ORIGINAL ARTICLE

The Role of Artificial Intelligence Algorithms in Challenging Diagnostic Cases – Between Potential and Real Clinical Support

Agata Doligalska-Dolina[EF], Klaudia Ziółkowska[BD], Maciej J. Wróbel[ACG]

Department of Otolaryngology and Laryngological Oncology, Collegium Medicum, Nicolaus Copernicus University, Bydgoszcz, Poland

ABSTRACT:

Introduction: Extranodal natural killer/T-cell lymphoma, nasal type (ENKTL), is a rare EBV-associated malignancy characterized by destructive tumors in the nasal cavity, often causing delayed diagnosis due to tissue necrosis. Artificial intelligence (AI) has the potential to assist in complex diagnostic challenges.

Aim: This study aimed to evaluate the diagnostic performance and clinical relevance of publicly accessible AI-driven platforms in identifying ENKTL based on a real, challenging clinical case.

Materials and methods: Clinical data from a 26-year-old male patient with protracted and complicated symptoms initially suggesting infectious conditions were used. Eight AI symptom assessment platforms were tested for diagnostic outputs using identical anonymized case inputs at early and advanced disease stages. Diagnostic suggestions, alignment with the final histopathological diagnosis, and quality of recommendations were analyzed.

Results: Six of eight platforms provided valid differential diagnoses consistent with clinical data. Early-stage data mainly generated infectious diagnoses such as peritonsillar abscess. Advanced-stage data shifted diagnoses toward invasive fungal, autoimmune, and malignant diseases, with platforms such as Doctronic and Symptomate explicitly including lymphoma. Diagnostic pathways varied, with recommendations for comprehensive laboratory tests, imaging, biopsies, and immunological assays. AI systems demonstrated oncologic vigilance when EBV viral data were included, emphasizing clinician oversight for accurate diagnosis.

Conclusions: AI-based diagnostic platforms show promise in supporting preliminary assessment and differential diagnosis in rare, complex cases such as ENKTL. While AI can guide diagnostic reasoning and propose investigations, it does not replace clinical judgment. Further development and validation of AI tools are needed to enhance their reliability and integration into medical practice.

KEYWORDS: clinical decision support, diagnostic algorithms, diagnostic delay, EBV-associated lymphoma, extranodal NK/T-cell lymphoma

STRESZCZENIE:

Wprowadzenie: Pozawęzłowy chłoniak z komórek NK/T typu nosowego (ENKTL) to rzadka nowotworowa choroba związana z wirusem EBV, która charakteryzuje się powstawaniem niszczących guzów w jamie nosowej i często opóźnionym rozpoznaniem z powodu martwicy tkanek. Sztuczna inteligencja (AI) wykazuje potencjał jako narzędzie wspomagające rozwiązywanie złożonych problemów diagnostycznych.

Cel: Celem niniejszego badania była ocena skuteczności diagnostycznej i adekwatności klinicznej ogólnodostępnych platform opartych na sztucznej inteligencji w rozpoznawaniu ENKTL na podstawie rzeczywistego, trudnego przypadku klinicznego.

Received: 07.04.2026
Accepted: 20.04.2026
Published: 00.00.2026

CITE THIS ARTICLE AS:
Doligalska-Dolina A., Ziolkowska K., Wrobel M.J.: The Role of Artificial Intelligence Algorithms in Challenging Diagnostic Cases – Between Potential and Real Clinical Support; Otolaryngol Pol 2026; 80 (2): 39–45; DOI: 10.5604/01.3001.0055.7287

AUTHORS' CONTRIBUTION:
A. Conceptualization
B. Data Curation
C. Formal Analysis
D. Funding Acquisition
E. Investigation
F. Methodology
G. Project Administration

CORRESPONDING AUTHOR:
Agata Doligalska-Dolina MD (ORCID: 0009-0005-2989-3927); Department of Otolaryngology and Laryngological Oncology, Collegium Medicum, Nicolaus Copernicus University;
Marii Skłodowskiej-Curie street 9, 85-094 Bydgoszcz, Poland;
E-mail: doli285@wp.pl

COPYRIGHT:
Some rights reserved: Polish Society of Otorhinolaryngologists Head and Neck Surgeons. Published by Index Copernicus Sp. z o.o.

