal CT images using DL algorithms to detect otosclerosis [84]. They reported that the DL algorithm, AlexNet, achieved 89 % accuracy in diagnosing otosclerosis based on CT images. Ogawa et al. analyzed temporal bone CT images of 6663 normal inner ears and 113 malformations to detect and localize abnormalities using 3D variational autoencoder [85]. They reported significant discrimination between normal and abnormal cases, with 92.0 % specificity and 99.1 % sensitivity.

4.2. *Influential factors*

Some clinical data affect the disease state and are called influential risk or prognostic factors. Classical statistical analysis mainly aims to identify these factors in patients with certain disorders, in contrast to the main purpose of ML, which is to predict outcomes. However, some ML algorithms, such as RF, can identify factors important to prediction, resulting in the determination of influential factors.

Profant et al. analyzed patients with vestibular schwannomas and applied a DT algorithm to identify influential factors in decision-making [86]. They reported that the tumor size, speech discrimination score, and hearing threshold were the main influential factors. Gadot et al. used tree-based algorithms to predict active treatment and influential factors in decision-making and concluded that ML achieved 80–88 % accuracy in predicting active treatment. They identified maximum tumor dimension, age at presentation, Koos grade, and progressive symptoms to be the most influential factors [87]. Tang et al. surveyed 32,465 patients in the National Inpatient Sample, applied Akaike Information Criterion-based model selection, and created a customized risk stratification score that included 5 variables (age $\geq$60 years, hydrocephalus, preoperative cranial nerve palsies, diabetes mellitus, and hypertension), predicting high mortality [88].

4.3. *Surgical outcome*

Surgery plays a pivotal role in the treatment of ear diseases, with procedures such as cochlear implantation [89–92], active middle ear implantation [93,94], and tympanoplasty [95–98] achieving significant success in improving patients' quality of life. Despite these advances, predicting surgical outcomes remains challenging because of the complex interplay between multiple patient characteristics and disease states [99–101]. Traditional analysis methods, such as linear regression [102–104] or risk stratification [105–107], have limited ability to provide accurate predictions under such complex circumstances. In response to these challenges, researchers are increasingly resorting to ML for such tasks owing to its unparalleled ability to construct robust predictive models.

ML algorithms offer a promising approach to achieve accurate predictions over a spectrum of surgical treatments. These algorithms have the potential to redefine our ability to predict surgical outcomes [108, 109], owing to their ability to identify intricate patterns and relationships in large datasets. ML promises to revolutionize the field by enabling a personalized approach to treatment, tailored to the unique characteristics of each patient [110,111]. As we delve deeper into the intersection of surgery and technology, the horizon of medical research expands, offering exciting potential for personalized and optimized treatment of ear diseases. This evolving landscape is expected to usher in a new era of healthcare, in which the fusion of surgical expertise and technological innovation opens doors to unprecedented advances in patient care [112].

4.3.1. *Cochlear implantation*

Cochlear implantation (CI) is an effective treatment for severe-to-profound hearing loss. Using this device, patients can hear sounds, understand their meaning, and communicate verbally, all of which improve their quality of life. However, the postoperative performance of cochlear implants remains variable among patients with different backgrounds. Identification of factors influencing postoperative outcomes has been a long-standing problem, and age, residual hearing thresholds, duration of hearing loss, surgical technique, and electrode type have been reported as potential prognostic factors. Despite this research, predicting postoperative outcomes remains difficult.

Recently, an ML approach was applied to solve this problem and establish prediction models. Tan et al. used SVM in supervised and semi-supervised models to analyze functional MRI and predict language performance 2 years after CI [113]. They reported that their ML program achieved an accuracy of 81.3 % and an area under the curve (AUC) of 0.97. Another study reported that a ML program could predict postoperative performance with an accuracy of 95.4 % in the classification task [114]. The NN used in this study was accurate, and another algorithm, XGBoost, was used to examine the prognostic factors and identify over 20 influential variables, including preoperative sentence test performance, age, and tinnitus handicap inventory.

With the expansion of the criteria for CI and its applied to patients with residual low frequency hearing or new devices, hearing preservation has become one of the main concerns following implantation. Zeitler et al. used several ML algorithms to predict the lowest quartile of postoperative changes in acoustic hearing [115]. They compared the algorithms and concluded that an RF algorithm achieved the best performance (AUC = 0.83).

