

Literature Review**Image-based artificial intelligence for otitis media effusion diagnosis:
a systematic review and meta-analysis**

Dwi Khoirriyani, Nova Nasikhatussoray, Muyassaroh
Hermina Banyumanik Hospital, Semarang, Indonesia

ABSTRACT

Background: Otitis media with effusion (OME) is one of the most common ear conditions encountered in primary healthcare, particularly in facilities without supporting audiometry or tympanometry. OME is often misdiagnosed or delayed in treatment because it requires specialised skills to interpret otoscopic findings. Artificial Intelligence (AI) can help to simplify the diagnosis of OME so that appropriate management can be provided immediately. **Purpose:** To determine the performance of AI based on otoscope results in the diagnosis of OME. **Method:** A systematic review search based on the PRISMA flowchart was conducted in three databases (PubMed, Cochrane Library, ScienceDirect) until July 2024. Data extraction was performed to assess the sensitivity and specificity of AI for the diagnosis of OME. The risk of bias was assessed using QUADAS-2. All data were analyzed using Review Manager 5.4 and MetaDTA 2.0. **Result:** Twelve studies with 20,452 otoscope results were included in this review. The meta-analysis showed that AI based on otoscope results had a high sensitivity of 0.93 (95% CI; 0.87-0.96), and a high specificity of 0.96 (95% CI; 0.94-0.98). In pediatric patients, this otoscope-based AI was also shown to diagnose OME with high sensitivity and specificity of 0.93 (95% CI; 0.64-0.99); 0.94 (95% CI; 0.86-0.97). **Conclusion:** AI demonstrates high sensitivity and specificity performance for the diagnosis of OME in both adult and pediatric patients. AI shows promising potential to help identify OME in primary care settings before referral for further management, and to avoid unnecessary antibiotic use. Further research is needed.

Keywords: artificial intelligence, otitis media with effusion, otoscope

ABSTRAK

Latar belakang: Otitis Media Efusi (OME) adalah salah satu penyakit telinga yang banyak dijumpai di fasilitas kesehatan layanan primer, terutama pada fasilitas kesehatan yang tidak memiliki peralatan audiometri atau timpanometri. OME sering salah terdiagnosis atau terlambat dalam penanganannya karena membutuhkan keterampilan khusus dalam interpretasi hasil otoskop. Artificial Intelligence (AI) dapat membantu untuk mempermudah penegakkan diagnosis OME sehingga tatalaksana yang tepat dapat segera diberikan. **Tujuan:** Untuk mengetahui performa dari AI berdasarkan hasil otoskop dalam diagnosis OME. **Metode:** Pencarian tinjauan sistematis berdasarkan diagram alur PRISMA dilakukan pada tiga basis data (PubMed, Cochrane Library, ScienceDirect) sampai Juli 2024. Ekstraksi data dilakukan untuk menilai sensitivitas dan spesifisitas dari AI untuk diagnosis OME. Penilaian risiko bias dilakukan dengan QUADAS-2. Semua data dianalisis dengan Review Manager 5.4 and MetaDTA 2.0. **Hasil:** Dua belas studi dengan 20,452 hasil otoskop dimasukkan dalam penelitian ini. Metanalisis menunjukkan AI berbasis hasil otoskop memiliki sensitivitas tinggi sebesar 0,93 (95% CI; 0,87-0,96) dan spesifisitas tinggi sebesar 0,96 (95% CI; 0,94-0,98). Pada pasien anak, AI berbasis otoskop ini juga terbukti dapat mendiagnosis OME dengan sensitivitas dan spesifisitas tinggi sebesar 0,93 (95% CI; 0,64-0,99); 0,94 (95% CI; 0,86-0,97). **Kesimpulan:** AI menunjukkan performa sensitivitas dan spesifisitas yang tinggi untuk diagnosis OME pada pasien dewasa maupun anak. AI menunjukkan potensi yang menjanjikan untuk membantu deteksi OME di fasilitas layanan primer sebelum dirujuk untuk tatalaksana lebih lanjut, dan menghindari penggunaan antibiotik yang tidak diperlukan. Penelitian lebih lanjut masih perlu dilakukan.

