International Journal of Pediatric Otorhinolaryngology 205 (2026) 112847

Contents lists available at ScienceDirect

International Journal of Pediatric Otorhinolaryngology

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

An artificial intelligence classifier as a screening tool to rule out otitis media in children

Simo Nuuttila [a,*], Antti Vallin [b], Tuomas Klockars [b], Aino Ruohola [a], Miia Laine [c,d], Lotta E. Ivaska [e], Paula A. Tähtinen [a]

[a] Department of Paediatrics and Adolescent Medicine, Turku University Hospital and University of Turku, Kiinamyllynkatu 4-8, Turku, 20520, Finland
[b] Sibbo Medical Devices Oy, Hitsaajankatu 22, Helsinki, 00810, Finland
[c] Department of Clinical Microbiology, Turku University Hospital, Kiinamyllynkatu 4-8, Turku, 20520, Finland
[d] Institute of Biomedicine, University of Turku, Kiinamyllynkatu 4-8, Turku, 20520, Finland
[e] Department of Otorhinolaryngology – Head and Neck Surgery, Turku University Hospital and University of Turku, Kiinamyllynkatu 4-8, Turku, 20520, Finland

ARTICLE INFO

Keywords:
Otitis media
AI
Diagnostic classifier

ABSTRACT

Objective: Acute otitis media is the most common bacterial infection among children and a significant global health burden. Despite its high incidence, diagnostic accuracy is poor. The objective of this study was to evaluate whether an artificial intelligence classifier can rule out otitis media in children based on a tympanic membrane image.

Methods: Artificial intelligence analysis of tympanic membrane images was carried out on images gathered as part of a randomized double-blind study. 793 tympanic membrane images were analyzed with an AI classifier. Images were obtained from children aged 6 to 35 months participating in a trial investigating the efficacy of amoxicillin-clavulanate for acute otitis media. The primary outcome was the sensitivity, specificity and accuracy of the classifier.

Results: All four variants of the artificial intelligence classifier showed excellent sensitivity for an abnormal ear (96% to 100%), and areas under the curves were respectively high (0.83-0.92). After a change in image normalization due to an initially poor image quality, the performance of the best variant improved to a specificity of 73%, and sensitivity remained high (92%).

Conclusions: Our study suggests that an artificial intelligence classifier at a primary level can rule out otitis media in children. This may eliminate the need for a physician's visit in the great majority of suspected acute otitis media cases in children with healthy ears. Further research in a parent-led setting is needed to measure the real-world impact of automatic classifiers.

1. Introduction

Acute otitis media (AOM) is a sudden infection of the middle ear accompanied by acute signs of illness and tympanic membrane (TM) bulging [1]. It is the most common bacterial infection among young children [2]. Due to its high incidence rate, AOM can be considered one of the most significant conditions contributing to public morbidity, especially among young children. Annually over 15 million pediatric AOM cases are diagnosed in the United States alone, resulting in over 4 billion USD in added health care costs [3].

Despite its high prevalence among young children, reaching diagnosis in this age group is particularly challenging, largely due to practical limitations such as the fussiness of the child and a narrow ear canal which is easily obscured by cerumen but difficult to clean.

When parents suspect their child may have AOM and symptoms are not severe, they typically first meet with a triage nurse or physician's assistant (PA). At this phase, a thorough anamnestic history is taken, and the child's overall condition is evaluated. If the child is in poor condition, they are quickly escalated to a physician, while a child in overall good condition may be discharged with follow-up instructions. However, triage professionals are typically not well trained in otoscopy, resulting in most suspected AOM cases being directly escalated to a physician.

* Corresponding author. Department of Paediatrics and Adolescent Medicine, Turku University Hospital, Kiinamyllynkatu 4-8, FIN-20520, Turku, Finland.
E-mail addresses: scgnuu@utu.fi (S. Nuuttila), antti.vallin@sibbomedicaldevices.com (A. Vallin), tuomas.klockars@sibbomedicaldevices.com (T. Klockars), aino.ruohola@utu.fi (A. Ruohola), miia.laine@varha.fi (M. Laine), lesaar@utu.fi (L.E. Ivaska), paula.tahtinen@utu.fi (P.A. Tähtinen).

https://doi.org/10.1016/j.ijporl.2026.112847
Received 11 March 2026; Received in revised form 30 April 2026; Accepted 8 May 2026
Available online 9 May 2026

The diagnosis of AOM is based on taking a thorough history and performing reliable otoscopy. Traditional physician-performed otoscopy, largely relying on the appearance of the TM and in particular its bulging [4,5], has low diagnostic accuracy and largely depends on the level of expertise [5–10]. Novel measures for diagnosing AOM, such as optical coherence tomography and low-coherence interferometry have been developed, but they lack clinical feasibility [11].