Currently, CI is used in patients with various etiologies, including genetic hearing loss, viral infection, and sudden sensorineural hearing loss. Among the etiologies of hearing loss, cochlear nerve deficiency (CND) is challenging. Patients with CND have fewer cochlear nerve fibers and insufficient spiral ganglion cells that are targeted by CI stimulation [116]. Initially, CND was considered to be a contraindication to CI [117]. However, recent studies have suggested that CND patients could benefit from CI [118–120], although the performance was lower than in patients without CND [121]. Therefore, the indication of CI in patients with CND remains controversial. Lu et al. attempted to solve this problem by predicting postoperative CI outcomes using an ML technique. They used the SVM algorithm and achieved 71 % accuracy and 0.71 AUC in predicting hearing rehabilitation, and 93 % accuracy and 0.94 AUC in predicting speech rehabilitation [122].

4.3.2. *Active middle ear implantation*

Active middle ear implants, such as the Vibrant Soundbridge, are used in patients with sensorineural hearing loss or conductive and mixed hearing loss [123]. It is useful for improving patients' speech intelligibility; however, predicting postoperative speech outcome has traditionally been difficult because of the diversity in patients' background conditions. Koyama et al. applied ML to predict postoperative outcomes and reported that RF algorithms achieved better prediction accuracy than classical regression models [124]. The application of ML to determine the outcome of active middle ear implantation is now being researched. However, the number of cases is limited, and the scarcity of cases may be an obstacle to the use of ML in this analysis.

4.3.3. *Tympanoplasty for chronic otitis media*

COM is a common ear disease [125], whose treatment primarily includes medication, e.g., antibiotics. However, patients who cannot be successfully treated with medication or who have conductive hearing loss due to COM complications can be candidates for surgical treatment.

The surgical approach for COM is tympanoplasty, which can improve hearing levels. However, in some cases, hearing recovery cannot be achieved due to the patient's condition. Many researchers have attempted to clarify hearing recovery before surgery; however, this remains unclear. Koyama et al. applied ML techniques to predict postoperative hearing outcomes and reported that an RF algorithm achieved 81.5 % accuracy in binary classification and 63.1 % accuracy in multiclass classification, thereby outperforming the classical numerical classification model [126].

4.3.4. Vestibular schwannoma resection

Vestibular schwannomas are benign tumors arising from Schwann cells. Surgical approaches include middle fossa, retrosigmoid, and stereotactic radiosurgery [127]. In such cases, the hearing outcome or preservation after surgery is one of the main concerns, and ML has been used to predict postoperative hearing levels.

Cha et al. analyzed 52 operative cases of vestibular schwannomas and developed a postoperative hearing level prediction program [128]. They reported that the NN achieved 90 % prediction accuracy. Dixon et al. studied 144 patients with vestibular schwannomas treated using the middle fossa approach, and reported that an RF model exhibited perfect accuracy in predicting hearing preservation [129].

Postoperative dizziness is another concern in this case. Suresh et al. used 1137 cases to predict postoperative dizziness [130]. They reported that logistic regression, linear discriminant analysis, and gradient boosting achieved high performance in binary and multiple classifications with AUCs of 0.89 and 0.86, respectively.

5. Future directions of research

ML has been and is being applied in many areas of otology; however, some challenges remain in advancing these studies, especially regarding the prediction of postoperative outcomes. Till date, study participants have been primarily recruited from single institutes. Therefore, the input data have borne selection biases and the amount of data has been limited. ML tends to enable more robust and accurate prediction as the amount of input data increases—thus, collection of data under fixed criteria or as a fixed data type is desirable. Therefore, high-quality datasets are required to further develop these research areas.

6. Ethical considerations

Despite the great possible benefits of ML advancement, there remain some ethical considerations. ML can easily lose fairness, transparency, security and interpretability, and these factors must be secured in ML when applying [131]. ML system is based on the input data, therefore if input data has been biased, then output or the result have possibility to be biased, which lose fairness and can distract the decision making. ML sometimes produce large different result using similar input and lose reproducibility and transparency. In order to ensure transparency, standard protocols are required in creating ML [132]. Some ML algorithms are difficult to understand how outputs come out and elicit detailed explanation. This loss of transparency in the algorithm is called the “black box” problem [133]. To overcome this problem, ML should be open to the users how the algorithm works and what data it was built on. However, ML is built using the patients’ information, and ensuring transparency could lead to the disclosure of the information and violate their privacy.

In surgical prediction, the results predicted by ML will play an important role in the decision making of both surgeons and patients, but ML prediction has several considerations above. Therefore, considering these ethical aspects, ML prediction should be the assistive role for healthcare provider or surgeon and the patients, not to override or replace them.