Kata kunci: artificial intelligence, *otitis media efusi*, *otoskop*

Correspondence address: Dwi Khoirriyani. Hermina Banyumanik Hospital, Semarang, Indonesia.
Email: dwikhoirriyani@yahoo.com

INTRODUCTION

Otitis media with effusion (OME) is a common primary care condition without typical symptoms. OME is most commonly found in pediatric patients with complaints of hearing loss, but in some cases, it occurs in adults. The diagnosis of OME is essentially clinical, based on otoscopy and supported by audiometry/tympanometry results. In healthcare facilities without audiometry or tympanometry equipment, it is vital to have expertise in assessing OME findings on otoscopy.¹ Otitis media is prevalent worldwide, with approximately 80% of children experiencing otitis media during their lifetime, of which 80-90% will experience OME.² Many studies have been conducted to determine the diagnosis and management of OME. If OME is not properly diagnosed and managed, it can lead to serious conditions such as hearing loss, impaired cognitive function, unnecessary surgery, overuse of antibiotics, and other complications. Nearly 80% of OME cases are treated with antibiotics that actually unwarranted, leading to antibiotic resistance. In addition, accurate diagnosis is necessary to prevent side effects and manage them appropriately.³

Diagnosing OME is challenging as it often manifests without acute symptoms such as pain or fever. Therefore, there is a need for specialized facilities and training for medical personnel who may encounter cases of OME. Training in diagnosis is available to general practitioners and pediatricians who frequently encounter these cases. It includes using an otoscope to detect OME (70% success rate).⁴ Failing to recognize OME leads to misdiagnosing and mismanaging OME. If OME is not treated, complications

like mastoiditis and hearing loss can occur, reducing quality of life.¹ The need for new approaches to diagnose accurately is urgent.

Over the past 5 years, artificial intelligence (AI) has been increasingly used in medicine for diagnosis. AI processes human-created data with multiple algorithms, producing highly accurate predictions.⁵ Conventional otoscopic imaging is highly dependent on the skill of the examiner to diagnose OME. Thus, the integration of deep learning algorithms into oto-endoscopic imaging is highly relevant.⁶ The use of AI in medicine has grown significantly, especially in studies involving endoscopic imaging, due to advances in computer science. The convenience of AI through automated diagnosis of OME cases is still not applicable to actual conditions due to the uncertainties associated with deep learning, which is a major limitation.

This study was conducted to evaluate the performance of AI based on otoscope results in the diagnosis of OME. Based on the study results, we discuss the improvement measures in this study and suggestions for future research.

METHOD

A systematic review was conducted using the PRISMA flowchart. Ethical approval and patient consent were not required for this study.

A systematic review of databases (PubMed, ScienceDirect, and Cochrane) was conducted from original articles (uploaded until July 2024) related to the use of AI to diagnose OME using otoscopic images. The search was limited to patients who were alive,

and reviews were uploaded in English. The medical and non-medical keywords used in the search were “artificial intelligence” AND “machine learning” AND “deep learning” AND “neural network” AND “otitis media” AND “middle ear”.

The inclusion criteria were full-text articles describing the development of autonomous and supervised OME image classification algorithms using the above AI. The exclusion criteria were articles that did not use AI approaches, used imaging modalities other than otoscopy, or reviews, or editorial articles.

The quality of the included studies was assessed using the Quality Assessment and Diagnostic Accuracy Studies (QUADAS2) tool. The QUADAS2 assessment tool consists of patient selection, index test, reference standard, flow and timing. The authors considered whether the subjects used to train the algorithm were applicable to the target population.

For the Index Test, the authors considered whether the algorithm’s classifying categories were applicable to the intended use. The authors considered whether the methods used were in line with clinical practice.

A systematic review and meta-analysis were done with the primary outcome being performance in diagnosing OME. The secondary outcome was performance in diagnosing OME in pediatric patients

All data were analyzed using Review Manager 5.4 and MetaDTA 2.0.

RESULT

Study selection

The search strategy yielded a total of 154 relevant articles and abstracts. After full-text review, 12 articles met the inclusion criteria.