OPEN ACCESS:
The content of the journal „Polish Society of Otorhinolaryngologists Head and Neck Surgeons" is circulated on the basis of the Open Access which means free and limitless access to scientific data.

CREATIVE COMMONS CC BY:
This material is available under the Creative Commons – Attribution NonCommercial 4.0 International (CC BY-NC 4.0). The full terms of this license are available on: https://creativecommons.org/licenses/by-nc/4.0/Legalcode

COMPETING INTERESTS:
The authors declare that they have no competing interests.

Materiały i metody: Wykorzystano dane kliniczne dotyczące 26-letniego pacjenta płci męskiej, u którego obserwowano przedłużające się i powikłane objawy, początkowo sugerujące choroby zakaźne. Przetestowano osiem opartych na AI platform oceny objawów i oceniono uzyskane wyniki diagnostyczne sformułowane przez poszczególne platformy w oparciu na identycznych, zanonimizowanych danych dotyczących przedmiotowego przypadku we wczesnym i zaawansowanym stadium choroby. Przeanalizowano sugestie diagnostyczne, zgodność z ostateczną diagnozą histopatologiczną oraz jakość zaleceń.

Wyniki: Sześć spośród ośmiu platform dostarczyło prawidłowe rozpoznania różnicowe zgodne z danymi klinicznymi. Dane z wczesnego etapu choroby wskazywały głównie na schorzenia zakaźne, takie jak ropień okołomigdałkowy. Dane dotyczące zaawansowanego stadium choroby wiązały się z przesunięciem spektrum diagnostycznego w kierunku rozpoznania chorób grzybiczych o przebiegu inwazyjnym, chorób autoimmunologicznych oraz nowotworów złośliwych, przy czym w wynikach generowanych przez platformy takie jak Doctronic i Symptomate wyraźnie uwzględnione zostały chłoniaki. Proponowane schematy diagnostyczne były zróżnicowane i obejmowały zalecenia dotyczące kompleksowych badań laboratoryjnych, badań obrazowych, biopsji oraz testów immunologicznych. Platformy oparte na sztucznej inteligencji wykazały czujność onkologiczną po uwzględnieniu w danych wejściowych informacji o zakażeniu EBV, co podkreśla znaczenie, jakie dla postawienia trafnego rozpoznania ma nadzorowanie procesu przez lekarza klinicystę.

Wnioski: Platformy diagnostyczne oparte na AI wydają się obiecującym narzędziem wspomagającym wstępną ocenę i diagnostykę różnicową w rzadkich, złożonych przypadkach, takich jak ENKTL. Chociaż AI może wspomagać proces diagnostyczny i sugerować dalsze badania, nie zastępuje ona oceny klinicznej. Konieczne jest dalsze doskonalenie i walidacja narzędzi opartych na AI w celu zwiększenia ich wiarygodności i ułatwienia ich wdrożenia w praktyce medycznej.

SŁOWA KLUCZOWE: algorytmy diagnostyczne, chłoniak związany z wirusem EBV, opóźnienie diagnostyczne, pozawęzłowy chłoniak z komórek NK/T, system wspomagania decyzji klinicznych

ABBREVIATIONS

AI – artificial intelligence
CMV – cytomegalovirus
CRP – C-reactive protein
CT – computed tomography
DDGP – cisplatin, dexamethasone, gemcitabine, and pegaspargase
DNA – deoxyribonucleic acid
EBER – Epstein–Barr virus–encoded small RNAs
EBV – Epstein–Barr virus
ENKTL – extranodal natural killer/T-cell lymphoma
ESR – erythrocyte sedimentation rate
GDPR – General Data Protection Regulation
GPA – granulomatosis with polyangiitis
HIPAA – Health Insurance Portability and Accountability Act
HIV – human immunodeficiency virus
MRI – magnetic resonance imaging
PET/CT – positron emission tomography–computed tomography
SOC 2 – System and Organization Controls 2

INTRODUCTION

Extranodal natural killer/T-cell lymphoma, nasal type (ENKTL), is a rare Epstein–Barr virus (EBV)-associated lymphoproliferative malignancy. The disease occurs more frequently in East Asia, Central and South America, and is relatively uncommon in Europe [1]. It is characterized by the formation of destructive tumors in the nasal cavity and nasopharynx, typically presenting with nasal obstruction, epistaxis, mucosal inflammation, sinusitis, and ulceration of the nasal septum and adjacent tissues [2]. The low incidence contributes to diagnostic delays, although time is a critical factor in oncological conditions. As the disease progresses, histopathological verification becomes increasingly challenging due to tissue necrosis and secondary disintegration [1, 3].