7. Conclusions

In this review, we presented a brief overview of the history and concept of ML, discussed some representational algorithms, and surveyed the applications of ML in otology. We then discussed ML-based surgical outcome prediction in some surgical treatments, such as CI, active middle ear implant, tympanoplasty, and vestibular schwannoma resection, while clarifying future directions of research.

In conclusion, with advancements in computer science, ML has been applied in many healthcare studies, including otology, with great success. ML exhibits advantages in terms of prediction performance compared to classical statistical analyses. Appropriate algorithms depend on the research question and often require additional data. High-quality datasets collected under fixed criteria and as fixed data types should be prepared to advance these studies further.

Declaration of competing interest

The author has no conflict of interest to disclose regarding this article.

Acknowledgment

This work was supported by JSPS KAKENHI Grant number JP22K19578.

References

[1] Ramesh AN, Kambhampati C, Monson JRT, Drew PJ. Artificial intelligence in medicine. Ann R Coll Surg Engl 2004;86(5):334–8.
[2] Shortliffe EH, Davis R, Axline SG, Buchanan BG, Green CC, Cohen SN. Computer-based consultations in clinical therapeutics: explanation and rule acquisition capabilities of the MYCIN system. Comput Biomed Res 1975;8(4):303–20.
[3] Greenhill AT, Edmunds BR. A primer of artificial intelligence in medicine. Tech Innov Gastrointest Endosc 2020;22(2):85–9.
[4] Kaul V, Enslin S, Gross SA. History of artificial intelligence in medicine. Gastrointest Endosc 2020;92(4):807–12.
[5] Sutton RS, Barto AG. Reinforcement learning: an introduction. 2nd ed. Massachusetts: The MIT Press; 2018.
[6] Stember JN, Shalu H. Reinforcement learning using Deep*Q* networks and *Q* learning accurately localizes brain tumors on MRI with very small training sets. BMC Med Imaging 2022;22(1):224.
[7] Capurro D, Yetisgen M, van Eaton E, Black R, Tarczy-Hornoch P. Availability of structured and unstructured clinical data for comparative effectiveness research and quality improvement: a multisite assessment. EGEMS 2014;2(1):1079.
[8] Kho AN, et al. Electronic medical records for genetic research: results of the eMERGE consortium. Sci Transl Med 2011;3:79re1.
[9] Abhyankar S, Demner-Fushman D, Callaghan FM, McDonald CJ. Combining structured and unstructured data to identify a cohort of ICU patients who received dialysis. J Am Med Inform Assoc 2014;21(5):801–7.
[10] Wang Y, Ng K, Byrd RJ, Hu J, Ebadollahi S, Daar Z, et al. Early detection of heart failure with varying prediction windows by structured and unstructured data in electronic health records. In: Annual international conference of the IEEE engineering in medicine and biology society (EMBC); 2015. p. 2530–3.
[11] Vest JR, Grannis SJ, Haut DP, Halverson PK, Menachemi N. Using structured and unstructured data to identify patients’ need for services that address the social determinants of health. Int J Med Inform 2017;107:101–6.
[12] Alloghani M, Al-Jumeily D, Mustafina J, Hussain A, Aljaaf AJ. A systematic review on supervised and unsupervised machine learning algorithms for data science. Supervised and unsupervised learning for data science. Springer; 2020. p. 3–21.
[13] Mavrogiorgou A, Kiourtis A, Kleftakis S, Mavrogiorgos K, Zafeiropoulos N, Kyriazis D. A catalogue of machine learning algorithms for healthcare risk predictions. Sensors (Basel) 2022;22(22):8615.
[14] Ferdous M, Debnath J, Chakraborty NR. Machine learning algorithms in healthcare: a literature survey. In: 11th International conference on computing, communication and networking technologies (ICCCNT). IEEE; 2020. p. 1–6.
[15] Jayatilake SMDAC, Ganegoda GU. Involvement of machine learning tools in healthcare decision making. J Healthc Eng 2021;2021:6679512.
[16] Tibshirani R. Regression shrinkage and selection via the Lasso. J R Stat Soc Ser B 1996;58(1):267–88.
[17] Musoro JZ, Zwinderman AH, Puhan MA, ter Riet G, Geskus RB. Validation of prediction models based on Lasso regression with multiply imputed data. BMC Med Res Methodol 2014;14:116.
[18] Fujino Y, Murata H, Mayama C, Asaoka R. Applying “Lasso” regression to predict future visual field progression in glaucoma patients. Invest Ophthalmol Vis Sci 2015;56(4):2334–9.
[19] Zhang X, Xu Q, Yang L, Sun G, Liu G, Lian C, et al. Dynamic risk prediction models for different subtypes of hypertensive disorders in pregnancy. Front Surg 2022;9:1005974.
[20] Ou Q, Jin W, Lin L, Lin D, Chen K, Quan H. LASSO-based machine learning algorithm to predict the incidence of diabetes in different stages. Aging Male 2023;26(1):2205510.
[21] Silverman BW, Jones MC, Fix E, Hodges JL. An important contribution to nonparametric discriminant analysis and density estimation: commentary on Fix and Hodges (1951). Int Stat Rev 1989;57:233–8.
[22] Cover T, Hart P. Nearest neighbor pattern classification. IEEE Trans Inf Theory 1967;13:21–7.