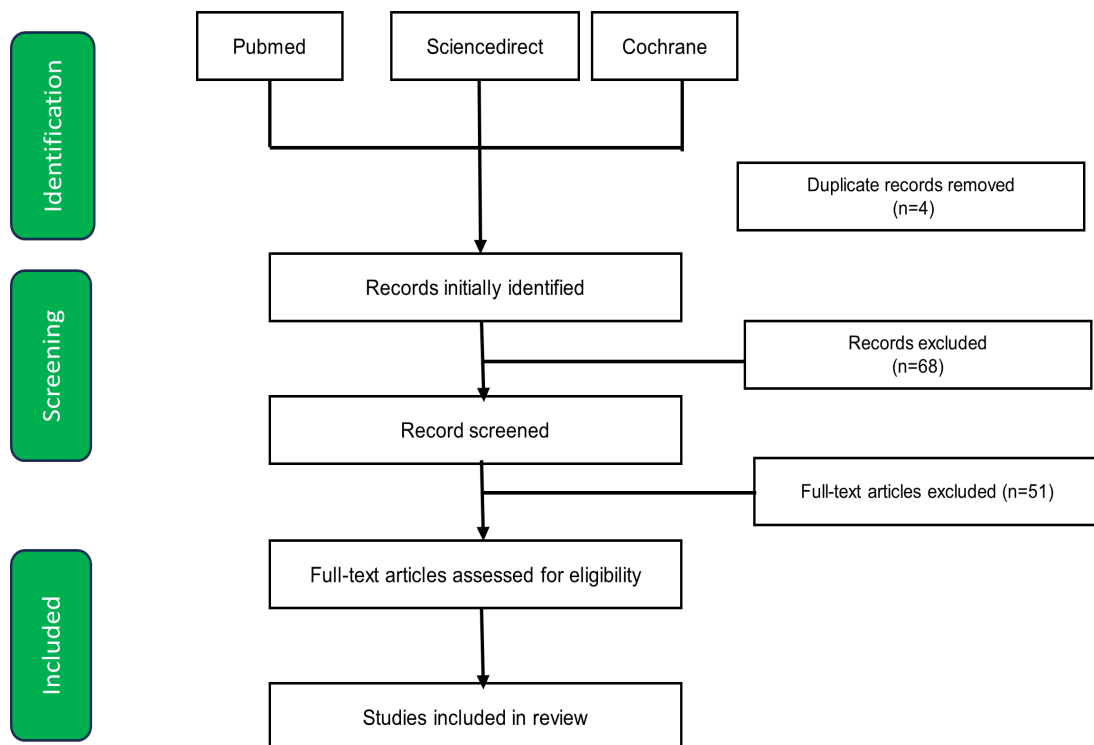


Figure 1. PRISMA flow chart

Study characteristics

All studies used otoscope image data as an input variable. Table 1 summarized the characteristics of the included studies. Nine articles obtained otoscope results in an ENT

polyclinic, one article obtained otoscope results in the operating room just prior to myringotomy, and two articles mentioned only outpatients in a hospital.

Table 1. Characteristics

Article	Country	AI model	Diagnosis by	Inclusion (Age)	Output	N	Prevalence	Performance (Sen/Spec)
Byun et al., 2021	Korea	ResNet/ NA	3 otolaryngologists	NA	Normal, OME, COM, Cholesteatoma	71	23%	0.94/1.00
Choi et al., 2022	Korea	EfficientNet-B/ Pytorch	2 otolaryngologists	NA	OME, COM, normal	6630	24%	0.92/0.99
Khan et al., 2020	Korea	DenseNet161/ PyTorch	2 otolaryngologists	NA	OME, Normal, COM	2484	36%	0.93/0.97
Viscario et al., 2022	Chili	VGG-16 neural network/ Python	Otolaryngologist	NA	OME, COM, Normal, Earwax plug	2000	25%	0.96/0.88
Dubois et al., 2024	France	Inception-v2/ Python	1 otolaryngologist	5-18 years old	Normal, earwax plug, Perforation of TM, OME, CCR, External otitis, Tympanosclerosis, AOM, Osteoma, Corpus alienium, Graft TM	326	3%	0.58/0.98
Cai et al., 2021	China	ResNet 50/ Python 3.6 in keras	Otolaryngologist	NA	OME, CSOMa, CSOMi	6066	45%	0.93/0.98
Tran et al., 2018	Vietnam	MTISRC	Otolaryngologist	6 months-12 years old	OME, AOM	214	53%	0.93/0.89
Byun et al., 2022	Korea	Teachable machine	Otolaryngologist	NA	Normal, OME, COM, Cholesteatoma	80	25%	0.70/0.92
Crowson et al., 2020	Massachusetts	ResNet-34/ Python	Otolaryngologist who performed the myringotomy	<18 years old	OME, Normal	69	61%	0.83/0.81