The clinical course of the 26-year-old male patient described in this report was consistent with previously published data. His diagnostic process extended over several months, and ENKTL was ultimately identified only after numerous complications that masked the underlying malignancy. The final diagnosis was established following repeated multidisciplinary consultations and extensive deliberation beyond the core expertise of the attending physicians. Given the prolonged and complex diagnostic pathway, a crucial question arises: in the era of available artificial intelligence (AI)-based decision support tools, could such technologies be employed to assist in narrowing the differential diagnosis or at least to generate clinically meaningful guidance toward the final recognition of disease?

AI has been rapidly integrated into various domains of medicine, including communication tools (e.g., chatbots) and decision-support applications that provide diagnostic suggestions based on reported symptoms [4]. Exploring innovative digital approaches to mitigate diagnostic delays seems rational, yet the question remains whether such implementations are sufficiently mature to offer genuine clinical value.

In this context, and based on the authors' direct experience with a diagnostically challenging case, an exploratory assessment was undertaken to examine the performance of publicly accessible AI-driven diagnostic platforms. These are tools that both patients and clinicians might intuitively consult for assistance in identifying the potential causes and nature of a complex clinical presentation. The case – medically intricate and still ongoing – represents not only a clinical struggle for survival but also a real-world test of the diagnostic capabilities of current AI systems. The study aims to evaluate whether contemporary AI platforms can effectively shorten the diagnostic process and serve as practical support in clinical decision-making [5].

AIM

The primary objective of this study was to evaluate whether, and to what extent, currently available AI-based medical platforms are capable of generating diagnostic hypotheses aligned with a patient's actual clinical condition – that is, whether these systems can correctly identify the probable cause of disease and produce a diagnosis consistent with the final hospital-confirmed outcome. The research adopted a theoretical and exploratory design focused on the assessment of online digital health tools utilizing artificial intelligence.

MATERIALS AND METHODS

An anonymized dataset derived from a real clinical case was used, encompassing the course of illness and the final histopathological diagnosis upon which treatment was based.

For the purpose of this study, a set of publicly available AI-powered symptom assessment platforms was selected. These systems are designed to generate probable disease diagnoses based on user-input symptoms. To simulate a realistic patient search behavior, the following query was entered into the Google search engine: "What's wrong with me? – AI." This phrase was chosen to mirror the natural language typically used by patients seeking an explanation for their symptoms using AI tools. The search yielded 15 potential platforms, of which eight were ultimately included based on two criteria: full functionality and accessibility, and free-of-charge availability. The final selection included: Docus (Docus AI, USA), Doctronic (Doctronic Ltd., UK), MedSync (MedSync AI, USA), Symptochat (Symptochat SRL, Romania), UbieHealth (Ubie Inc., Japan), SymptomChecker.io (independent project, producer not specified), Symptomate (Infermedica Sp. z o.o., Poland), and MDandMe (AuxHealth Inc., USA).

Clinical information regarding the patient's medical history was extracted from hospital records and anonymized. To ensure consistency across all tested systems, identical case descriptions were entered into each platform. The dataset was divided into two distinct clinical stages. Early-stage data, symptoms, and findings at initial presentation to an otolaryngologist. Advanced-stage data, complex findings during hospitalization, including complications and advanced medical interventions.

For each platform, the AI system was asked to provide up to five differential diagnoses, allowing standardized comparison of results and establishing clear evaluation criteria for cross-platform analysis. The diagnostic outputs and clinical recommendations generated by each AI platform were analyzed for: accuracy and relevance of suggested diagnoses; alignment with the final hospital-confirmed diagnosis; and quality and depth of the diagnostic reasoning and recommendations provided.