Previous studies have reported very high diagnostic accuracy when digital otoscopy images are analyzed by artificial intelligence (AI) and neural network algorithms [12–18]. A key advantage in leveraging AI as a decision support tool is that, compared to other novel modalities, it largely relies on existing and widely accessible hardware and technology.

The use of digital otoscopy and AI as a decision support and triage tool for nurses or physician's assistants may improve the selection of patients who benefit from a physician's evaluation, leading to a more efficient use of scarce resources.

The objective of our study was to examine the diagnostic accuracy of an AI classifier in categorizing TM images as either normal or abnormal. Based on these findings, our goal was to determine the possibility of ruling out otitis media (OM) in children by combining digital otoscopy with an AI classifier.

2. Materials and methods

2.1. Patient population

For our analysis, 793 TM images gathered from 100 children in a primary care setting were selected. Images were obtained from children aged 6 to 35 months who were participating in a large randomized, double-blind, placebo-controlled trial investigating the efficacy of amoxicillin-clavulanate for AOM (ClinicalTrials.gov ID NCT00299455) [19]. The images were taken by trained otoscopists on days 1, 3, 8, 15, 30 and 60, thus representing different stages of OM recovery.

The images were collected with a digital video otoscope (Jedmed, St. Louis, MO, USA). The trial was conducted at a primary care level at Turku health care center between 2006 and 2009, and the study protocol was approved by the Ethical Committee of the Hospital District of Southwest Finland. Written informed consent to participate in the trial was obtained from parents.

2.2. Study design

To form the pre-specified gold standard (Fig. 1), TM images were classified by three specialized physicians (two specialists in ear, nose and throat (ENT) diseases TK and LEI; one specialist in pediatrics PAT) as normal or abnormal. Abnormal classification included all cases which could not be classified as normal due to pathology, i.e. images of AOM or otitis media with effusion (OME), but also images where the TM could not be classified normal due to poor image quality or cerumen build-up. The physicians based their classification on the visual appearance of the TM, focusing on the hallmark features of middle ear pathologies, such as bulging, retraction or opacification of the membrane, visible fluid behind the membrane or in severe cases TM bullae or perforation Where the experts assessment of the categorization was controversial, a conclusion was reached by discussion, and finally by a vote, if a consensus was not achieved. Of the 793 images in the study, all three specialists were unanimous in the classification of 685 images (86%), 107 (13%) were classified by discussion and 1 (0.1%) required a vote.

Once categorized, the 793 TM images were analyzed by the four variants of an AI classifier. In the same way as the physicians, they classified TM findings as either normal or abnormal. The classification by AI was recorded and compared to the gold standard.

The primary outcome was the sensitivity, specificity and diagnostic accuracy of the AI classifier, and ultimately the area under the receiver operating characteristic (ROC) curve.

Our study followed the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD + AI) reporting guideline [20].

2.3. AI classifier

An AI classifier was developed by Sibbo Medical Devices (Helsinki, Finland) to classify TM images either as normal or abnormal (Fig. 2). The classifier is based on Multi-Axis Vision Transformer architecture, combining Convolutional Neural Networks and Vision Transformers, implemented with the PyTorch framework. The training dataset consisted of 2700 TM images labelled by a panel of 7 ENT specialists, and 300 additional images were used as a test set. The training and test datasets consisting of a total of 3000 TM images which were independent from those analyzed in this study. Four different variants were trained on the training data with conventional transfer learning

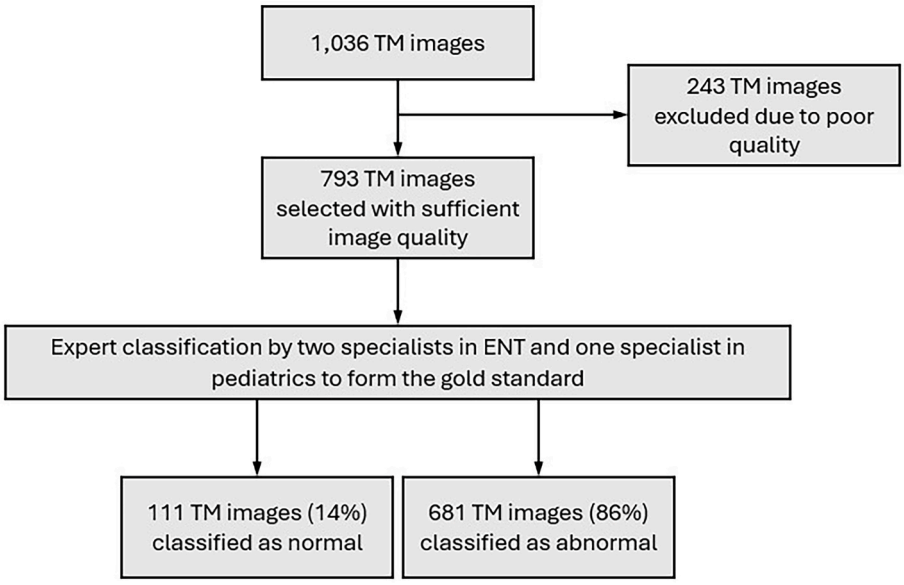

Fig. 1. Forming the gold standard. The gold standard was set by two specialists in ear, nose and throat diseases, and one specialist in pediatrics.