Article	Country	AI model	Diagnosis by	Inclusion (Age)	Output	N	Prevalence	Performance (Sen/Spes)
Chen et al., 2022	Taiwan	NASNetMobile (TL), MobileNetV2(TL), DenseNet201(TL)/ Python	3 otolaryngologists	NA	Normal, AOM, Acute myringitis, CSOM, OME, Perforation of TM, Earwax plug, Tube placement, Retraction, Otitomycosis	2095	10%	0.98/0.99
Shim et al., 2023	Korea	ResNet50, DenseNet201, InceptionV3, Inception-ResNetV2/ Python	Otolaryngologist	<15 years old	Normal, OME	230	51%	0.87/0.93
Noda et al., 2024	Japan	GPT-4V	Otolaryngologist	NA	AOM, COM, Cholesteatoma, OME	187	26%	1.00/0.95

Risk of bias

Figures 2 and 3 showed the risk of bias and applicability concerns in all studies using the QUADAS-2 method. All studies had a low risk of bias and applicability concerns in

patient selection and the index test. One study had a low risk of bias and concerns about applicability in all domains. One study had a high risk of bias in flow and timing.

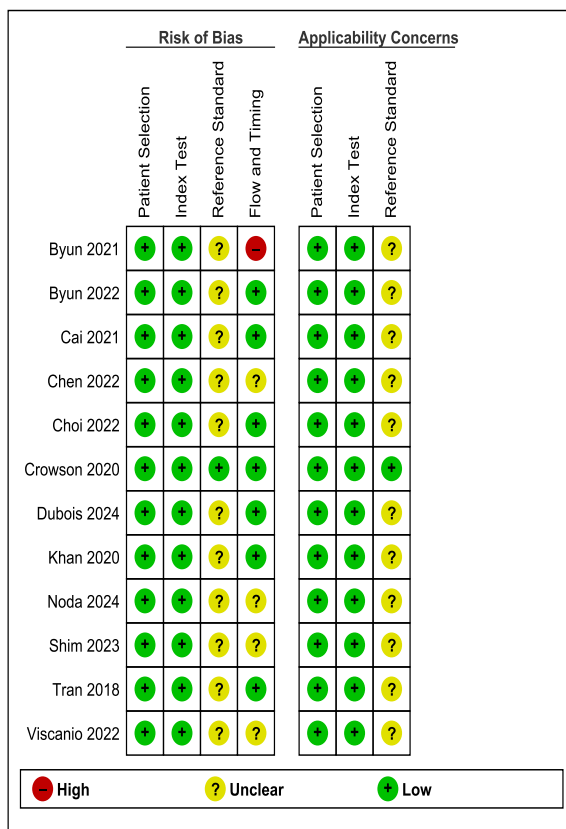


Figure 2. Traffic light plot from QUADAS-2

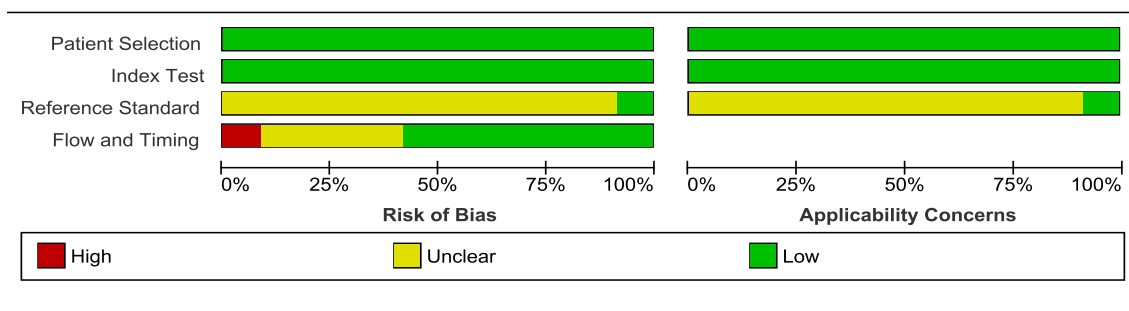


Figure 3. Summary plot from QUADAS-2

OME versus non-OME

Five studies showed high rates of sensitivity and specificity.^{7-9,12,16} Byun et al.⁷ showed that using AI models with ResNet18+Shuffle resulted in higher sensitivity and specificity compared to other models. The use of EfficientNet-B in Choi et al.⁸ showed sensitivity and specificity rates of 0.92 [95% CI; 0.90-0.93] and 0.99 [95% CI; 0.98-0.99], respectively. Khan et al.⁹ showed that using an AI model in the form

of DenseNet161 resulted in higher sensitivity and specificity rates than other AI models. ResNet 50 is an AI model used in the study by Cai et al.¹² with sensitivity and specificity values of (0.98 [0.95,0.99], 0.99[0.99,1.00]). Whereas in the research of Chen et al.¹⁶ used three AI models, i.e. NASNetMobile (TL), MobileNetV2 (TL), DenseNet201 (TL), to analyze OME images and other tympanic membrane conditions.