Summary

The anonymized case concerned a 26-year-old male admitted to the Department of Otolaryngology as emergency due to a suspected right-sided peritonsillar abscess. Clinical examination revealed extensive inflammatory infiltration of the lateral pharyngeal wall and right tonsil, accompanied by fever, dysphagia, and severe sinusitis. Despite broad-spectrum antibiotic therapy and abscess drainage, the patient's condition deteriorated, developing hospital-acquired pneumonia and cervical phlegmon, as revealed on CT imaging. Tracheotomy and repeated drainage procedures were performed, followed by wound revision due to bleeding from the tracheostomy site.

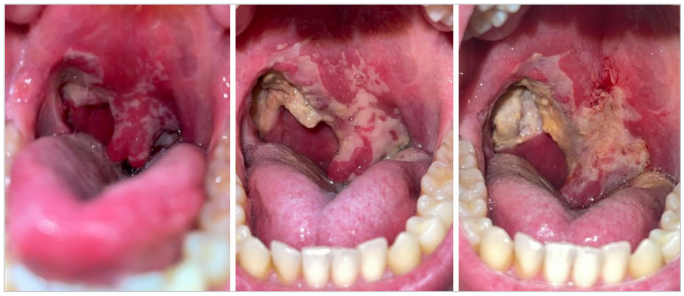

Fig. 1. Ulcerations within the oral cavity and pharynx – evaluation over consecutive weeks.

During hospitalization, high levels of EBV DNA indicated active infectious mononucleosis. The disease course was further complicated by pneumomediastinum and respiratory failure secondary to iatrogenic bronchial perforation, requiring intensive care management. After antifungal and corticosteroid therapy, the patient's condition improved sufficiently for discharge with outpatient follow-up.

Two months later, the patient was readmitted for recurrent sinus inflammation and ulcerative lesions of the pharynx. Despite medical and surgical interventions, necrotic changes of the palate and nasopharynx persisted, accompanied by ongoing EBV viremia (Fig. 1.). Extensive histopathological reevaluation confirmed nasal-type extranodal NK/T-cell lymphoma, stage IV (PET/CT). The patient was subsequently referred to the Department of Hematology and Transplantology, where DDGP chemotherapy was initiated, followed by qualification for allogeneic hematopoietic stem cell transplantation from a related donor.

RESULTS

Of the eight selected AI-driven diagnostic platforms, six generated valid diagnostic outputs consistent with the study design. Docus, Doctronic, Symptomate, and MDandMe provided complete responses with five suggested diagnoses each. MedSync and Symptochat produced partial outputs, whereas UbieHealth and SymptomChecker.io failed to return any results, despite the detailed case descriptions.

Application Performance Summary

Across the active systems, all platforms displayed appropriate data protection notices, including compliance with SOC 2 (System and Organization Controls 2), HIPAA (Health Insurance Portability and Accountability Act), and GDPR (General Data Protection Regulation) standards. Most tools immediately directed the user to an AI-based chat interface conducting a structured medical interview. These interviews generally followed conventional

INDEX COPERNICUS INTERNATIONAL

Tab. I. Summary of platforms used.

	MANUFACTURER / COUNTRY	TYPE OF USER INTERACTION	SCOPE OF DIAGNOSTICS	ADDITIONAL FEATURES	CERTIFICATION / COMPLIANCE	NOTABLE CHARACTERISTICS
Docus	Docus AI Inc., USA	Text-based chat with elements of empathy; conversations in two modes: with or without the patient's medical history	Mainly acute inflammatory conditions, infections, infectious complications	Patient profile, interpretation of test results, consultations with physicians in various specialties	SOC 2, HIPAA, GDPR	High level of personalization; realistic language; empathetic messages ("I'm really sorry")
Doctronic	Doctronic Ltd., UK	Advanced AI chat, SOAP notes, detailed questioning, medication matching	Acute conditions, infections, oral and throat diseases, opportunistic infections	Option to book appointments, explanation of medication mechanisms, quick physician consultation (up to 30 min)	HIPAA	High detail of recommendations; therapeutic and pharmacological components; risk of overestimation of AI role by patients
MedSync	MedSync AI, USA	Simple interface, conversation in SOAP note format	Typical upper respiratory tract infections, basic diagnostics	Short diagnostics, patient education	HIPAA	Simplified structure; accurate basic diagnoses; clear disclaimer that it does not replace a physician
Symptochat	Symptochat SRL, Romania	Text and voice chat (speech recognition)	Infectious diseases, acute conditions (e.g., dyspnea)	Information on diseases, complications, treatments; follow-up questions about medications	No certification data available	Intuitive mobile app; limited personalization; emphasis on acute conditions
MDandMe	AuxHealth Inc., USA	Natural dialogue with AI under the name *Arora*; open conversation without character limit	Chronic diseases, infections, analysis of symptoms and medical history	Test result analysis, health diary, pain intensity assessment (scale 1–10)	HIPAA	High flexibility and personalization; AI verifies understanding of symptoms; contextual analysis
Symptomate	Infermedica Sp. z o.o., Poland	Chat with dynamically generated questions or an extended version (Chat BETA – user's own questions)	Acute and chronic conditions, infections, internal medicine	AI chat BETA version, multilingual availability	Certified Class I Medical Device (EU); HIPAA, GDPR	The only platform with EU certification; strong emphasis on data security and medical quality