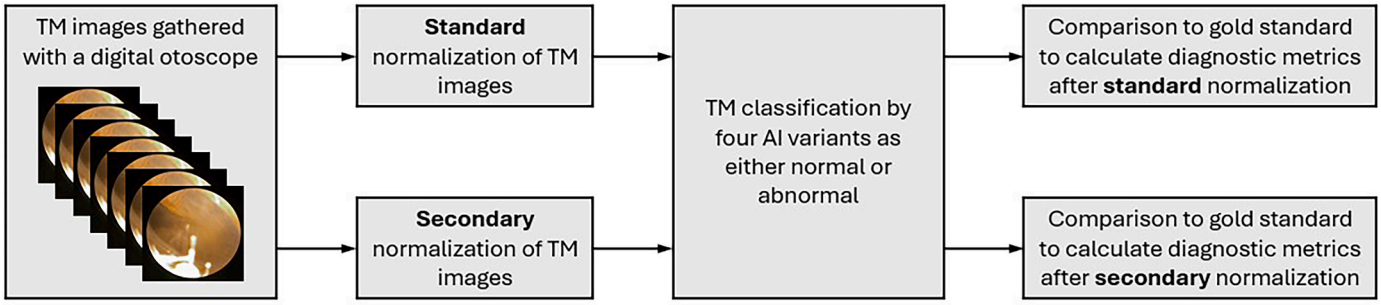

Fig. 2. Illustration of image normalization and operating logic of the AI classifier. Image normalization was performed as a pre-processing step. Due to technical discrepancies between the AI training and test data sets, secondary normalization was carried out to improve the performance of the classifier.

methods. A binary classification output was generated by each variant, with a 50% probability threshold used for classification, consistent with prior studies [12]. A detailed description of the function and key differences of the different variants is available in the supplementary material (Supplement 1).

2.4. Image normalization

Image normalization is a pre-processing step during AI training and inference, which maps the numerical values of image pixel intensities to a range from −1 to 1. In deep learning, it consists of calculating the mean and standard deviation of the image channels for all training data and using those values to normalize all images for the AI classifier in the future. This process helps the training speed and convergence of the AI classifier. In this study, a standard normalization was performed for all images before initial analyses.

The image set used in the study was taken during the years 2006-2009. It differed markedly in e.g. coloring and resolution when compared to the more modern image set the AI classifier was trained on. This resulted in initially poor specificity for abnormal ears after standard normalization.

To minimize this image quality bias, we calculated specific normalization values for the study dataset and performed a secondary normalization to make the datasets more similar in pixel distribution. As expected, using these values yields noticeably better specificity. The process of image normalization and basic function of the classifier is presented in Fig. 2.

2.5. Statistical analysis

We calculated sensitivity, specificity and diagnostic accuracy for the four variants of the AI classifier after both standard and secondary normalization. Diagnostic accuracy is defined as the proportion of test results that are correct. In addition to these calculations, we constructed ROC curves and calculated the areas under the curves (AUC). Analyses were performed with Python software, version 3.12.5, and R software, version 4.5.1.

3. Results

In a set of 793 TM images of which 111 were normal, all four variants proved to have very good sensitivity for an abnormal ear after standard normalization (96% to 100% depending on the variant). The specificity of the variants was clearly inferior, with the three best variants ranging from 36% to 44%, and with the outlying variant 2 as low as 4% (Table 1). Accuracy ranged from 87% to 90%.

Secondary normalization resulted in a clearly better performance of all four variants (Table 2). The sensitivity for an abnormal ear remained high (89% to 95% depending on the variant) and the specificity clearly improved (45% to 73% depending on the variant) (Table 2). Accuracy ranged from 85% to 90%.

The change in normalization produced no change in the AUCs for any of the variants, ranging from 0.83 to 0.92 after standard normalization and 0.85 to 0.92 after secondary normalization. High AUCs demonstrate that the AI classifier has an excellent ability to differentiate between normal and abnormal ears. ROC curves and AUC's are shown in Fig. 3.

4. Discussion

In this study, we compared the accuracy of an AI classifier in ruling out otitis media in young children. Our results show that excellent sensitivity for ruling out OM can be achieved and a simple change in normalization of the TM images yielded noticeably better specificity for the AI classifier.

