Adv Ther (2025) 42:2559–2568
https://doi.org/10.1007/s12325-025-03198-4

COMMENTARY

The Limitations of Artificial Intelligence in Head and Neck Oncology

Karthik N. Rao · Veronica Fernandez-Alvarez · Orlando Guntinas-Lichius · M. P. Sreeram · Remco de Bree · Luiz P. Kowalski · Arlene Forastiere · Pia Pace-Asciak · Juan P. Rodrigo · Nabil F. Saba · Ohad Ronen · Ewa Florek · Gregory W. Randolph · Alvaro Sanabria · Jan B. Vermorken · Ehab Y. Hanna · Alfio Ferlito

Received: January 27, 2025 / Accepted: April 4, 2025 / Published online: April 29, 2025

ABSTRACT

Artificial intelligence (AI) is revolutionizing head and neck oncology, offering innovations in tumor detection, treatment planning, and patient management. However, its integration into clinical practice is hindered by several limitations. These include clinician mistrust due to a lack of understanding of AI mechanisms, biases in algorithm development, and the potential over-reliance on technology, which may undermine clinical expertise. Data-related challenges, such as inconsistent quality and limited representativeness of datasets, further complicate AI's application. Ethical, legal, and privacy concerns also pose significant barriers. Addressing these

This paper was written by members and invitees of the International Head and Neck Scientific Group.

K. N. Rao (✉) · M. P. Sreeram
Department of Head and Neck Oncology, Sri Shankara Cancer Foundation, Bangalore, India
e-mail: karthik.nag.rao@gmail.com

V. Fernandez-Alvarez
Department of Vascular and Endovascular Surgery, Hospital Universitario Central de Asturias, Oviedo, Spain

O. Guntinas-Lichius
Department of Otorhinolaryngology, Jena University Hospital, Jena, Germany

R. de Bree
Department of Head and Neck Surgical Oncology, University Medical Center Utrecht, Utrecht, The Netherlands

L. P. Kowalski
Head and Neck Surgery and Otorhinolaryngology Department, A C Camargo Cancer Center, Sao Paulo, Brazil

L. P. Kowalski
Head and Neck Surgery, Faculty of Medicine, University of Sao Paulo, Sao Paulo, Brazil

A. Forastiere
The Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, Baltimore, USA

P. Pace-Asciak
Department of Otolaryngology, Head and Neck Surgery, University of Toronto, Toronto, ON, Canada

J. P. Rodrigo
Instituto Universitario de Oncología del Principado de Asturias, University of Oviedo, Oviedo, Spain

J. P. Rodrigo
Department of Otolaryngology, Hospital Universitario Central de Asturias, Oviedo, Spain

J. P. Rodrigo
CIBERONC, Madrid, Spain

N. F. Saba
Department of Hematology and Medical Oncology, Winship Cancer Institute, Emory University School of Medicine, Atlanta, Georgia

issues through transparent AI systems, clinician education, and clear regulations is essential for ensuring responsible, equitable use in head and neck oncology. This manuscript explores the limitations of AI in head and neck oncology.

Keywords: Artificial intelligence; Head and neck neoplasms; Limitations; Algorithm bias; Clinical decision support systems

Key Summary Points

Artificial intelligence (AI) algorithms in head and neck oncology often function as"black boxes,"making it difficult for clinicians to understand their decision-making processes, leading to skepticism and reduced trust in their reliability.

Bias in AI algorithms arises from non-representative datasets and unconscious prejudices, leading to disparities in care, particularly for under-represented populations such as women and low-resource settings.

Over-reliance on AI may erode clinical expertise, as healthcare professionals might depend too heavily on automated insights, potentially overlooking critical contextual factors and nuanced patient-specific considerations.

High-quality, annotated datasets required for AI training are often lacking, with variations in imaging protocols and clinical practices posing challenges to the generalizability and reliability of AI models across diverse populations.

The financial burden of AI implementation, including infrastructure, validation, and ongoing updates, restricts access to well-funded institutions, exacerbating healthcare disparities in resource-limited settings.