medical frameworks, and the final reports were frequently formatted according to the SOAP (Subjective, Objective, Assessment, Plan) documentation standard. Although stylistic differences were observed among the responses, the diagnostic reasoning and suggested management pathways were broadly comparable across systems. In all cases, the platforms advised further evaluation by a specialist clinician. Notably, Docus and Doctronic also offered the possibility of scheduling an online medical consultation. Each diagnostic process began with the collection of basic demographic information: age and sex, followed by symptom-oriented questioning, with varying degrees of interactivity. Most chat-based systems allowed users to freely phrase or redirect the conversation, while MedSync and Symptochat followed a rigid question structure without such flexibility. All platforms provided options for maintaining a personal health record, including data on chronic diseases, medications, and allergies, though the mode of presentation differed substantially among systems (Tab. I.).

Analysis of Diagnostic Suggestions

When presented with the initial symptom set (before hospital admission), the majority of platforms identified peritonsillar abscess and acute tonsillitis as the most probable diagnoses. Additional suggestions included necrotizing ulcerative tonsillitis (Doctronic), parapharyngeal space infection (MDandMe), and secondary bacterial infection following COVID-19 (Docus). Some systems, such as Doctronic, extended the differential diagnosis to include invasive fungal infections or opportunistic infections. Other platforms added inflammatory processes of the paranasal sinuses (Docus, MDandMe, Symptomate) and acute otitis media (MDandMe, Symptomate), which, although relevant, were not directly linked to the final oncological diagnosis.

After introducing the second dataset (hospitalization stage with complications), the differential diagnoses shifted toward chronic, systemic, and invasive conditions. The most frequent suggestions included chronic invasive fungal sinusitis, with aspergillosis or mucormycosis specifically indicated by Doctronic. Several systems proposed granulomatosis with polyangiitis (GPA) and other autoimmune vasculitides or chronic granulomatous diseases associated with immune dysfunction. Multiple platforms (Doctronic, MDandMe, Symptomate) also listed necrotizing fasciitis among potential complications, while Docus mentioned necrotizing ulcerative gingivitis, and MDandMe included malignant otitis media or external ear infection in its differential set (Tab. II.).

Diagnostic and Therapeutic Recommendations

All analyzed AI systems emphasized a dual-track diagnostic approach: rapid identification of acute infection and local complications within the head and neck region, alongside differentiation for chronic, autoimmune, fungal, and malignant etiologies. The depth and prioritization of recommended investigations varied between systems. The extended dialogue concerning diagnostic recommendations was conducted in a non-standardized manner and varied depending on the differential

Tab. II. Comparison of results generated by the platforms.