Diagnostic otoscopy is challenging for most physicians. Several studies suggest that traditional physician-performed otoscopy when diagnosing AOM exhibits a low diagnostic accuracy, ranging from as low as 25% to 88% at best, depending on medical specialty and experience [5–10]. Additionally, the diagnosis of AOM cannot be predicted based on symptom quality, duration or severity [21–23]. New and efficient ways for reaching the correct diagnosis are called for, which is why we applied an AI classifier to a clinical setting.

Our study shows similar sensitivity and overall accuracy compared to previous AI studies, in which sensitivity and diagnostic accuracy has ranged from 91% to 96% and 80% to 93%, respectively [12–18]. ROC curves and the respective AUCs in our study were slightly better than in a similar study with AUCs ranging from 0.77 to 0.92 [17]. A recent study reported a clearly superior specificity (94%) and an equal sensitivity

Table 1
Results after standard normalization. Sensitivity, specificity, diagnostic accuracy with their corresponding 95% confidence intervals, and AUC's were calculated for each variant of the AI classifier.

	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Accuracy, % (95% CI)	AUC	Considered normal, n (%)
Gold standard	100	100	100	1.0	111 (14)
Variant 1	98 (97-99)	36 (27-46)	89 (87-91)	0.92	53 (7)
Variant 2	100 (99-100)	4 (1-9)	87 (84-89)	0.86	4 (1)
Variant 3	96 (94-97)	44 (35-54)	89 (87-91)	0.83	75 (10)
Variant 4	98 (97-99)	41 (31-50)	90 (88-92)	0.90	58 (7)

Table 2
Results after secondary normalization. Sensitivity, specificity, diagnostic accuracy with their corresponding 95% confidence intervals, and AUC's were calculated for each variant of the AI classifier.

	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Accuracy, % (95% CI)	AUC	Considered normal, n (%)
Gold standard	100	100	100	1.0	111 (14)
Variant 1	95 (94-97)	57 (48-66)	90 (88-92)	0.92	94 (12)
Variant 2	94 (92-96)	45 (36-54)	87 (85-89)	0.85	91 (12)
Variant 3	89 (86-91)	63 (54-72)	85 (83-88)	0.86	146 (18)
Variant 4	92 (90-94)	73 (65-81)	89 (87-92)	0.91	135 (17)

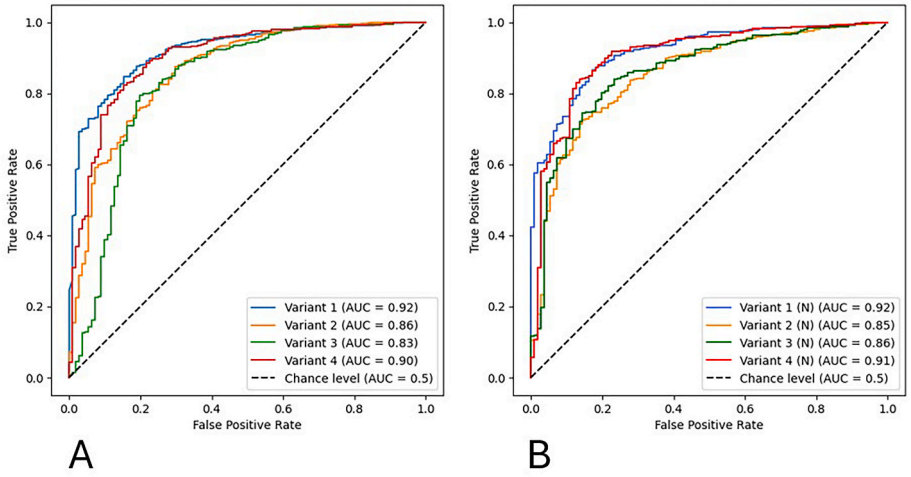

Fig. 3. ROC curves and corresponding AUCs for four artificial intelligence variants A) after standard normalization and B) after secondary normalization.

(94%) compared to ours. However, a key difference is that the tool used in the study was developed to classify findings as either AOM or non-AOM (including both healthy ears and e.g. OME) [12], compared to our normal vs. abnormal approach. Focusing on the identification of a single condition with clearly defined features may yield better sensitivity and specificity compared to our approach of ruling out all types of pathologies. Additionally, the study used short videos of the TM as input for the classifier, which is possibly a more optimal form of input than still images [11,24–26].

Compared to other AI studies[12–18], our study setting placed emphasis on a primary care and real-world setting. Our dataset consisted of a wide variety of otoscopic findings that in some cases can neither be classified as AOM or OME, nor are they completely normal. Additionally, some of the images were visually suboptimal due to e.g. cerumen remains, poor lighting or positioning of the digital otoscope. These features improve the relevance of the dataset as it better represents a real-world primary level setting. On the other hand, this impaired the classifier's specificity as few patients have a completely normal TM finding soon after AOM diagnosis, and many non-pathological artefacts were likely flagged as abnormal features by the classifier.