Table 2. Forest plot of the sensitivity and specificity performance of the AI algorithm in detecting OME

Study	TP	FN	FP	TN	Sensitivity (95% CI)	Specificity (95% CI)
Byun et al., 2021	16	1	0	54	0.94 [0.73–0.99]	1.00 [0.93–1.00]
Choi et al., 2022	1498	132	66	4934	0.92 [0.90–0.93]	0.99 [0.98–0.99]
Choi et al., 2022	831	66	45	1542	0.93 [0.91–0.94]	0.97 [0.96–0.98]
Viscaino et al., 2022	480	20	179	1321	0.96 [0.94–0.97]	0.88 [0.86–0.90]
Dubois et al., 2024	7	5	6	308	0.58 [0.32–0.81]	0.98 [0.96–0.99]
Tran et al., 2018	105	9	11	90	0.93 [0.87–0.96]	0.89 [0.82–0.94]
Byun et al., 2022	14	6	5	55	0.70 [0.48–0.85]	0.92 [0.82–0.96]
Cai et al., 2021	2542	186	71	3267	0.93 [0.92–0.94]	0.98 [0.97–0.98]
Crowson et al., 2021	35	7	5	22	0.83 [0.69–0.92]	0.81 [0.63–0.92]
Chen et al., 2022	201	4	13	1877	0.98 [0.95–0.99]	0.99 [0.99–1.00]
Noda et al., 2024	42	6	10	129	0.88 [0.75–0.94]	0.93 [0.87–0.96]
Shim et al., 2023	118	0	5	107	1.00 [0.97–1.00]	0.96 [0.90–0.98]
					0.92 [0.86, 0.95]	0.97 [0.93, 0.98]

*TP=True Positive; FN=False Negative; FP=False Positive; TN=True Negative

OME and non-OME in pediatric patients

Four studies used pediatric patient samples as subjects in their studies.^{11,13,15,17} Dubois et al.¹¹ showed a sensitivity of 0.58 [95% CI; 0.32-0.81] and a specificity of 0.98 [95% CI; 0.96-0.99]. Another study by Tran et al.¹³ found sensitivity and specificity of 0.93 [95% CI; 0.87-0.96] and 0.89 [95% CI;

0.82-0.94], respectively. In contrast, the study by Crowson et al.¹⁵ showed sensitivity and specificity rates of 0.83 [95% CI; 0.69-0.92] and 0.81 [95% CI; 0.63-0.92].

The study by Shim et al.¹⁷ showed a very high sensitivity of 1.00 [95% CI; 0.97-1.00] with a specificity of 0.96 [95% CI; 0.90-0.98].