PLATFORM	FIRST DATA GROUP (BEFORE ADMISSION TO HOSPITAL)	SECOND DATA GROUP (DURING HOSPITALIZATION, WITH THE OCCURRENCE OF COMPLICATIONS)
Docus	1. Peritonsillar abscess	1. Chronic invasive fungal sinusitis
	2. Acute bacterial sinusitis	2. Granulomatosis with vasculitis (Wegener's granulomatosis)
	3. Tonsillitis	3. Drug-resistant bacterial infection
	4. Exacerbation of allergic rhinitis	4. Necrotizing ulcerative gingivitis (NUG)
	5. Secondary bacterial infection after COVID-19	5. Necrotizing fasciitis / deep neck space infection
Doctronic	1. Peritonsillar abscess	1. Invasive fungal infection (e.g., aspergillosis, mucormycosis)
	2. Parapharyngeal space abscess	2. Granulomatosis with vasculitis
	3. Acute tonsillitis	3. Malignant neoplasm (e.g., lymphoma or squamous cell carcinoma)
	4. Necrotizing ulcerative tonsillitis or pharyngitis	4. Chronic granulomatous disease / primary immunodeficiency
	5. Invasive fungal or opportunistic infection	5. Necrotizing fasciitis of the neck
MDandMe	1. Peritonsillar abscess	1. Chronic invasive fungal sinusitis
	2. Acute tonsillitis	2. Necrotizing fasciitis of the pharyngeal region
	3. Acute otitis media	3. Malignant otitis externa or media
	4. Exacerbation of chronic sinusitis	4. Granulomatous disease (e.g., Wegener's granulomatosis)
	5. Parapharyngeal space infection	5. Chronic bacterial superinfection with abscess formation
Symptomate	1. Peritonsillar abscess	1. Chronic invasive fungal sinusitis
	2. Chronic sinusitis with acute exacerbation	2. Necrotizing fasciitis of the head and neck
	3. Acute otitis media	3. Malignant neoplasm (e.g., squamous cell carcinoma)
	4. Allergic rhinitis with secondary infection	4. Autoimmune disease
	5. Tonsillitis or pharyngitis	5. Chronic bacterial infection with multidrug-resistant organisms

diagnoses proposed by each AI system. Each session with a given platform followed a distinct conversational pathway, leading to different diagnostic reasoning trajectories. The recommended investigations differed in their prioritization and level of clinical sophistication. During the initial phase, most platforms (Docus, Doctronic, Symptomate, MDandMe) recommended baseline laboratory tests – complete blood count with differential, C-reactive protein (CRP), and erythrocyte sedimentation rate (ESR) – as well as microbiological cultures from throat swabs and ulcerated lesions for bacterial and fungal pathogens. In cases of atypical or prolonged symptoms, immunological and allergy testing were also suggested. Radiological imaging was consistently prioritized. All major systems advised contrast-enhanced CT of the paranasal sinuses and neck (Docus, Doctronic, MDandMe, Symptomate) to assess for deep neck space infection or phlegmon, with further escalation to MRI, ultrasound, or PET/CT (MDandMe, Doctronic) in persistent or necrotic disease with detectable EBV viremia, to evaluate potential malignancy or disease progression. In severe or atypical courses, Doctronic and Symptomate recommended advanced microbiological workups, including fungal assays (galactomannan, 1,3-β-D-glucan) and tuberculosis cultures, as well as histopathological and microbiological biopsy (Docus, Doctronic, MDandMe, Symptomate). All platforms underscored the relevance of immunological testing, including lymphocyte immunophenotyping (CD3, CD4, CD8, CD19, NK cells), neutrophil function tests (NBT, DHR – Docus, Doctronic), complement system assays (C3, C4, CH50 – Doctronic), autoantibody panels (ANCA, ANA, RF), and quantitative immunoglobulins – particularly in suspected immunodeficiency (Doctronic, MDandMe).

Doctronic, Symptomate, and MDandMe explicitly recommended EBV/CMV/HIV serology, as well as EBV DNA quantification in plasma in cases of persistent necrosis and infiltration, emphasizing the need for oncologic vigilance.

These systems also advocated biopsy of necrotic or ulcerative lesions with full histopathological and immunohistochemical assessment, including EBER in situ hybridization in suspected EBV-associated NK/T-cell lymphoma.

Among all evaluated tools, Doctronic directly included lymphoma as a possible diagnosis; Symptomate (beta version) also recognized the possibility of malignancy and recommended a diagnostic pathway involving MRI, CT, and biopsy.

Although Docus, MDandMe, MedSync, and Symptochat did not explicitly suggest lymphoma, the range of investigations they proposed (e.g., PET/CT and tissue biopsy) could reasonably guide clinicians toward an oncological diagnosis.