Other modalities for screening middle ear fluid, such as tympanometry, optical coherence tomography and acoustic reflectometry, have presented a wide range of diagnostic metrics roughly in line with our classifier [27–30]. These modalities are based on the detection of middle ear fluid, and similar to our classifier, cannot differentiate between AOM and OME [31,32]. Contrary to some novel modalities, digital otoscopy combined with an AI classifier largely relies on existing and widely accessible hardware and technology and hence comes with lower acquisition and operating costs.

The four variants of the classifier had different structures and architectures. After a secondary normalization, variant 4 with a 92% sensitivity for an abnormal ear, 73% specificity, 89% accuracy and 0.91 AUC would likely perform best as a real-life screening tool. In practice, if used at a primary level by nurses or PAs, the tool could rule out OM in 73% of children with suspected AOM who have a healthy ear. This population could be discharged with follow-up instructions with no need to visit the physician, assuming no other reasons for a visit are present. This reduction in the demand for a scarce resource is a significant change compared to the current state where this patient group typically meets directly with the physician. Our deduction is based on the assumption that without the help of an AI classifier, all patients presenting with suspected AOM based on symptoms would meet a physician and undergo traditional otoscopy.

Our AI classifier represents a binary (normal vs. abnormal) classifier to be used as a screening tool for nurses and PAs. When screening is performed at a primary level by non-physicians, the tool is required to have excellent sensitivity for recognizing pathologies, to some extent at the cost of an inferior specificity. We believe this is a necessary calibration for a screening tool to minimize the false negative rate and the risk of missing a potentially treatable middle ear pathology, such as AOM. Non-physician health care professionals would be required to undergo training on performing safe otoscopy on a potentially struggling toddler to avoid e.g. damage to the ear canal and ensure high quality of images.

We believe that our approach of classifying tympanic membrane findings as normal or abnormal, regardless of the specific pathology, is more practical as a screening tool for non-physician health care professionals than attempting to identify precise diagnoses. As there is currently no clear consensus on which AOM patients most benefit from antimicrobials, classifiers cannot be trained to differentiate this either. This further justifies the use of our binary classification approach as patients in overall good condition with a negative test result may be provided with clinical clearance.

The 92% sensitivity of our classifier can be considered a very good diagnostic metric for a clinical test. However, it's important to note that the reliability of diagnostic accuracy metrics largely depends on the prevalence of the condition [33]. In practice, the use of our classifier would result in an 8% false negative rate. We argue that the potential risks of not identifying AOM in a child in overall good condition are minimal, as the majority of AOM are likely to heal without antimicrobial

therapies [34], and severe complications of untreated AOM are extremely rare [35]. Patients with persisting symptoms despite previous normal TM classification should be examined by a physician, as well as patients presenting with overall poor condition or significant parental concern [36].

Our study has certain limitations. The images used in the study were collected during 2006-2009. As a result, their technical quality is inferior compared to modern images which the AI classifier was trained on. The difference in the image quality is a likely explanation for the initially low specificity for abnormal ears after standard normalization, ranging from 4% to 44% depending on the variant. In other words, the TM images were classified as abnormal in part due to poor image quality and not actual TM pathology.

To overcome this problem, we performed a secondary normalization using specific normalization values, i.e. forcing the datasets to be more similar. This relatively simple change in image normalization resulted in a clear increase in specificity for an abnormal ear, which now ranged from 45% to 73%.

This highlights one major problem of AI image analyses: the reliability of the classifier is dependent on the similarity of the test and training datasets. Different digital otoscopes may require device-specific normalization procedures to minimize image quality bias. Thus, in theory, otoscope LED wavelength or even the color of the ear speculum are likely to impact the classifier results. This could be a significant obstacle to the widespread adoption of artificial intelligence in digital otoscopes.

Our classifier is developed to differentiate a normal TM from an abnormal one but cannot differentiate which exact TM features it has based its classification on. This reduces its value as a tool for educational purposes. Additionally, the distribution of the images in the test set was skewed towards abnormal, as only 14% of images were classified as normal using the gold standard.

Our large dataset consisting of TM images from children aged 6 to 35 months served ideally for the purpose of evaluating TM abnormalities, as AOM is most common and most difficult to diagnose in this age group.

5. Conclusions

Our study suggests that an AI classifier can reliably recognize and rule out OM in children. If used as a screening tool at a primary level, the best variant of our classifier could reduce the need for a physician's evaluation in 73% of suspected OM cases in children who have a healthy ear; this could potentially result in significant cost savings in a very large patient population. Patients with persisting symptoms despite previous normal TM classification should be examined by a physician, as well as patients presenting with overall poor condition or significant parental concern. Future research should focus on optimizing the classifier's performance to reduce the count of false positives without decreasing the tool's high sensitivity, validating the classifier in a multicenter or device-diverse setting, testing the classifier in a parent-led setting and evaluating the performance of AI classifiers against less experienced otoscopists.