THE LIMITATIONS OF ARTIFICIAL INTELLIGENCE IN HEAD AND NECK ONCOLOGY

Artificial intelligence (AI) has made transformative inroads into healthcare, particularly in oncology, where its ability to analyze large datasets and provide actionable insights has opened new avenues for diagnosis, treatment, and prognosis. In head and neck oncology, AI applications have demonstrated potential in areas such as tumor detection, classification, treatment planning, survival prediction and patient education [1–3]. However, the practical integration of AI into this complex field is not without challenges. While its promise is undeniable, there are limitations

O. Ronen
Department of Otolaryngology, Head and Neck Surgery, Galilee Medical Center, Affiliated with Azrieli Faculty of Medicine, Bar-Ilan University, Safed, Israel

E. Florek
Laboratory of Environmental Research, Department of Toxicology, Poznan University of Medical Sciences, 60-806 Poznan, Poland

G. W. Randolph
Massachusetts Eye and Ear Infirmary, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

A. Sanabria
Department of Surgery, School of Medicine, Universidad de Antioquia/Hospital Universitario San Vicente Fundación-CEXCA Centro de Excelencia en Enfermedades de Cabeza y Cuello, Medellín, Colombia

J. B. Vermorken
Department of Medical Oncology, Antwerp University Hospital, Edegem, Belgium

J. B. Vermorken
Faculty of Medicine and Health Sciences, University of Antwerp, Antwerp, Belgium

E. Y. Hanna
Department of Head and Neck Surgery, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

A. Ferlito
Coordinator of the International Head and Neck Scientific Group, 35100 Padua, Italy

that should inform its developing utilization and necessitate a balanced and cautions approach.

LACK OF UNDERSTANDING AND TRUST

The significant barriers to the adoption of AI in head and neck oncology are the possible inconsistencies in the data used to construct the instruments and the limited understanding of its mechanisms among healthcare professionals. The inherent complexity of AI technology often creates a disconnect for clinicians without a technical background, fostering skepticism and mistrust regarding its reliability [4]. Many AI algorithms operate as "black boxes," generating outputs without providing transparent reasoning or explanations, which makes their integration into clinical workflows difficult where trust and explainability are essential [5]. It should also be noted that the prompt used is of great importance [6]. Small changes in prompts and different prompt techniques can generate different responses [7, 8]. The generative AI has a tendency to confabulate, filling gaps with jargon when it lacks definitive answers, and its overconfidence in these responses undermines reliability in clinical decision-making [9]. For clinicians, comprehension of the decision-making process behind AI models is critical to ensure informed clinical decisions. The absence of such transparency results in reluctance to utilize AI tools, as clinicians may fear potential errors or biases that could adversely affect patient care [10]. Additionally, the lack of comprehensive education on AI principles, capabilities, and limitations across medical disciplines exacerbates this hesitation. This gap sustains resistance to AI adoption, impeding the integration of these innovative technologies into routine clinical practice [11]. An emerging solution to tackle the black box problem in AI is Explainable AI (XAI) [12, 13].

Implementing XAI can enhance transparency by providing interpretable outputs, increasing clinician trust in AI-driven decisions. Additionally, integrating AI education into medical training can improve understanding, reducing skepticism and fostering safer adoption in clinical practice.

BIAS IN AI ALGORITHMS

The risk of bias presents a critical limitation in the application of AI in head and neck oncology, with profound implications for equitable utilization. Bias can arise from various sources, including the unconscious prejudices of developers, the representativeness of training datasets, and inherent flaws in algorithmic design [14]. In head and neck oncology, these biases can exacerbate existing disparities, leading to skewed outcomes. For example, training datasets often over-represent individuals from specific geographic, ethnic, or socioeconomic groups, limiting the generalizability of the algorithm. An AI model trained predominantly on Western populations may perform poorly in predicting outcomes or recommending treatments for patients in low- and middle-income countries, where the prevalence, presentation, and management of head and neck cancers differ significantly [15]. It should be noted that external validation is an issue that needs to be addressed whenever clinical trial results are considered in patient management, even without AI involvement [16]. Furthermore, imaging data in head and neck oncology presents additional challenges, as variability in imaging protocols, scanner technologies, and operator expertise can compromise the consistency and quality of training data [17]. Additionally, the under-representation of women in datasets should be considered within the bias segment. While geographic, ethnic, and socioeconomic factors are often discussed, the predominance of male participants in clinical trials and medical studies remains a significant issue [18, 19]. The absence or limited presence of sex-disaggregated data can result in skewed results, highlighting the need for balanced representation in study populations [20, 21]. Acknowledging lead time and over-diagnosis bias is also essential for the accuracy of all prognostic deep learning models [22]. These issues are amplified when there are concerns about the validity of research findings, increasing the risk of diagnostic inaccuracies and flawed treatment guidance when AI is used across diverse clinical settings [7, 9, 23].