DISCUSSION

Numerous digital platforms and applications dedicated to diagnostic

support are currently available in clinical and consumer health contexts. Most existing studies focus on tools such as ChatGPT. However, the number of newly emerging platforms continues to grow, and attention should also be given to those appearing among the top search results for queries such as "What's wrong with me? – AI" or "What disease do I have? – AI" [6–10]. These tools exist in both free and subscription-based formats, offering either basic access or extended professional functionalities. Although originally designed for patients, many of these platforms can serve as adjunctive diagnostic aids for healthcare professionals when faced with diagnostically complex cases. All analyzed systems encouraged users to provide detailed health information, ensuring compliance with current data privacy standards.

Ultimately, six digital platforms were evaluated in this study: Docus, Doctronic, MedSync, Symptochat, MDandMe, and Symptomate. Each platform employed a conversational interface to collect symptom data, generate differential diagnoses, and emphasize the necessity of professional medical consultation. Across all systems, compliance with HIPAA and GDPR regulations was explicitly stated.

Key differences between platforms were observed in the degree of personalization, granularity of diagnostic reasoning, and integration with professional medical services.

During the first stage of testing, all platforms predominantly identified upper respiratory bacterial infections such as peritonsillar abscess, tonsillitis, or sinusitis. After expanding the dataset to include complex clinical findings, the diagnostic suggestions evolved toward chronic, autoimmune, fungal, or malignant pathologies. The most notable increase in diagnostic complexity was observed in Doctronic and Symptomate, both of which included opportunistic infections, autoimmune diseases, and malignant processes in their differential diagnoses. Meanwhile, Docus and MDandMe retained an infectious focus but expanded their differential scope to encompass granulomatous diseases and necrotizing soft tissue infections of the neck. Considering the specific case analyzed, the AI-generated diagnostic reasoning consistently converged around two key pathways: severe infectious complications secondary to immune dysfunction, and potential malignant disease.

The final hospital diagnosis, nasal-type extranodal NK/T-cell lymphoma, confirmed that the underlying malignancy arose within the immune system itself, illustrating the intertwined clinical presentation of infectious and oncological processes. This dual-path diagnostic challenge exemplifies the limitations of current AI systems, which often struggle with cases where symptomatology overlaps across immunological, infectious, and malignant domains. A particularly noteworthy observation was made with the Doctronic system, which explicitly recommended repeat biopsy following an initially negative histopathological result. This aligns with evidence that advanced AI models can incorporate probabilistic reasoning and uncertainty estimation into clinical suggestions – an essential feature for diagnostic safety. However, not all systems demonstrated this capability, highlighting disparities in algorithmic design and the importance of model validation and clinical reliability testing before implementation in healthcare settings.

A comparative analysis between Doctronic and Docus revealed that Doctronic offered greater diagnostic depth, better prioritization of investigations, and a more comprehensive differential list, suggesting superior utility for specialists managing diagnostically ambiguous or rare conditions.

Another key finding was the platforms' oncologic vigilance in response to EBV DNA positivity. Systems that incorporated viral data promptly shifted diagnostic reasoning toward lymphoma-related pathologies, underscoring the high dependence of AI outputs on the completeness and quality of input data. When EBV information was omitted, lymphoma was rarely included in the differential diagnosis. This highlights the critical importance of clinician oversight and precise data entry – without which AI systems may overlook life-threatening conditions.

It is essential to emphasize in studies of this kind that the implementation of artificial intelligence in medicine requires an ethical framework grounded in fairness, transparency, reliability, accountability, privacy, and empathy [11]. Ensuring the confidentiality of patient data is equally crucial, as the risk of unauthorized access can significantly undermine patient trust [12]. The credibility and trustworthiness of AI systems are indispensable prerequisites for encouraging patients to share their data and to engage confidently with such platforms.

CONCLUSIONS

All analyzed AI-based diagnostic tools demonstrated potential value in the preliminary assessment of patient health status. Among them, Docus and Doctronic proved to be the most comprehensive, integrating symptom assessment, diagnostic reasoning, and options for physician consultation within a single platform. Symptomate was distinguished by its adherence to European medical standards and official certification, while MDandMe offered a highly natural conversational flow and advanced personalization of user interaction. Conversely, MedSync and Symptochat served primarily as tools for self-monitoring and health education, providing accessible, easy-to-use interfaces for basic self-assessment.