[23] Ehsani R, Drabløs F. Robust distance measures for kNN classification of cancer data. Cancer Inform 2020;19:1–9. 1176935120965542.
[24] Singh D, Singh B. Investigating the impact of data normalization on classification performance. Appl Soft Comput 2020;97:105524.
[25] Liao Y, Vemuri VR. Use of k-nearest neighbor classifier for intrusion detection. Comput Secur 2002;21(5):439–48.
[26] Dong DS, Kandel S. Machine learning for breast cancer diagnosis and classification using hand-crafted features. In: Proceedings of the 12th international conference on soft computing and pattern recognition (SoCPaR 2020) 12. Springer International Publishing; 2021. p. 364–73.
[27] Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. In: Proceedings of the 5th annual ACM workshop on computational learning theory. ACM Press; 1992. p. 144–52.
[28] Drucker H, Burges CJ, Kaufman L, Smola A, Vapnik V. Support vector regression machines. In: Mozer MC, Joradn MI, Petsche T, editors. Advances in neural information processing systems 9. Cambridge: MIT Press; 1997. p. 155–61.
[29] Basudhar A, Dribusch C, Lacaze S, Missoum S. Constrained efficient global optimization with support vector machines. Struct Multidiscip Optim 2012;46: 201–21.
[30] Sweilam NH, Tharwat AA, Abdel Moniem NK. Support vector machine for diagnosis cancer disease: a comparative study. Egypt Inform J 2010;11(2):81–92.
[31] Zhang F, Petersen M, Johnson L, Hall J, O'Bryant SE. Recursive support vector machine biomarker selection for Alzheimer's disease. J Alzheimers Dis 2021;79 (4):1691–700.
[32] Maltarollo VG, Kronenberger T, Espinoza GZ, Oliveira PR, Honorio KM. Advances with support vector machines for novel drug discovery. Expert Opin Drug Discov 2019;14(1):23–33.
[33] Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol 2017;2(4):230–43.
[34] Doupe P, Faghmous J, Basu S. Machine learning for health services researchers. Value Health 2019;22(7):808–15.
[35] Efron B, Hastie T. Computer age statistical inference. Cambridge University Press; 2016.
[36] Morgan Kaufmann. Data Preprocessing. In: Han J, Kamber M, Pei J, editors. Data mining concepts and techniques. 3rd ed. Elsevier; 2011. p. 83–124.
[37] Breiman L. Random forests. Mach Learn 2001;45:5–32.
[38] Rigatti SJ. Random forest. J Insur Med 2017;47(1):31–9.
[39] Qi Y. Random forest for bioinformatics. In: Zhang C, Ma Y, editors. Ensemble machine learning: methods and applications. New York: Springer; 2012. p. 307–23.
[40] Wu Y, Fang Y. Stroke prediction with machine learning methods among older Chinese. Int J Environ Res Public Health 2020;17(6):1828.
[41] Song W, Zhou X, Duan Q, Wang Q, Li Y, Li A, et al. Using random forest algorithm for glomerular and tubular injury diagnosis. Front Med (Lausanne) 2022;9: 911737.
[42] Sreejith S, Khanna Nehemiah H, Kannan A. Clinical data classification using an enhanced SMOTE and chaotic evolutionary feature selection. Comput Biol Med 2020;126:103991.
[43] Blagus R, Lusa L. SMOTE for high-dimensional class-imbalanced data. BMC Bioinform 2013;14:106.
[44] Yuan KC, Tsai LW, Lee KH, Cheng YW, Hsu SC, Lo YS, et al. The development an artificial intelligence algorithm for early sepsis diagnosis in the intensive care unit. Int J Med Inform 2020;141:104176.
[45] Ridgeway G. The state of boosting. Comput Sci Stat 1999;31:172–81.
[46] Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining-KDD 2016, San Francisco; 2016. p. 785–94.