Ensuring diverse, representative datasets and implementing bias-detection tools can mitigate AI bias in head and neck oncology. Rigorous external validation and fairness-aware algorithms enhance reliability. Collaboration between developers, clinicians, and policymakers is key to equitable AI application.

OVER-RELIANCE ON AI AND THE EROSION OF HUMAN EXPERTISE

The increasing integration of AI in head and neck oncology raises significant concerns about potential over-reliance on technology at the expense of clinical expertise. While AI demonstrates remarkable capabilities in pattern recognition and data processing, it lacks the contextual understanding and adaptive judgment that experienced clinicians bring to patient care [24]. Over-dependence on AI could erode critical thinking and decision-making skills among healthcare providers, particularly in high-stakes scenarios like addressing intra-operative complications or managing unexpected treatment outcomes [25]. This over-reliance or cognitive bias is important in research and clinical decision-making. AI generated information should complement, not supplant, clinical judgement [26]. These situations often demand a level of intuition and flexibility that AI systems cannot replicate [27]. Furthermore, the limitations of AI being bound by its programming and the constraints of its training data make it ill-suited for handling atypical cases or novel clinical presentations [28]. Not all considerations taken into account during multidisciplinary team meetings are measurable, and the nuances of specific patient characteristics are of utmost importance when tailoring a unique treatment plan for a patient [29, 30]. AI should be regarded as an additional support tool rather than a replacement for clinical decision-making by qualified experts. Ultimately, it is the medical professional who bears the legal and ethical responsibility for decisions that can have significant consequences for patient outcomes.

A potential solution to mitigate over-reliance on AI is to implement a hybrid decision-making model where AI serves as an adjunct rather than a replacement for clinical expertise. Regular AI–human cross-validation, mandatory clinician oversight, and continuous medical education on AI limitations can help maintain critical thinking and clinical judgment. Additionally, structured AI audit systems can ensure that AI recommendations are reviewed and validated before influencing key treatment decisions.

DATA-RELATED CHALLENGES

The success of AI in head and neck oncology relies heavily on the availability, quality, and representativeness of data, yet several critical challenges limit progress in this domain. Robust AI models require large, high-quality, and annotated datasets tailored to head and neck oncology [31]. Unfortunately, the availability of such datasets is often limited, hindering the development of AI systems that can generalize effectively across diverse patient populations and clinical scenarios [32]. Datasets are often marked by significant variability in imaging modalities, clinical practices, and patient demographics. This heterogeneity complicates the development of standardized AI algorithms, making it challenging to achieve consistent performance across different healthcare settings [33]. Moreover, the reliability of ground truth data is essential. Sometimes, a real gold standard is lacking, but a high variability in annotation may also exist. A survey among pathologists showed a heterogeneous landscape of annotation practices with minimal standardization in computational pathology [34]. Incomplete or poorly annotated datasets undermine the reliability of AI models, increasing the risk of inaccurate predictions and potentially leading to suboptimal clinical outcomes [35]. Ensuring consistent data quality is paramount for the safe and effective implementation of AI in clinical practice.

Another challenge to adopting AI algorithms for clinical evaluation is the lack in most cases of independent or external validation. In the few studies where external validations were conducted, several concerns emerged. These included issues such as the similarity of external datasets, the minimum dataset size required for validation, acceptable performance metrics, and the specific procedures used for the validation process, whether truly independent or otherwise [31].

Multi-institutional collaborations and federated learning can enhance data availability and standardization. Uniform annotation guidelines and AI-driven harmonization improve dataset consistency. Mandatory external validation ensures AI model reliability across diverse clinical settings.