Data sharing statement

Deidentified participant data will be made available upon reasonable request from the corresponding author.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

No generative AI or AI-assisted technologies were used in the manuscript preparation process.

Funding

The original trial was funded by the Fellowship Award of the European Society for Paediatric Infectious Diseases (to AR) and by grants from the Foundation for Paediatric Research; Research Funds from Specified Government Transfers; the Jenny and Antti Wihuri Foundation; the Paulo Foundation; the Maud Kuistila Memorial Foundation; the Emil Aaltonen Foundation; the Finnish Cultural Foundation, Varsinais-Suomi Regional Fund; the Turku University Hospital Research Foundation; and the Finnish-Norwegian Medical Foundation.

CRediT authorship contribution statement

Simo Nuuttila: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing – original draft, Writing – review & editing. **Antti Vallin:** Data curation, Formal analysis, Project administration, Software, Visualization, Writing – review & editing. **Tuomas Klockars:** Investigation, Methodology, Validation, Writing – review & editing. **Aino Ruohola:** Data curation, Funding acquisition, Investigation, Writing – review & editing. **Miia Laine:** Data curation, Investigation, Writing – review & editing. **Lotta E. Ivaska:** Data curation, Investigation, Methodology, Writing – review & editing. **Paula A. Tähtinen:** Data curation, Formal analysis, Investigation, Methodology, Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Tuomas Klockars and Antti Vallin reports a relationship with Sibbo Medical Devices, Helsinki, Finland that includes: board membership, consulting or advisory, employment, equity or stocks, and travel reimbursement. Tuomas Klockars has patent licensed to Patent for digital otoscope used in the study. Dr Klockars is a founder and major share holder of Sibbo Medical Devices. Mr Vallin is an employee of Sibbo Medical Devices.Drs Nuuttila, Ruohola, Laine, Ivaska and Tähtinen reported no conflicts of interest. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijporl.2026.112847.

References

[1] A.S. Lieberthal, A.E. Carroll, T. Chonmaitree, T.G. Ganiats, A. Hoberman, M. A. Jackson, M.D. Joffe, D.T. Miller, R.M. Rosenfeld, X.D. Sevilla, R.H. Schwartz, P. A. Thomas, D.E. Tunkel, The diagnosis and management of acute otitis media, Pediatrics 131 (2013) e964–e999, https://doi.org/10.1542/peds.2012-3488.

[2] A. Vergison, R. Dagan, A. Arguedas, J. Bonhoeffer, R. Cohen, I. Dhooge, A. Hoberman, J. Liese, P. Marchisio, A.A. Palmu, G.T. Ray, E.A.M. Sanders, E.A. F. Simões, M. Uhari, J. van Eldere, S.I. Pelton, Otitis media and its consequences: beyond the earache, Lancet Infect. Dis. 10 (2010) 195–203, https://doi.org/10.1016/S1473-3099(10)70012-8.

[3] T. Hu, Y. Song, N. Done, S. Mohanty, Q. Liu, E.M. Sarpong, E. Lemus-Wirtz, J. Signorovitch, T. Weiss, Economic burden of acute otitis media, pneumonia, and invasive pneumococcal disease in children in the United States after the introduction of 13-valent pneumococcal conjugate vaccines during 2014-2018, BMC Health Serv. Res. 23 (2023) 398, https://doi.org/10.1186/s12913-023-09244-7.

[4] N. Shaikh, A. Hoberman, H.E. Rockette, M. Kurs-Lasky, Development of an algorithm for the diagnosis of otitis media, Acad. Pediatr. 12 (2012) 214–218, https://doi.org/10.1016/j.acap.2012.01.007.

[5] N. Shaikh, M.K. Stone, M. Kurs-Lasky, A. Hoberman, Interpretation of tympanic membrane findings varies according to level of experience, Paediatr. Child Health 21 (2016) 196–198, https://doi.org/10.1093/pch/21.4.196.