Table 3. Forest plot of the sensitivity and specificity performance of the AI algorithm in detecting OME in pediatric patients

Study	TP	FN	FP	TN	Sensitivity (95% CI)	Specificity (95% CI)
Dubois et al., 2024	7	5	66	308	0.58 [0.32, 0.81]	0.98 [0.96, 0.99]
Tran et al., 2018	105	8	11	90	0.93 [0.87, 0.96]	0.89 [0.82, 0.94]
Crowson et al., 2021	35	7	5	22	0.83 [0.69, 0.92]	0.81 [0.63, 0.92]
Shim et al., 2023	118	0	5	107	1.00 [0.97, 1.00]	0.96 [0.90, 0.98]
					0.98 [0.64, 0.99]	0.94 [0.86, 0.97]

*TP=True Positive; FN=False Negative; FP=False Positive; TN=True Negative

DISCUSSION

This meta-analysis evaluated the use of apps with AI technology to detect and diagnose OME using otoscope images. The 12 studies showed that otoscope-based AI had a high sensitivity of 0.93 (95% CI; 0.87-0.96) and a high specificity of 0.96 (95% CI; 0.94-0.98), when comparing OME and non-OME. This provided evidence that otoscope-based AI could help diagnose OME with high accuracy. One study showed that the performance of AI diagnosis was higher than that of human direct examination, with a high accuracy of 0.91 (95%CI: 90.1-91.3%) in distinguishing between normal and abnormal tympanic membrane images.¹⁹

Otoscope-based AI uses tympanic membrane image data obtained from the ears of patients with middle ear disorders. AI has the ability to capture characteristic images such as color, shape, and histogram of gradient (HOG) of the tympanic membrane, and then processes the obtained data with certain algorithms to classify abnormalities of the tympanic membrane.⁷

A study by Shim et al.¹⁷ showed that the use of AI could improve the accuracy of otologists in diagnosing the presence of OME. A telemedicine system with AI could help patients in areas with a shortage of specialists or for facilities that have difficulty making referrals.²⁰ Early detection of OME through AI can also contribute to the reduction of

unnecessary use of antibiotics and thus to the reduction of antimicrobial resistance.²¹

This otoscope-based AI study also found that OME was diagnosed in pediatric patients with high sensitivity and specificity of 0.93 (95% CI; 0.64-0.99); 0.94 (95% CI; 0.86-0.97). The accuracy of OME diagnosis by pediatric otologists is strongly influenced by the operator's technique and the child's cooperation during tympanometry and acoustic reflectometry examinations, as both devices require a precise view of the ear canal and a stationary position for several seconds to minutes, which is very difficult to achieve in children.¹³ Therefore, the use of AI to detect the presence of OME in pediatric patients can be used as a tool in daily practice.

In the majority of systematic reviews, the diagnosis of OME was based on a diagnosis by an otolaryngologist. One study specifically used intraoperative otoscopic images of children undergoing myringotomy for recurrent OMA or OME. The authors indicated that this approach could be the gold standard for detecting OME.¹⁵

Limitations of this study included the variability of the AI model, the gold standard method of OME diagnosis, the otoscopes used, and the various non-OME outcomes included in the study. AI has the potential to improve accuracy, time and resource savings in the diagnosis of OME so that further management can be appropriate. Further

research is needed using myringotomy as the gold standard and focusing more on differentiating the type of OME itself.

In conclusion, AI has been shown to have good performance in the diagnosis of OME. AI has the potential to improve the accuracy of OME diagnosis in primary care settings prior to referral for further management. Future research is needed to differentiate OME by type. AI in diagnosing OME can help primary care doctors detect and treat ear disorders early to prevent complications.