COST AND ACCESSIBILITY

The widespread adoption of AI in head and neck oncology may be limited to well-funded institutions, leaving underserved populations at a disadvantage. Mainly due to the cost, implementing AI technologies in clinical settings remains a significant barrier. Developing, validating, and deploying AI algorithms require substantial financial resources, technical expertise, and infrastructure [36–38]. These costs can be prohibitive for smaller healthcare institutions or those in resource-limited settings, exacerbating disparities in access to advanced cancer care. Moreover, the maintenance and updating of AI systems to reflect new evidence and guidelines further contribute to their cost [39].

Cloud-based AI models, public–private partnerships, and open-source frameworks can reduce costs and improve accessibility for resource-limited institutions.

ETHICAL AND LEGAL CONSIDERATIONS

The integration of AI presents complex ethical and legal challenges. Key issues include informed consent, data ownership, algorithm accountability, and liability for errors [39, 40]. For instance, if an AI system provides incorrect treatment recommendations, it is unclear whether the responsibility lies with the clinician using the system, the healthcare institution implementing it, or the developers who designed it [41]. It is important to highlight the difficulty in challenging AI results and its system. For AI to succeed, mechanisms for challenging both the results and the methods must be readily available and widely accessible [42, 43]. These ambiguities highlight the need for creation and implementation clear regulatory frameworks to define accountability and to mitigate risks [43].

Ethical concerns also arise from its potential to influence critical decisions, such as treatment eligibility or prognosis, which can profoundly affect patient outcomes. Ensuring transparency, fairness, and accountability is paramount to avoid biases or inequities in care delivery. Furthermore, patient privacy regulations like General Data Protection Regulation (GDPR) and the Health Insurance Portability and Accountability Act (HIPAA), while vital for safeguarding data, pose significant challenges by limiting data sharing and multi-institutional collaborations [44]. Moreover, AI software is considered to be a medical device, potentially hindering medical device development, launch, and marketing brought by specific regulations [45]. The approval of AI-enabled medical devices varies globally, with the United States Food and Drug Administration (FDA) overseeing the USA, the National Medical

Products Administration (NMPA) regulating China, and the European Union (EU) relying on Notified Bodies for Conformité Européenne (CE) marking. To address AI-enabled devices, authorities have introduced specific regulations and guidelines, reflecting a universal challenge [46]. This restriction hinders the creation of comprehensive, diverse datasets necessary for training robust AI models and AI based devices. In this scenario, utilizing real-world data that reflect everyday clinical practices may be significantly more reliable than relying on potentially biased literature derived from successful case series and positive clinical trials. The FDA has begun refining regulations for approving AI tools, recognizing challenges beyond training datasets, such as continuous learning from new patient data. This raises critical questions about when significant updates warrant re-approval [47]. Addressing these ethical and legal concerns is essential for the responsible and equitable deployment of AI in clinical practice.

An additional factor to consider is the possibility that AI systems are susceptible to cyber attacks, which can compromise patient data and the integrity of medical decision-making. Threats such as data re-identification, model extraction, and adversarial attacks highlight the vulnerabilities of these technologies [48].

Clear regulatory frameworks should define accountability, data ownership, and liability in AI use in healthcare. AI models must ensure transparency with explainability features and mechanisms to challenge outputs. Standardized global guidelines and cyber security measures are essential for ethical AI deployment.

EVOLVING CANCER BIOLOGY

Despite significant advancements in AI, the incomplete understanding of cancer biology remains a critical challenge in head and neck oncology. This limitation hinders the development of AI models capable of reliably predicting patient outcomes or of determining optimal treatment strategies [49]. AI algorithms are fundamentally reliant on the quality and scope of the data they are trained on, and they lack the ability to infer biological mechanisms or relationships that are not explicitly represented within these datasets [50].

Head and neck cancers have extensive heterogeneity in genetic, molecular, and clinical features, which exemplifies this complexity [51]. Accurately modeling such diversity demands an intricate understanding of the underlying disease biology, which continues to evolve. The inability to capture this dynamic complexity in AI frameworks underscores the importance of advancing cancer biology research in tandem with AI innovation to ensure that these technologies can be applied effectively and responsibly in clinical settings [52].

Integrating multi-omics data with AI can enhance modeling of cancer heterogeneity. Collaboration with cancer biologists will refine algorithms to reflect evolving disease mechanisms. Generation 2 AI systems incorporate advanced self-learning and adaptability, allowing continuous refinement from new data and clinical experiences. These models can dynamically adjust to evolving cancer biology, improving predictive accuracy and treatment recommendations.