Across all systems, a consistent diagnostic pattern was observed: confirmation of an infectious etiology coupled with sustained oncologic and immunologic vigilance, particularly in cases of atypical, chronic, or necrotizing disease courses. Importantly, only Doctronic and Symptomate explicitly included malignancy. Meanwhile, Docus and MDandMe, although not explicitly identifying malignancy, proposed a comprehensive set of investigations that would effectively direct clinicians toward the correct oncological diagnosis.

Digital diagnostic platforms open a wide spectrum of opportunities in AI-assisted medicine. Their analytical potential is considerable, and they may serve as practical support tools for both patients and clinicians, particularly in complex or rare disease scenarios where diagnostic uncertainty is high. Nevertheless, at the current stage of technological development, their output should not be interpreted as a definitive diagnosis or used as the sole basis for therapeutic decision-making. Human clinical judgement remains paramount.

REFERENCES

1. Kimura H., de Leval L., Cai Q., Kim W.S.: EBV-associated NK and T-cell lymphoid neoplasms. Curr Opin Oncol., 2022; 34(5): 422–431. doi: 10.1097/CCO.0000000000000889.
2. Schuler A., Smith E., Lowe L., Helfrich Y.: Extranodal natural killer/T-cell lymphoma, nasal type: A rare but critical diagnosis. JAAD Case Rep., 2017; 3(3): 225–227. doi: 10.1016/j.jdcr.2017.02.017.
3. Daines B.S., Varman R.M., Nguyen T.Q.: Extranodal natural killer/T-cell lymphoma nasal type. Proc (Bayl Univ Med Cent)., 2022; 35(5): 672–674. doi: 10.1080/08998280.2022.2071074.
4. Gilbert S., Mehl A., Baluch A. et al.: How accurate are digital symptom assessment apps for suggesting conditions and urgency advice? A clinical vignettes comparison to GPs. BMJ Open., 2020; 10(12): e040269. doi: 10.1136/bmjopen-2020-040269.
5. Singh N., Jain M., Kamal M.M. et al.: Technological paradoxes and artificial intelligence implementation in healthcare. An application of paradox theory. Technol Forecast Soc Change., 2024; 198: 122967. doi: 10.1016/j.techfore.2023.122967.
6. Yalla G.R., Hyman N., Hock L.E. et al.: Performance of Artificial Intelligence Chatbots on Glaucoma Questions Adapted From Patient Brochures. Cureus., 2024; 16(3): e56766. doi: 10.7759/cureus.56766.
7. Jain S., Naicker D., Raj R. et al. Computational Intelligence in Cancer Diagnostics: A Contemporary Review of Smart Phone Apps, Current Problems, and Future Research Potentials. Diagnostics (Basel)., 2023; 13(9): 1563. doi: 10.3390/diagnostics13091563.
8. Al-Ashwal F.Y., Zawiah M., Gharaibeh L., Abu-Farha R., Bitar A.N.: Evaluating the Sensitivity, Specificity, and Accuracy of ChatGPT-3.5, ChatGPT-4, Bing AI, and Bard Against Conventional Drug-Drug Interactions Clinical Tools. Drug Healthc Patient Saf., 2023; 15: 137–147. doi: 10.2147/DHPS.S425858.
9. Hirosawa T., Kawamura R., Harada Y. et al.: ChatGPT-Generated Differential Diagnosis Lists for Complex Case-Derived Clinical Vignettes: Diagnostic Accuracy Evaluation. JMIR Med Inform., 2023; 11: e48808. doi: 10.2196/48808.
10. Chu B., Modi N.D., Menz B.D. et al.: Generative AI's healthcare professional role creep: a cross-sectional evaluation of publicly accessible, customised health-related GPTs. Front Public Health., 2025; 13: 1584348. doi: 10.3389/fpubh.2025.1584348.
11. Siala H., Wang Y.: SHIFTing artificial intelligence to be responsible in healthcare: A systematic review. Soc Sci Med., 2022; 296: 114782. doi: 10.1016/j.socscimed.2022.114782.
12. Wiljer D., Hakim Z.: Developing an Artificial Intelligence-Enabled Health Care Practice: Rewiring Health Care Professions for Better Care. J Med Imaging Radiat Sci., 2019; 50(4 Suppl 2): S8–S14. doi: 10.1016/j.jmir.2019.09.010.