This is a repository copy of *Exploring the role of artificial intelligence in oral pathology: diagnostic and prognostic implications*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/228742/

Version: Accepted Version

Article:

Alsanie, I.S. orcid.org/0000-0003-3037-9181, Qannam, A., Bello, I.O. orcid.org/0000-0003-2085-3185 et al. (1 more author) (2025) Exploring the role of artificial intelligence in oral pathology: diagnostic and prognostic implications. Journal of Oral Pathology & Medicine. ISSN 0904-2512

https://doi.org/10.1111/jop.70002

© 2025 The Authors. Except as otherwise noted, this author-accepted version of a journal article published in Journal of Oral Pathology & Medicine is made available via the University of Sheffield Research Publications and Copyright Policy under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Exploring the Role of Artificial Intelligence in Oral Pathology: Diagnostic and Prognostic Implications

Running title: AI Implications in Oral Pathology

Keywords: Artificial Intelligence, Oral Pathology, Computational Pathology, Diagnostic and Prognostic Analysis.

Ibrahim S. Alsanie,[1] Ahmed Qannam,[1] Ibrahim O. Bello, [1,2] Syed Ali Khurram.[3]

[1] Department of Oral Medicine and Diagnostic Sciences, College of Dentistry, King Saud University, Riyadh, Saudi Arabia.

[2] Department of Pathology, Haartman Institute, University of Helsinki, Helsinki, Finland.

[3] School of Clinical Dentistry, University of Sheffield, Faculty of Health, United Kingdom.

Corresponding Author: Ibrahim S Alsanie,

Assistant professor, Department of Oral Medicine and Diagnostic Sciences, College of Dentistry, King Saud University, Riyadh, Saudi Arabia.

Office AA22, Ground floor, College of Dentistry, King Saud University, Riyadh, Saudi Arabia, Tel: 00966114698424

Email: ialsanie@ksu.edu.sa, ORCID ID: 0000-0003-3037-9181

Authors emails

Ahmed Qannam, (qannam@ksu.edu.sa)

Ibrahim O. Bello, (ibello@ksu.edu.sa), ORCID ID: 0000-0003-2085-3185

Syed Ali Khurram, (s.a.khurram@sheffield.ac.uk), ORCID ID: 0000-0002-0378-9380

Conflict of interest statement

The authors declare that they have no conflicts of interest.

Acknowledgment

The authors would like to thank the College of Dentistry Research Centre and Deanship of Scientific Research at King Saud University, Saudi Arabia for funding this research project.

Author contributions

ISA worked on conceptualization of the study, data curation and manuscript preparation, editing and review. IOB has participated in study concept and design, data curation and manuscript editing and review. AQ has participated in data acquisition and manuscript editing and review. SAK has participated in data acquisition and manuscript editing and review.

Ethics Approval

The study was approved by Health and Care Research Wales ethics (reference: 20/WS/0017, 28/01/2020).

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon request.

Abstract:

Objectives: This narrative review aims to briefly discuss the concepts of artificial intelligence (AI), explore its role in oral pathology and oncology, and shed light on AI's potential in addressing diagnostic and prognostic challenges in the field. Additionally, future perspectives of AI were postulated.

Methods: A literature search using PubMed, EMBASE, Web of Science, and Scopus for studies published before February 2025 related to AI usage in the assessment of oral diseases such as oral potentially malignant disorders and transformation to oral cancer, oral squamous cell carcinoma, salivary gland tumours and odontogenic tumours.

Results: Numerous investigative efforts have been directed at AI applications for diagnostic, risk assessment and prognostication of oral diseases, although these are still relatively in their infancy, and more still needs to be done to refine these tools to make them overcome all the challenges that are being encountered.

Conclusions: While AI does not seek to replace oral pathologists but supplement their efforts, it has an immense potential to revolutionize oral pathology practice in the years to come. AI applications would greatly improve diagnostic accuracy, enhance epidemiological surveillance of diseases, drive the profession towards more personalized patient care and ultimately improve patient outcomes beyond the current level.