[6] R.M. Rosenfeld, Diagnostic certainty for acute otitis media, Int. J. Pediatr. Otorhinolaryngol. 64 (2002) 89–95, https://doi.org/10.1016/s0165-5876(02)00073-3.
[7] M.E. Pichichero, M.D. Poole, Assessing diagnostic accuracy and tympanocentesis skills in the management of otitis media, Arch. Pediatr. Adolesc. Med. 155 (2001) 1137, https://doi.org/10.1001/archpedi.155.10.1137.
[8] M.E. Pichichero, Diagnostic accuracy, tympanocentesis training performance, and antibiotic selection by pediatric residents in management of otitis media, Pediatrics 110 (2002) 1064–1070, https://doi.org/10.1542/peds.110.6.1064.
[9] W.S. Jones, P.H. Kaleida, How helpful is pneumatic otoscopy in improving diagnostic accuracy? Pediatrics 112 (2003) 510–513, https://doi.org/10.1542/peds.112.3.510.
[10] A. Sorrento, M.E. Pichichero, Assessing diagnostic accuracy and tympanocentesis skills by nurse practitioners in management of otitis media, J. Am. Acad. Nurse Pract. 13 (2001) 524–529, https://doi.org/10.1111/j.1745-7599.2001.tb00019.x.
[11] S. Esposito, S. Bianchini, A. Argentiero, R. Gobbi, C. Vicini, N. Principi, New approaches and technologies to improve accuracy of acute otitis media diagnosis, Diagnostics 11 (2021) 2392, https://doi.org/10.3390/diagnostics11122392.
[12] N. Shaikh, S.J. Conway, J. Kovacevic, F. Condessa, T.R. Shope, M.A. Haralam, C. Campese, M.C. Lee, T. Larsson, Z. Cavdar, A. Hoberman, Development and validation of an automated classifier to diagnose acute otitis media in children, JAMA Pediatr. 178 (2024) 401–407, https://doi.org/10.1001/jamapediatrics.2024.0011.
[13] M.G. Crowson, D.W. Bates, K. Suresh, M.S. Cohen, C.J. Hartnick, "Human vs Machine" validation of a deep learning algorithm for pediatric middle ear infection diagnosis, Otolaryngol. Head Neck Surg. 169 (2023) 41–46, https://doi.org/10.1177/01945998221119156.
[14] Z. Wu, Z. Lin, L. Li, H. Pan, G. Chen, Y. Fu, Q. Qiu, Deep learning for classification of pediatric otitis media, Laryngoscope 131 (2021) E2344–E2351, https://doi.org/10.1002/lary.29302.
[15] Z. Zhong, X. Guo, D. Jia, H. Zheng, Z. Wu, X. Wang, Artificial intelligence as an auxiliary tool in pediatric otitis media diagnosis, Int. J. Pediatr. Otorhinolaryngol. 187 (2024) 112154, https://doi.org/10.1016/j.ijporl.2024.112154.
[16] A. Kuruvilla, N. Shaikh, A. Hoberman, J. Kovačević, Automated diagnosis of otitis media: vocabulary and grammar, Int. J. Biomed. Imag. 2013 (2013) 327515, https://doi.org/10.1155/2013/327515.
[17] T.-T. Tran, T.-Y. Fang, V.-T. Pham, C. Lin, P.-C. Wang, M.-T. Lo, Development of an automatic diagnostic algorithm for pediatric otitis media, Otol. Neurotol. 39 (2018) 1060–1065, https://doi.org/10.1097/MAO.0000000000001897.
[18] C.-K. Shie, H.-T. Chang, F.-C. Fan, C.-J. Chen, T.-Y. Fang, P.-C. Wang, A hybrid feature-based segmentation and classification system for the computer aided self-diagnosis of otitis media, Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. 2014 (2014) 4655–4658, https://doi.org/10.1109/EMBC.2014.6944662.
[19] P.A. Tähtinen, M.K. Laine, P. Huovinen, J. Jalava, O. Ruuskanen, A. Ruohola, A placebo-controlled trial of antimicrobial treatment for acute otitis media, N. Engl. J. Med. 364 (2011) 116–126, https://doi.org/10.1056/NEJMoa1007174.
[20] G.S. Collins, K.G.M. Moons, P. Dhiman, R.D. Riley, A.L. Beam, B. Van Calster, M. Ghassemi, X. Liu, J.B. Reitsma, M. van Smeden, A.-L. Boulesteix, J. C. Camaradou, L.A. Celi, S. Denaxas, A.K. Denniston, B. Glocker, R.M. Golub, H. Harvey, G. Heinze, M.M. Hoffman, A.P. Kengne, E. Lam, N. Lee, E.W. Loder, L. Maier-Hein, B.A. Mateen, M.D. McCradden, L. Oakden-Rayner, J. Ordish, R. Parnell, S. Rose, K. Singh, L. Wynants, P. Logullo, TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods, BMJ 385 (2024) e078378, https://doi.org/10.1136/bmj-2023-078378.