[47] Zhang YW, Feng B, Chen Y, Liao W, Guo C. Fault diagnosis method for oil-immersed transformer based on XGBoost optimized by genetic algorithm. Electric Power Autom Equip 2021;41(2):200–6.
[48] Nobre J, Neves RF. Combining principal component analysis, discrete wavelet transform and XGBoost to trade in the financial markets. Expert Syst Appl 2019; 125:181–94.
[49] Wang Y, Guo Y. Forecasting method of stock market volatility in time series data based on mixed model of ARIMA and XGBoost. China Commun 2020;17(3): 205–21.
[50] Ben JS, Stef N, Carmona P. Bankruptcy prediction using the XGBoost algorithm and variable importance feature engineering. Comput Econ 2023;61(2):715–41.
[51] Qu WL, Li YQ, Zhou L. Application of XGBoost algorithm in diabetic blood glucose prediction. J Jilin Norm Univ Nat Sci Ed 2019;40(4):118–25.
[52] Romeo L, Frontoni E. A unified hierarchical XGBoost model for classifying priorities for COVID-19 vaccination campaign. Pattern Recognit 2022;121: 108197.
[53] Liu Y, Qiao M. Heart disease prediction based on clustering and XGBoost algorithm. Comput Syst Appl 2019;28(1):228–32.
[54] Yuan LX, Gu YJ, Zhao DP. Research on abnormal user detection technology in social network based on XGBoost method. Appl Res Comput 2020;37(3):814–7.
[55] Zhang P, Jia Y, Shang Y. Research and application of XGBoost in imbalanced data. Int J Distrib Sens Netw 2022;18(6):1–10. 15501329221106935.
[56] Gurney K. An introduction to neural networks. 1st ed. CRC Press; 2017.
[57] Abiodun OI, Jantan A, Omolara AE, Dada KV, Mohamed NA, Arshad H. State-of-the-art in artificial neural network applications: a survey. Heliyon 2018;4(11): e00938.
[58] Choi RY, Coyner AS, Kalpathy-Cramer J, Chiang MF, Campbell JP. Introduction to machine learning, neural networks, and deep learning. Transl Vis Sci Technol 2020;9(2):14.
[59] James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning: with applications in R. New York: Springer; 2013.
[60] Hastie T, Tibshirani R, Friedman JH. The elements of statistical learning: data mining, inference, and prediction. 2nd ed. New York, NY: Springer; 2009.
[61] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436–44.
[62] Goodfellow I, Bengio YCA, Courville A. Deep learning. 1st ed. The MIT Press; 2016.
[63] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE 1998;86(11):2278–324.
[64] Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. Science 2006;313:504–7.
[65] Kaehler A, Bradski GR. Learning OpenCV 3: computer vision in C++ with the OpenCV library. Sebastopol, CA: O'Reilly Media; 2017.
[66] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM 2017;60(6):84–90.
[67] Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition; arXiv. 2014;1409.1556.
[68] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2015. p. 1–9.
[69] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 770–8.
[70] Ronneberger O., Fischer P., Brox T. U-net: convolutional networks for biomedical image segmentation. In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Springer International Publishing. Proceedings, part III 18 p. 234–41.
[71] Chan HP, Samala RK, Hadjiiski LM, Zhou C. Deep learning in medical image analysis. Adv Exp Med Biol 2020;1213:3–21.
[72] Egger J, Gsaxner C, Pepe A, Pomykala KL, Jonske F, Kurz M, et al. Medical deep learning—A systematic meta-review. Comput Methods Programs Biomed 2022; 221:106874.
[73] Lee JY, Choi SH, Chung JW. Automated classification of the tympanic membrane using a convolutional neural network. Appl Sci 2019;9(9):1827.