PATH FORWARD: ADDRESSING LIMITATIONS

To fully harness the potential of AI in head and neck oncology, several strategies are essential. Education and training programs for clinicians and healthcare professionals can bridge the gap between technical expertise and medical practice, fostering better understanding of AI capabilities, limitations, and applications [53]. By encouraging interdisciplinary collaboration, these programs can help build trust in AI tools. Efforts to mitigate bias in AI algorithms are also crucial for ensuring equity in cancer care, which can be achieved by using diverse and representative datasets in training, along with regular audits and independent validations to identify and address biases [54]. Transparency and explainability in AI systems are vital for gaining the trust of both clinicians and patients. Adopting clear methodologies and open reporting standards will make AI tools more interpretable and accountable, facilitating their integration into clinical workflows [55]. To safeguard against this opacity of black-box models in the integration of AI into clinical workflows, the integration of XAI technologies should be implemented. These models can present the reasoning behind decisions, enabling clinicians to assess their reliability and patients to make informed choices [12, 13, 56]. In head and neck oncology, integrating XAI could help clinicians understand the rationale behind predictions, such as tumor grading or treatment response, fostering greater acceptance of these technologies [48, 49, 57]. Furthermore, establishing regulatory frameworks to address ethical and legal challenges, such as data privacy, algorithm accountability, and liability, is necessary to ensure responsible use of AI in oncology [58]. The use of cryptographic techniques, such as homomorphic encryption and secure multiparty computation, allow data sharing without exposing raw patient data, safeguarding patient privacy during algorithm development [48]. Multi-institutional collaborations with clinical experts also play a critical role in advancing AI by enabling the creation of large, high-quality datasets, sharing of resources, and pooling of knowledge [59]. The success of AI in clinical practice hinges on the quality of the data used to train algorithms and the rigor of their evaluation. Current studies often rely on single-center datasets with limited external validation, compromising the generalizability of AI models. Establishing standardized protocols for data curation and adherence to reporting frameworks like "Prediction model Risk Of Bias ASsessment Tool" (PROBAST) [60], which is a tool designed to assess the risk of bias in studies that develop or validate predictive models, and "Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis" (TRIPOD) [61], which is a set of guidelines that ensures transparent reporting of multivariable predictive models, can significantly enhance the reliability of AI research.

These collaborative efforts are key to ensuring the equitable and responsible implementation of AI, which can ultimately drive innovation and improve patient outcomes in head and neck oncology.

CONCLUSION

AI holds a significant promise for transforming head and neck oncology practice worldwide, offering new tools to support specialists in diagnosis, treatment planning, and patient management. However, its integration into clinical practice is fraught with challenges, from biases and data limitations to ethical and legal concerns. Addressing these limitations requires a concerted effort from researchers, clinicians, and policymakers to ensure that AI is deployed responsibly and equitably. With ongoing advancements in AI and cautious clinical integration, it has the potential to enhance clinical decision-making and patient outcomes.

Author Contributions. Conceptualization: Karthik N Rao; Data collection: Karthik N Rao, Sreeram M.P; Data synthesis: Karthik N Rao, Sreeram M.P; Data analysis: Karthik N Rao; Drafting the article: Karthik N Rao; Critical revision of the article: Karthik N Rao, Veronica Fernandez-Alvarez, Orlando Guntinas-Lichius, Sreeram M.P, Remco de Bree, Luiz P. Kowalski, Arlene Forastiere, Pia Pace-Asciak, Juan P Rodrigo, Nabil F Saba, Ohad Ronen, Ewa Florek, Gregory W Randolph, Alvaro Sanabria, Jan B Vermorken, Ehab Y Hanna, Alfio Ferlito; Final approval of the version to be published: Karthik N Rao, Veronica Fernandez-Alvarez, Orlando Guntinas-Lichius, Sreeram M.P, Remco de Bree, Luiz P. Kowalski, Arlene Forastiere, Pia Pace-Asciak, Juan P Rodrigo, Nabil F Saba, Ohad Ronen, Ewa Florek, Gregory W Randolph, Alvaro Sanabria, Jan B Vermorken, Ehab Y Hanna, Alfio Ferlito.