1. Introduction

1.1 Diagnostic and prognostic challenges in oral pathology

Oral lesions can manifest as lumps, ulcers or discolorations of soft tissues or as expansion or incidental discovery of a radiolucent lesion of hard tissues. Diagnosis of oral lesions is often challenging because of similar clinical presentation. The lesions often crucially require comprehensive clinical data and precise interpretation of supplementary tests such as radiographic imaging to make an accurate diagnosis.[1] Histologically, multiple lesions also share common characteristics e.g. presence of mucous cells and keratinization in cystic lesions, or ameloblastoma-like epithelium in other lesions and need careful consideration of clinical details, histological features and the expertise of experienced pathologists.[1]

1.2 Potential of AI in pathology

Artificial Intelligence (AI) holds great promise in histopathology, despite being presently envisaged as a support system for pathologists, and not as a substitute.[2] Utilizing AI techniques, pathologists can receive support in different aspects of the diagnostic procedure, such as image analysis, pattern identification, and decision-making. AI algorithms can analyse whole slide images (WSI) and assist in the segmentation of different tissue components. This can help automate time-consuming tasks, such as identifying tumour boundaries or distinguishing between normal and abnormal tissues as well as benign and malignant tumours.[3,4] Tissue microarrays have been used in some studies[5] and this will be briefly discussed in section 3.1.

Moreover, other imaging modalities, such as confocal microscopy, show significant potential, especially when integrated with AI for detecting oral epithelial dysplasia (OED) and oral squamous cell carcinoma (OSCC). A recent study demonstrated that confocal laser endomicroscopy combined with convolutional neural networks (CNNs) successfully analysed 9,168 in vivo images from 59 patients, achieving high diagnostic accuracy (AUC 0.90–0.96). These findings highlight the potential of AI-powered confocal microscopy to provide rapid and precise real-time diagnostic triage for high-risk oral mucosal diseases.[6]

AI models have the capability to be trained in identifying particular tumour types or cellular abnormalities by recognizing patterns and features within histopathology images. By examining patient data, medical records, and imaging results, AI algorithms have the capacity to propose potential diagnoses or provide risk assessments or prognosis thereby facilitating informed treatment options.[7-9]

AI is able to support quality assurance efforts by identifying potential errors or discrepancies in histopathology slides, thereby enhancing the precision and dependability of diagnoses. Additionally, AI can optimize workflow procedures, including slide scanning and data management, resulting in improved efficiency for pathologists.[10,11] For AI to be incorporated into clinical histopathology, there is need for thorough validation and standardisation requiring extensive, meticulously curated datasets and collaborative efforts among pathologists, computer scientists, and clinicians.[12]

Since most previous studies often concentrate on the use of AI in histopathology at the expense of key concepts in AI, the objective of this narrative review is to briefly discuss the latter in an explainable manner followed by exploration of the role of AI in oral pathology and oncology.

2. Artificial Intelligence (AI)

2.1 Tools and components of AI

The utilization of AI technologies has undergone significant growth over the past decade, and it is crucial to grasp the associated terminology in order to help and rationalize understanding. The tools and components related to AI are defined or explained in Table 1.

2.2 AI algorithms

Machine learning (ML) training can be broadly categorised into supervised and unsupervised learning. In unsupervised learning, the computer can uncover concealed patterns in input data without any prior knowledge or training. However, this requires a substantial amount of data for the algorithms to independently analyse and identify differences. Conversely, supervised learning involves training the machine to classify data based on previously known input and output, such as diagnostic annotations on histology slides. While this method requires the provision of labelled/annotated data, it strikes a balance between the size of the dataset and prediction accuracy, making it suitable for smaller datasets despite being more labour-intensive.[13,14]