[74] Cha D, Pae C, Seong SB, Choi JY, Park HJ. Automated diagnosis of ear disease using ensemble deep learning with a big otoendoscopy image database. EBioMedicine 2019;45:606–14.
[75] Wu Z, Lin Z, Li L, Pan H, Chen G, Fu Y, et al. Deep learning for classification of pediatric otitis media. Laryngoscope 2021;131(7):E2344–51.
[76] Byun H, Lee SH, Kim TH, Oh J, Chung JH. Feasibility of the machine learning network to diagnose tympanic membrane lesions without coding experience. J Pers Med 2022;12(11):1855.
[77] Zeng J, Kang W, Chen S, Lin Y, Deng W, Wang Y, et al. A deep learning approach to predict conductive hearing loss in patients with otitis media with effusion using otoscopic images. JAMA Otolaryngol Head Neck Surg 2022;148(7):612–20.
[78] Carey GE, Jacobson CE, Warburton AN, Biddle E, Mannarelli G, Wilson M, et al. Machine learning for vestibular schwannoma diagnosis using audiometrie data alone. Otol Neurotol 2022;43(5):e530–4.
[79] Wang YM, Li Y, Cheng YS, He ZY, Yang JM, Xu JH, et al. Deep learning in automated region proposal and diagnosis of chronic otitis media based on computed tomography. Ear Hear 2020;41(3):669–77.
[80] Eroğlu O, Eroğlu Y, Yıldırım M, Karlıdag T, Çınar A, Akyiğit A, et al. Is it useful to use computerized tomography image-based artificial intelligence modelling in the differential diagnosis of chronic otitis media with and without cholesteatoma? Am J Otolaryngol 2022;43(3):103395.
[81] Ayral M, Türk Ö, Can Ş, Esen D, Topçu I, Akil F, et al. How advantageous is it to use computed tomography image-based artificial intelligence modelling in the differential diagnosis of chronic otitis media with and without cholesteatoma? Eur Rev Med Pharmacol Sci 2023;27(1):215–23.
[82] Su R, Song J, Wang Z, Mao S, Mao Y, Wu X, et al. Application of high resolution computed tomography image assisted classification model of middle ear diseases based on 3D-convolutional neural network. Zhong Nan Da Xue Xue Bao Yi Xue Ban 2022;47(8):1037–48.
[83] Takahashi M, Noda K, Yoshida K, Tsuchida K, Yui R, Nakazawa T, et al. Preoperative prediction by artificial intelligence for mastoid extension in pars flaccida cholesteatoma using temporal bone high-resolution computed tomography: a retrospective study. PLoS One 2022;17(10).
[84] Fujima N, Andreu-Arasa VC, Onoue K, Weber PC, Hubbell RD, Setty BN, et al. Utility of deep learning for the diagnosis of otosclerosis on temporal bone CT. Eur Radiol 2021;31(7):5206–11.
[85] Ogawa M, Kisohara M, Yamamoto T, Shibata S, Ojio Y, Mochizuki K, et al. Utility of unsupervised deep learning using a 3D variational autoencoder in detecting inner ear abnormalities on CT images. Comput Biol Med 2022;147:105683.
[86] Profant O, Bureš Z, Balogová Z, Betka J, Fík Z, Chovanec M, et al. Decision making on vestibular schwannoma treatment: predictions based on machine-learning analysis. Sci Reports 2021;11(1):18376.
[87] Gadot R, Anand A, Lovin BD, Sweeney AD, Patel AJ. Predicting surgical decision-making in vestibular schwannoma using tree-based machine learning. Neurosurg Focus 2022;52(4):E8.
[88] Tang OY, Bajaj AI, Zhao K, Rivera Perla KM, Ying YM, Jyung RW, et al. Association of patient frailty with vestibular schwannoma resection outcomes and machine learning development of a vestibular schwannoma risk stratification score. Neurosurgery 2022;91(2):312–21.
[89] McRackan TR, Bauschard M, Hatch JL, Franko-Tobin E, Droghini HR, Nguyen SA, Dubno JR. Meta-analysis of quality-of-life improvement after cochlear