[21] M.K. Laine, P.A. Tähtinen, O. Ruuskanen, P. Huovinen, A. Ruohola, Symptoms or symptom-based scores cannot predict acute otitis media at otitis-prone age, Pediatrics 125 (2010) e1154–e1161, https://doi.org/10.1542/peds.2009-2689.
[22] T. Heikkinen, O. Ruuskanen, Signs and symptoms predicting acute otitis media, Arch. Pediatr. Adolesc. Med. 149 (1995) 26–29, https://doi.org/10.1001/archpedi.1995.02170130028006.
[23] M. Niemela, M. Uhari, K. Jounio-Ervasti, J. Luotonen, O.P. Alho, E. Vierimaa, Lack of specific symptomatology in children with acute otitis media, Pediatr. Infect. Dis. J. 13 (1994) 765–768, https://doi.org/10.1097/00006454-199409000-00002.
[24] S. Ngombu, H. Binol, M.N. Gurcan, A.C. Moberly, Advances in artificial intelligence to diagnose otitis media: state of the art review, Otolaryngol. Head Neck Surg. 168 (2023) 635–642, https://doi.org/10.1177/01945998221083502.
[25] C. Soares, W. Clifton, W.D. Freeman, Use of handheld video otoscopy for the diagnosis of acute otitis media: technical note, Cureus 11 (2019) e5547, https://doi.org/10.7759/cureus.5547.
[26] E.M. Alenezi, K. Jajko, A. Reid, A. Locatelli-Smith, C.S. McMahen, K.F. Tao, J. Marsh, T. Bright, P.C. Richmond, R.H. Eikelboom, C.G. Brennan-Jones, Clinician-rated quality of video otoscopy recordings and still images for the asynchronous assessment of middle-ear disease, J. Telemed. Telecare 29 (2023) 435–443, https://doi.org/10.1177/1357633X20987783.
[27] R. Neuberger, J. Carter, A. Dives, A.S.C. Ibarrola, M. Ixba, E. Lambert, Y.-C. Liu, A. H. Messner, Optical coherence tomography as an emerging technique for identifying pediatric middle ear effusion, Laryngoscope (2025), https://doi.org/10.1002/lary.70296.
[28] D.M. Schwartz, R.H. Schwartz, Validity of acoustic reflectometry in detecting middle ear effusion, Pediatrics 79 (1987) 739–742.
[29] T. Muderris, A. Yazıcı, S. Bercin, G. Yalçıner, E. Sevil, M. Kırıs, Consumer acoustic reflectometry: accuracy in diagnosis of otitis media with effusion in children, Int. J. Pediatr. Otorhinolaryngol. 77 (2013) 1771–1774, https://doi.org/10.1016/j.ijporl.2013.08.019.
[30] N. Erkkola-Anttinen, M.K. Laine, P.A. Tähtinen, A. Ruohola, Can changes in parentally measured acoustic reflectometry levels predict the middle ear status? Int. J. Pediatr. Otorhinolaryngol. 95 (2017) 72–74, https://doi.org/10.1016/j.ijporl.2017.02.002.
[31] K.K. Helenius, M.K. Laine, P.A. Tähtinen, E. Lahti, A. Ruohola, Tympanometry in discrimination of otoscopic diagnoses in young ambulatory children, Pediatr. Infect. Dis. J. 31 (2012) 1003–1006, https://doi.org/10.1097/INF.0b013e31825cac94.
[32] M.K. Laine, P.A. Tähtinen, K.K. Helenius, R. Luoto, A. Ruohola, Acoustic reflectometry in discrimination of otoscopic diagnoses in young ambulatory children, Pediatr. Infect. Dis. J. 31 (2012) 1007–1011, https://doi.org/10.1097/INF.0b013e31825caf45.
[33] M.H. Murad, L. Lin, H. Chu, B. Hasan, R.A. Alsibai, A.S. Abbas, R.A. Mustafa, Z. Wang, The association of sensitivity and specificity with disease prevalence: analysis of 6909 studies of diagnostic test accuracy, CMAJ 195 (2023) E925–E931, https://doi.org/10.1503/cmaj.221802.
[34] N. Shaikh, Otitis media in young children, N. Engl. J. Med. 392 (2025) 1418–1426, https://doi.org/10.1056/NEJMcp2400531.
[35] N.E. Smolinski, E.J. Djabali, J. Al-Bahou, A. Pomputius, P.J. Antonelli, A. G. Winterstein, Antibiotic treatment to prevent pediatric acute otitis media infectious complications: a meta-analysis, PLoS One 19 (2024) e0304742, https://doi.org/10.1371/journal.pone.0304742.
[36] H. Pöyry, J. Turunen, E. Ritola, S. Hartikainen, J. Palviainen, I. Liimatta, U. Koskela, T. Pokka, M. Renko, N. Paalanne, O. Helve, M.V. Tejesvi, T. Ruuska-Loewald, Parental ability to identify severe illnesses in their children, JAMA Netw. Open 9 (2026) e2559998, https://doi.org/10.1001/jamanetworkopen.2025.59998.