Funding. No funding or sponsorship was received for this study or publication of this article.

Data Availability. Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Declarations

Conflict of Interest. Karthik N Rao, Veronica Fernandez-Alvarez, Orlando Guntinas-Lichius, Sreeram M.P, Remco de Bree, Luiz P. Kowalski, Arlene Forastiere, Pia Pace-Asciak, Juan P Rodrigo, Nabil F Saba, Ohad Ronen, Ewa Florek, Gregory W Randolph, Alvaro Sanabria, Jan B Vermorken, Ehab Y Hanna, and Alfio Ferlito have nothing to disclose.

Ethical Approval. This article is based on previously conducted studies and does not contain any new studies with human participants or animals performed by any of the authors.

REFERENCES

1. Luchini C, Pea A, Scarpa A. Artificial intelligence in oncology: current applications and future perspectives. Br J Cancer. 2022;126(1):4–9.

2. Lee JC, Hamill CS, Shnayder Y, Buczek E, Kakarala K, Bur AM. Exploring the role of artificial intelligence chatbots in preoperative counseling for head and neck cancer surgery. Laryngoscope. 2024;134(6):2757–61.

3. Broggi G, Maniaci A, Lentini M, Palicelli A, Zanelli M, Zizzo M, et al. Artificial intelligence in head and neck cancer diagnosis: a comprehensive review with emphasis on radiomics, histopathological, and molecular applications. Cancers. 2024;16(21):3623.

4. Hassan M, Kushniruk A, Borycki E. Barriers to and facilitators of artificial intelligence adoption in health care: scoping review. JMIR Hum Factors. 2024;29(11): e48633.

5. Kiseleva A, Kotzinos D, De Hert P. Transparency of AI in healthcare as a multilayered system of accountabilities: between legal requirements and technical limitations. Front Artif Intell. 2022;30(5): 879603.

6. Marchi F, Bellini E, Iandelli A, Sampieri C, Peretti G. Exploring the landscape of AI-assisted decision-making in head and neck cancer treatment: a comparative analysis of NCCN guidelines and ChatGPT responses. Eur Arch Oto-Rhino-Laryngol. 2024;281(4):2123–36.

7. Rao KN, Arora RD, Dange P, Nagarkar NM. NLP AI models for optimizing medical research: demystifying the concerns. Indian J Surg Oncol. 2023;14(4):854–8.

8. Bansal P. Prompt engineering importance and applicability with generative AI. J Comput Commun. 2024;12(10):14–23.

9. Neha F, Bhati D, Shukla DK, Amiruzzaman M. ChatGPT: transforming healthcare with AI. AI. 2024;5(4):2618–50.

10. Khan B, Fatima H, Qureshi A, Kumar S, Hanan A, Hussain J, et al. Drawbacks of artificial intelligence and their potential solutions in the healthcare sector. Biomed Mater Devices N Y N. 2023;8:1–8.

11. Ahmed MI, Spooner B, Isherwood J, Lane M, Orrock E, Dennison A. A systematic review of the barriers to the implementation of artificial intelligence in healthcare. Cureus. 2023;15(10):e46454.

12. Shephard AJ, Bashir RMS, Mahmood H, Jahanifar M, Minhas F, Raza SEA, et al. A fully automated and explainable algorithm for predicting malignant transformation in oral epithelial dysplasia. NPJ Precis Oncol. 2024;28(8):137.

13. Dörrich M, Hecht M, Fietkau R, Hartmann A, Iro H, Gostian AO, et al. Explainable convolutional neural networks for assessing head and neck cancer histopathology. Diagn Pathol. 2023;18(1):121.

14. Ferrara E. Fairness and bias in artificial intelligence: a brief survey of sources, impacts, and mitigation strategies. Sci. 2024;6(1):3.

15. Chen RJ, Wang JJ, Williamson DFK, Chen TY, Lipkova J, Lu MY, et al. Algorithm fairness in artificial intelligence for medicine and healthcare. Nat Biomed Eng. 2023;7(6):719–42.

16. Dekkers OM, von Elm E, Algra A, Romijn JA, Vandenbroucke JP. How to assess the external validity of therapeutic trials: a conceptual approach. Int J Epidemiol. 2010;39(1):89–94.