ML models are often referred to as "shallow learning" because they typically have a limited number of layers or lack depth compared to more complex models. The training set consists of classes of objects, and the goal is to build a model that can predict the class of unknown examples. Deep learning (DL) utilizes multiple layers of convolutional neural networks (CNNs), which consist of an input layer, an output layer, and several intermediary hidden layers. This architecture enables the handling of intricate decision-making processes. DL and CNNs have an advantage of being able to automatically extract features directly from the data through a process called feature learning (Figure 1). In contrast, ML requires the programmer to manually engineer the feature set, explicitly specifying which features to consider. Additionally, DL offers the capability of reinforcement learning, where the algorithm learns to react to an environment by maximizing a reward function.[15]

2.3 Approaches and techniques

In supervised ML, there are several commonly used approaches such as detection, segmentation, classification, and path extraction/tessellation.

Detection focuses on identifying and locating objects within an image, allowing precise localisation. Classification involves dividing a dataset into groups based on specific features or characteristics, enabling the classification of new instances into the appropriate groups. Segmentation, on the other hand, aims to partition WSIs into smaller, meaningful regions of interest (ROI). This process helps extract relevant information and facilitates the grouping of data into distinct segments. Lastly, patch extraction/tessellation involves the creation of paths or interconnected patterns within the image, enabling the analysis of the spatial relationships between different elements.[13,15,16] These approaches include the segmentation of normal and cancerous tissues or benign and malignant tumours, as well as subtyping and grading of tumours. Given the large size and high dimensionality of WSIs, most AI model training methods for WSIs do not utilize the entire images as input. Instead, they divide the images into regions of interest and deconstruct them into smaller image patches for training. After making predictions or obtaining outcomes, these patches are then reassembled. Figure 2 illustrates the steps and workflow commonly employed in computational pathology for AI models, encompassing both ML and DL techniques.

Unsupervised learning can be used to identify hidden patterns in histopathological images and genomic data without predefined labels, aiding in tumour subtyping, anomaly detection, and biomarker discovery. However, this requires a huge amount of data. Common approaches include clustering techniques like k-means and hierarchical clustering for grouping similar tissue patterns, as well as dimensionality reduction methods such as principal component analysis (PCA) and autoencoders to enhance feature extraction for diagnostic and prognostic assessments.[13,16]

2.4 Training methods

2.4.1 Machine learning methods

Unsupervised learning makes use of diverse ML training techniques, including clustering, watershed, and Otsu[16-18] (Table 2). Supervised learning on the contrary, employs various training methods such as Decision Trees, Random Forests (RF), K-Nearest Neighbour (KNN), Support Vector Model (SVM), Bayesian classifiers, and Neural Networks[15,19-23] (Table 2).

2.4.2 Neural networks and deep learning methods

Various items can be classified using either artificial neural networks (ANN) and CNN, the two types of neural networks.[24] An ANN is a computational model that emulates the structure of the brain. It comprises interconnected artificial neurons that learn and predict by adapting connections based on training data. A CNN is a specialized ANN designed for analysing structured grid-like data. CNNs excel in tasks like image classification and object detection. A CNN has exhibited superior performance compared to ANN in numerous computer vision tasks, such as image classification, segmentation, and object detection.[15,25] The training methods type and subtypes are summarised in Table 2.

2.4.3 Algorithm Performance Assessment

Multiple performance metrics are available to assess the efficacy of algorithms, with some metrics specifically designed for particular applications. Threshold-based metrics, relying on a qualitative understanding of error, are utilized when the objective is to minimize the number of errors. In contrast, other metrics evaluate the model's performance when classifiers are employed to select the optimal instances from a dataset.[26] In object detection tasks, such as identifying mitotic figures in tissue, metrics like precision, recall, and the F1-score to evaluate detection performance are often utilized.[27] For segmentation tasks, where the goal is to classify each pixel into a category (e.g., tumour segmentation), Dice Similarity Coefficient (DSC) and Intersection over Union (IoU) are widely used to evaluate the overlap between predicted and ground-truth segmentation masks.[28] Furthermore, in classification tasks, such as distinguishing between various tissue types , metrics like accuracy, precision, recall, F1-score, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC) are employed for evaluating performance.[29] A brief explanation of performance assessment tools is also provided in Table 1.