REFERENCE

- Vanneste P, Vanneste P, Page C. Otitis media with effusion in children: pathophysiology, diagnosis, and treatment. A review. *J Otol.* 2019;14(2):33-39.
- Parmar S, Davessar JL, Singh G, Arora N, Kansal L, Singh J. Prevalence of otitis media with effusion in children with hearing loss. *Indian J Otolaryngol Head Neck Surg.* 2019;71(Suppl 2):1276-1281.
- Brennan-Jones CG, Whitehouse AJO, Calder SD, Da Costa C, Eikelboom RH, Swanepoel DW, et al. Does otitis media affect later language ability? A prospective birth cohort study. *J Speech Lang Hear Res.* 2020;63(7):2441-2452.
- Crowson MG, Hartnick CJ, Diercks GR, Gallagher TQ, Fracchia MS, Setlur J, et al. Machine learning for accurate intraoperative pediatric middle ear effusion diagnosis. *Pediatrics.* 2021;147(4):e2020034546. doi:10.1542/peds.2020-034546.
- Al-Antari MA. Advancements in artificial intelligence for medical computer-aided diagnosis. *Diagnostics (Basel).* 2024;14(12):1265. doi:10.3390/diagnostics14121265.
- Song D, Kim T, Lee Y, Kim J. Image-based artificial intelligence technology for diagnosing middle ear diseases: a systematic review. *J Clin Med.* 2023;12(18):5831.
- Byun H, Yu S, Oh J, Bae J, Yoon MS, Lee SH, et al. An assistive role of a machine learning network in diagnosis of middle ear diseases. *J Clin Med.* 2021;10(15):3198. doi:10.3390/jcm10153198.
- Choi Y, Chae J, Park K, Hur J, Kweon J, Ahn JH. Automated multi-class classification for prediction of tympanic membrane changes with deep learning models. *PLoS One.* 2022;17(10):e0275846. doi:10.1371/journal.pone.0275846.
- Khan MA, Kwon S, Choo J, Hong SM, Kang SH, Park IH, et al. Automatic detection of tympanic membrane and middle ear infection from oto-endoscopic images via convolutional neural networks. *Neural Netw.* 2020;126:384-394. doi:10.1016/j.neunet.2020.03.023.
- Viscaino M, Talamilla M, Maass JC, Henríquez P, Délano PH, Auat Cheein C, et al. Color dependence analysis in a CNN-based computer-aided diagnosis system for middle and external ear diseases. *Diagnostics (Basel).* 2022;12(4):917. doi:10.3390/diagnostics12040917.
- Dubois C, Eigen D, Simon F, Couloigner V, Gormish M, Chalumeau M, et al. Development and validation of a smartphone-based deep-learning-enabled system to detect middle-ear conditions in otoscopic images. *NPJ Digit Med.* 2024;7(1):162. doi:10.1038/s41746-024-01159-9.
- Cai Y, Yu JG, Chen Y, Liu C, Xiao L, Grais EM, et al. Investigating the use of a two-stage attention-aware convolutional neural network for the automated diagnosis of otitis media from tympanic membrane images: a prediction model development and validation study. *BMJ Open.* 2021;11(1):e041139. doi:10.1136/bmjopen-2020-041139.
- Tran TT, Fang TY, Pham VT, Lin C, Wang PC, Lo MT. Development of an automatic diagnostic algorithm for pediatric otitis media. *Otol Neurotol.* 2018;39(8):1060-1065. doi:10.1097/MAO.0000000000001897.
- 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. doi:10.3390/jpm12111855.
- Chen YC, Chu YC, Huang CY, Lee YT, Lee WY, Hsu CY, et al. Smartphone-based artificial intelligence using a transfer learning algorithm for the detection and diagnosis of middle ear diseases: a retrospective deep learning study. *EClinicalMedicine.* 2022;51:101543. doi:10.1016/j.eclinm.2022.101543.

16. Shim JH, Sunwoo W, Choi BY, Kim KG, Kim YJ. Improving the accuracy of otitis media with effusion diagnosis in pediatric patients using deep learning. *Bioengineering (Basel)*. 2023;10(11):1304.
17. Noda M, Yoshimura H, Okubo T, Koshu R, Uchiyama Y, Nomura A, et al. Feasibility of multimodal artificial intelligence using GPT-4 Vision for the classification of middle ear disease: qualitative study and validation. *JMIR AI*. 2024;3:e58342.
18. Habib AR, Kajbafzadeh M, Hasan Z, Wong E, Gunasekera H, Perry C, et al. Artificial intelligence to classify ear disease from otoscopy: a systematic review and meta-analysis. *Clin Otolaryngol*. 2022;47(3):401-413. doi:10.1111/coa.13925.
19. Ning AY, Cabrera CI, D'Anza B. Telemedicine in otolaryngology: a systematic review of image quality, diagnostic concordance, and patient and provider satisfaction. *Ann Otol Rhinol Laryngol*. 2021;130(2):195-204.
20. Wilkinson E. AI app that can diagnose acute otitis media may cut antibiotic use, say researchers. *BMJ*. 2024;384:q572. doi:10.1136/bmj.q572.