17. Willemink MJ, Koszek WA, Hardell C, Wu J, Fleischmann D, Harvey H, et al. Preparing medical imaging data for machine learning. Radiology. 2020;295(1):4–15.

18. Merone L, Tsey K, Russell D, Nagle C. Sex inequalities in medical research: a systematic scoping review of the literature. Womens Health Rep. 2022;3(1):49–59.

19. Daitch V, Turjeman A, Poran I, Tau N, Ayalon-Dangur I, Nashashibi J, et al. Underrepresentation of women in randomized controlled trials: a systematic review and meta-analysis. Trials. 2022;23(1):1038.

20. Straw I, Wu H. Investigating for bias in healthcare algorithms: a sex-stratified analysis of supervised machine learning models in liver disease prediction. BMJ Health Care Inform. 2022;29(1):e100457.

21. Peters SAE, Woodward M. A roadmap for sex- and gender-disaggregated health research. BMC Med. 2023;13(21):354.

22. Bretthauer M, Løberg M, Holme Ø, Adami HO, Kalager M. Deep learning and cancer biomarkers: recognising lead-time bias. The Lancet. 2021;397(10270):194.

23. Özer M. Is artifical intelligence hallucinating? Turk J Psychiatry. 2024;35(4):333–5.

24. Khosravi M, Zare Z, Mojtabaeian SM, Izadi R. Artificial intelligence and decision-making in healthcare: a thematic analysis of a systematic review of reviews. Health Serv Res Manag Epidemiol. 2024;5(11):23333928241234864.

25. Zhai C, Wibowo S, Li LD. The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. Smart Learn Environ. 2024;11(1):28.

26. Zwaan L. Cognitive bias in large language models: implications for research and practice. NEJM AI. 2024;1(12):AIe2400961.

27. Ahmad A, Premanandan S, Cajander Å, Langegård U, Uereten E, Tiblom EY. A qualitative study with informal caregivers and healthcare professionals for individuals with head and neck cancer on the usage of AI Chatbots. Stud Health Technol Inform. 2024;22(316):751–5.

28. Siala H, Wang Y. SHIFTing artificial intelligence to be responsible in healthcare: a systematic review. Soc Sci Med. 2022;1(296): 114782.

29. Ronen O, Robbins KT, de Bree R, Guntinas-Lichius O, Hartl DM, Homma A, et al. Standardization for oncologic head and neck surgery. Eur Arch Oto-Rhino-Laryngol. 2021;278(12):4663–9.

30. Shellenberger TD, Weber RS. Multidisciplinary team planning for patients with head and neck cancer. Oral Maxillofac Surg Clin N Am. 2018;30(4):435–44.

31. Mäkitie AA, Alabi RO, Ng SP, Takes RP, Robbins KT, Ronen O, et al. Artificial intelligence in head and neck cancer: a systematic review of systematic reviews. Adv Ther. 2023;40(8):3360–80.

32. Temple SWP, Rowbottom CG. Gross failure rates and failure modes for a commercial ai-based auto-segmentation algorithm in head and neck cancer patients. J Appl Clin Med Phys. 2024;25(6): e14273.

33. Najjar R. Redefining radiology: a review of artificial intelligence integration in medical imaging. Diagnostics. 2023;13(17):2760.

34. Montezuma D, Oliveira SP, Tolkach Y, Boor P, Haragan A, Carvalho R, et al. Annotation practices in computational pathology: a European Society of Digital and Integrative Pathology (ESDIP) survey study. Lab Investig J Tech Methods Pathol. 2024;28: 102203.

35. Aldoseri A, Al-Khalifa KN, Hamouda AM. Re-thinking data strategy and integration for artificial intelligence: concepts, opportunities, and challenges. Appl Sci. 2023;13(12):7082.

36. Das SK, Dasgupta RK, Roy SD, Shil D. AI in Indian healthcare: from roadmap to reality. Intell Pharm. 2024;2(3):329–34.

37. Maleki Varnosfaderani S, Forouzanfar M. The role of AI in hospitals and clinics: transforming healthcare in the 21st Century. Bioengineering. 2024;11(4):337.