implantation and associations with speech recognition abilities. Laryngoscope 2018;128(4):982–90.
[90] Gundacker G, Trales DE, Stefanescu HE. Quality of life and audiological benefits in pediatric cochlear implant users in romania: systematic review and cohort study. J Pers Med 2023;13(11):1610.
[91] Morettin M, Santos MJ, Stefanini MR, Antonio Fde L, Bevilacqua MC, Cardoso MR. Measures of quality of life in children with cochlear implant: systematic review. Braz J Otorhinolaryngol 2013;79(3):375–81.
[92] Gaylor JM, Raman G, Chung M, Lee J, Rao M, Lau J, Poe DS. Cochlear implantation in adults: a systematic review and meta-analysis. JAMA Otolaryngol Head Neck Surg 2013;139(3):265–72.
[93] Labassi S, Beliaeff M, Péan V, Van de Heyning P. The vibrant Soundbridge® middle ear implant: a historical overview. Cochlear Implants Int 2017;18(6): 314–23.
[94] Bruchhage KL, Leichtle A, Schönweiler R, Todt I, Baumgartner WD, Frenzel H, Wollenberg B. Systematic review to evaluate the safety, efficacy and economical outcomes of the Vibrant Soundbridge for the treatment of sensorineural hearing loss. Eur Arch Otorhinolaryngol 2017;274(4):1797–806.
[95] Lucidi D, Cantaffa C, Nocini R, Martone A, Alicandri-Ciufelli M, Marchioni D, Presutti L, Molinari G. Quality of life after surgical treatment for chronic otitis media: a systematic review of the literature. J Pers Med 2022;12(12):1959.
[96] Pap I, Kovács M, Bölcsföldi B, Szakács Z, Gerlinger I, Imreh B, Csongor A, Warta V, Szanyi I. Quality-of-life outcomes with endoscopic and microscopic type I tympanoplasty-a prospective cohort study. Eur Arch Otorhinolaryngol 2023;280 (10):4401–8.
[97] Pontillo V, Damiani M, Harib A, Sammali M, Graziano G, Quaranta N. Quality of life after cholesteatoma surgery: comparison between surgical techniques. Acta Otorhinolaryngol Ital 2022;42(3):293–9.
[98] Lucidi D, De Corso E, Paludetti G, Sergi B. Quality of life and functional results in canal wall down vs canal wall up mastoidectomy. Acta Otorhinolaryngol Ital 2019;39(1):53–60.
[99] Ardıç FN, Mengi E, Tümkaya F, Kara CO, Bir F. Correlation between surgical outcome and stage of acquired middle ear cholesteatoma: revalidation of the EAONO/JOS staging system. J Int Adv Otol 2020;16(1):34–9.
[100] Jung DJ, Lee HJ, Hong JS, Kim DG, Mun JY, Bae JW, Yoo MH, Lee KY. Prediction of hearing outcomes in chronic otitis media patients underwent tympanoplasty using ossiculoplasty outcome parameter staging or middle ear risk indices. PLoS One 2021;16(7):e0252812.
[101] Leung J, Wang NY, Yeagle JD, Chinnici J, Bowditch S, Francis HW, Niparko JK. Predictive models for cochlear implantation in elderly candidates. Arch Otolaryngol Head Neck Surg 2005;131(12):1049–54.
[102] van Dijk JE, van Olphen AF, Langereis MC, Mens LH, Brokx JP, Smoorenburg GF. Predictors of cochlear implant performance. Audiology 1999;38(2):109–16.
[103] Anderson CA, Wiggins IM, Kitterick PT, Hartley DEH. Pre-operative brain imaging using functional near-infrared spectroscopy helps predict cochlear implant outcome in deaf adults. J Assoc Res Otolaryngol 2019;20(5):511–28.
[104] Gantz BJ, Woodworth GG, Knutson JF, Abbas PJ, Tyler RS. Multivariate predictors of audiological success with multichannel cochlear implants. Ann Otol Rhinol Laryngol 1993;102(12):909–16.
[105] Tihanyi A, Speck I, Wolff K, Arnold P, Aschendorff A, Arndt S. Hearing outcome after tympanoplasty type III. Eur Arch Otorhinolaryngol 2022;279(9):4353–62.
[106] Cacco T, Africano S, Gaglio G, Carmisciano L, Piccirillo E, Castello E, Peretti G. Correlation between peri-operative complication in middle ear cholesteatoma surgery using STAMCO, ChOLE, and SAMEO-ATO classifications. Eur Arch Otorhinolaryngol 2022;279(2):619–26.
[107] Wang X, Guo J, Liu W, Chen M, Shao J, Zhang X, Ma N, Li Y, Peng Y, Zhang J. Comparison of the EAONO/JOS, STAMCO and ChOLE cholesteatoma staging systems in the prognostic evaluation of acquired middle ear cholesteatoma in children. Eur Arch Otorhinolaryngol 2022;279(12):5583–90.
[108] Goldenholz DM. Can machine learning solve this one? Clinical pitfalls in surgical outcome prediction. Epilepsia 2023;64(5):1190–4.
[109] Zanier O, Zoli M, Staartjes VE, Guaraldi F, Asioli S, Rustici A, et al. Machine learning-based clinical outcome prediction in surgery for acromegaly. Endocrine 2022;75(2):508–15.