38. Petersson L, Larsson I, Nygren JM, Nilsen P, Neher M, Reed JE, et al. Challenges to implementing artificial intelligence in healthcare: a qualitative interview study with healthcare leaders in Sweden. BMC Health Serv Res. 2022;22(1):850.

39. Brady AP, Allen B, Chong J, Kotter E, Kottler N, Mongan J, et al. Developing, purchasing, implementing and monitoring AI tools in radiology: practical considerations. A multi-society statement from the ACR, CAR, ESR, RANZCR and RSNA. Radiol Artif Intell. 2024;6(1):e230513.

40. Rodrigues R. Legal and human rights issues of AI: gaps, challenges and vulnerabilities. J Responsible Technol. 2020;1(4): 100005.

41. Habli I, Lawton T, Porter Z. Artificial intelligence in health care: accountability and safety. Bull World Health Organ. 2020;98(4):251–6.

42. Ledro C, Nosella A, Dalla PI. Integration of AI in CRM: challenges and guidelines. J Open Innov Technol Mark Complex. 2023;9(4): 100151.

43. Gilbert S, Kather JN. Guardrails for the use of generalist AI in cancer care. Nat Rev Cancer. 2024;24(6):357–8.

44. Gerke S, Minssen T, Cohen G. Ethical and legal challenges of artificial intelligence-driven healthcare. Artificial intelligence in healthcare. New York: Elsevier; 2020. p. 295–336.

45. Svempe L. Exploring impediments imposed by the medical device regulation EU 2017/745 on software as a medical device. JMIR Med Inform. 2024;5(12): e58080.

46. Liu Y, Yu W, Dillon T. Regulatory responses and approval status of artificial intelligence medical devices with a focus on China. NPJ Digit Med. 2024;18(7):255.

47. Center for Devices and Radiological Health. Artificial Intelligence and Machine Learning in Software as a Medical Device. FDA [Internet]. 2025 Jan 6 [cited 2025 Jan 16]; Available from: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device

48. Alonso A, Siracuse JJ. Protecting patient safety and privacy in the Era of artificial intelligence. Semin Vasc Surg. 2023;36(3):426–9.

49. Pham TD, Teh MT, Chatzopoulou D, Holmes S, Coulthard P. Artificial intelligence in head and neck cancer: innovations, applications, and future directions. Curr Oncol. 2024;31(9):5255–90.

50. Barbierato E, Gatti A. The challenges of machine learning: a critical review. Electronics. 2024;13(2):416.

51. Deshmukh A, Rao KN, Arora RD, Nagarkar NM, Singh A, Shetty OS. Molecular insights into oral malignancy. Indian J Surg Oncol. 2022;13(2):267–80.

52. Verlingue L, Boyer C, Olgiati L, Brutti Mairesse C, Morel D, Blay JY. Artificial intelligence in oncology: ensuring safe and effective integration of language models in clinical practice. Lancet Reg Health - Eur. 2024;1(46): 101064.

53. Sorte SR, Rawekar A, Rathod SB. Understanding AI in healthcare: perspectives of future healthcare professionals. Cureus. 2024;16(8):e66285.

54. Ueda D, Kakinuma T, Fujita S, Kamagata K, Fushimi Y, Ito R, et al. fairness of artificial intelligence in healthcare: review and recommendations. Jpn J Radiol. 2024;42(1):3–15.

55. Karalis VD. The integration of artificial intelligence into clinical practice. Appl Biosci. 2024;3(1):14–44.

56. Alonso C, Berg A, Kothari S, Papageorgiou C, Rehman S. Will the AI revolution cause a great divergence? J Monet Econ. 2022;1(127):18–37.

57. Asaadi S, Martins KN, Lee MM, Pantoja JL. Artificial intelligence for the vascular surgeon. Semin Vasc Surg. 2023;36(3):394–400.

58. Palaniappan K, Lin EYT, Vogel S. Global regulatory frameworks for the use of artificial intelligence (AI) in the healthcare services sector. Healthcare. 2024;12(5):562.

59. Broekhuizen T, Dekker H, de Faria P, Firk S, Nguyen DK, Sofka W. AI for managing open innovation: opportunities, challenges, and a research agenda. J Bus Res. 2023;1(167): 114196.

60. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med. 2019;170(1):51–8.

61. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;16(385): e078378.