[110] Senders JT, Staples PC, Karhade AV, Zaki MM, Gormley WB, Broekman MLD, Smith TR, Arnaout O. Machine learning and neurosurgical outcome prediction: a systematic review. World Neurosurg 2018;109:476–86. e1.
[111] Deo RC. Machine learning in medicine. Circulation 2015;132(20):1920–30.
[112] Bur AM, Shew M, New J. Artificial intelligence for the otolaryngologist: a state of the art review. Otolaryngol Head Neck Surg 2019;160(4):603–11.
[113] Tan L, Holland SK, Deshpande AK, Chen Y, Choo DI, Lu LJ. A semi-supervised Support vector machine model for predicting the language outcomes following cochlear implantation based on pre-implant brain fMRI imaging. Brain Behav 2015;5(12):e00391.
[114] Crowson MG, Dixon P, Mahmood R, Lee JW, Shipp D, Le T, et al. Predicting postoperative cochlear implant performance using supervised machine learning. Otol Neurotol 2020;41(8):e1013–23.
[115] Zeitler DM, Buchlak QD, Ramasundara S, Farrokhi F, Esmaili N. Predicting acoustic hearing preservation following cochlear implant surgery using machine learning. Laryngoscope 2024;134(2):926–36.
[116] Adunka OF, Jewells V, Buchman CA. Value of computed tomography in the evaluation of children with cochlear nerve deficiency. Otol Neurotol 2007;28(5): 597–604.
[117] Shelton C, Luxford WM, Tonokawa LL, Lo WWM, House WF. The narrow internal auditory canal in children: a contraindication to cochlear implants. Otolaryngol Head Neck Surg 1989;100(3):227–31.
[118] Kang WS, Lee JH, Lee HN, Lee KS. Cochlear implantations in young children with cochlear nerve deficiency diagnosed by MRI. Otolaryngol Head Neck Surg 2010; 143(1):101–8.
[119] Zhang Z, Li Y, Hu L, Wang Z, Huang Q, Wu H. Cochlear implantation in children with cochlear nerve deficiency: a report of nine cases. Int J Pediatr Otorhinolaryngol 2012;76(8):1188–95.
[120] Wei X, Lu S, Chen B, Chen J, Zhang L, Li Y, et al. Cochlear implantation programming characteristics and outcomes of cochlear nerve deficiency. Eur Arch Otorhinolaryngol 2023;280(10):4409–18.
[121] Yousef M, Mesallam TA, Garadat SN, Almasaad A, Alzhrani F, Alsanosi A, et al. Audiologic outcome of cochlear implantation in children with cochlear nerve deficiency. Otol Neurotol 2021;42(1):38–46.
[122] Lu S, Xie J, Wei X, Kong Y, Chen B, Chen J, et al. Machine learning-based prediction of the outcomes of cochlear implantation in patients with cochlear nerve deficiency and normal cochlea: a 2-year follow-up of 70 children. Front Neurosci 2022;16:895560.
[123] Maw J. The Vibrant Soundbridge: a Global Overview. Otolaryngol Clin North Am 2019;52(2):285–95.
[124] Koyama H, Mori A, Nagatomi D, Fujita T, Saito K, Osaki Y, et al. Machine learning technique reveals prognostic factors of vibrant soundbridge for conductive or mixed hearing loss patients. Otol Neurotol 2021;42(9):e1286–e92.
[125] Auinger P, Lanphear BP, Kalkwarf HJ, Mansour ME. Trends in otitis media among children in the United States. Pediatrics 2003;112(3 Pt 1):514–20.
[126] Koyama H, Kashio A, Uranaka T, Matsumoto Y, Yamasoba T. Application of machine learning to predict hearing outcomes of tympanoplasty. Laryngoscope 2023;133(9):2371–8.
[127] Ansari SF, Terry C, Cohen-Gadol AA. Surgery for vestibular schwannomas: a systematic review of complications by approach. Neurosurg Focus 2012;33(3): E14.
[128] Cha D, Shin SH, Kim SH, Choi JY, Moon IS. Machine learning approach for prediction of hearing preservation in vestibular schwannoma surgery. Sci Rep 2020;10(1):7136.
[129] Dixon PR, Wojdyla L, Lee J, Moshtaghi O, Claussen AD, Khan UA, et al. Machine learning to predict hearing preservation after middle cranial fossa approach for sporadic vestibular schwannomas. Otol Neurotol 2022;43(9):1072–7.
[130] Suresh K, Elkahwagi MA, Garcia A, Naples JG, Corrales CE, Crowson MG. Development of a predictive model for persistent dizziness following vestibular schwannoma surgery. Laryngoscope 2023;133(12):3534–9.
[131] Nwosu OI, Crowson MG, Rameau A. Artificial intelligence governance and otolaryngology-head and neck surgery. Laryngoscope 2023;133(11):2868–70.
[132] Crowson MG, Rameau A. Standardizing machine learning manuscript reporting in otolaryngology-head & neck surgery. Laryngoscope 2022;132(9):1698–700.
[133] Anderson M, Anderson SL. How should AI be developed, validated, and implemented in patient care? AMA J Ethics 2019;21(2):E125–30.