2.5 Use of AI in Diagnostic Oral Pathology

Diagnostic oral pathology can benefit from AI with algorithms being used to effectively organise and classify data such as patient demographics, clinical notes, and laboratory test results. AI can help differentiate lesional areas from normal tissue, in addition to being able to diagnose lesions as well as classifying them to subtypes or histological grades.

2.5.1 Oral Potentially Malignant Disorders (OPMDs)

The risk of OPMDs progressing to OSCC varies based on several factors, making its prediction difficult. Physicians evaluate each case individually to determine the best treatment approach. A biopsy followed by histological analysis remains the most reliable diagnostic method, particularly for detecting OED, which is characterized by abnormal cell morphology and structural changes in the oral epithelium.[30]

In 2012, Krishnan et al. developed a texture-based method that utilized the density and thickness of each epithelial layer to accurately segment oral submucous fibrosis (OSF) tissue from normal tissue.[31] Recently, some investigators used neural network architectures for segmentation of 1636 images of cancerous and precancerous lesions. They found that the models were able to classify cropped or entire images based on the presence or absence of lesions with satisfactory accuracy (F1-score 0.77). This reinforces the promising prospect of AI for lesion identification, in addition to identifying their associated specific pathology.[32]

Accurate grading of OED is crucial for effective patient management of the lesions. However, this task is particularly difficult due to the significant intra- and interobserver disagreement among pathologists.[30] This emphasizes the necessity for an automated system that can provide more precise and dependable grading. The findings of some investigators highlight the importance of utility of the irregular stratification of the epithelium, an architectural abnormality, as a valuable training factor for ML models in accurately predicting dysplasia grade.[33] As is often the case, the findings are limited by the small sample size and the exclusion of many important architectural and cytological features which can be added in future models.

2.5.2 Oral cancer

OSCC constitutes more than 90% of head and neck cancers, and has been on the rise in various regions worldwide.[34,35] A significant number of cases are diagnosed late leading to lower survival rates which has stood at approximately 50-55% for many years.[34] Its prevalence is influenced by factors such as lifestyle choices, cultural practices and the availability of healthcare services.[35]

The potential of AI in oral cancer lies in its ability to aid pathologists in accurate diagnosis, prognostication, and treatment through image analysis, pattern recognition, and decision support systems. An early AI study which showed a good potential for automated diagnosis and grading of OSCC using keratinization index was previously done.[36] The study was then followed up by the same group three years later.[3] This latter study incorporated a model that was able to accurately detect and segment various layers of the oral mucosa and subsequently employed a texture-based classification method to identify keratin pearls within the segmented keratin regions. The model demonstrated an outstanding accuracy of 96.88%.[3] Both studies had the limitation of small sample size, absence of cross-centre validation and lack of comparison with a DL method. In 2019, the same investigators created a computer-aided tool designed to identify and outline nuclei in oral histopathology images, thus assisting in the screening of OSCC. The authors employed a combination of texture analysis and ML methods to develop an algorithm capable of automatically segmenting nuclei from histological images with an excellent 94.22% dice coefficient.[37] This study had similar limitations to those that preceded it which could potentially restrict the applicability of its findings to different populations.

In another study, Rahman et al., employed microscopic images to categorize histological slides of normal and OSCC tissues. They utilized a grey-level co-occurrence matrix (GLCM) to analyse the texture features of the images, while feature extraction was carried out using histogram techniques. The classification was performed using a Linear SVM (Table 2), which achieved a remarkable accuracy of 100% and yielded satisfactory outcomes.[38] The training cohorts include cropped images of haematoxylin and eosin (H&E) slides, a limitation which may introduce bias and affect the representativeness of the training data.

Another study deployed four deep networks for comparison of OSCC segmentation using annotated data from The Cancer Genome Atlas (TCGA) dataset. Assessment of the performance of the networks showed that U-Net, when enhanced with ResNet50 as an encoder outperformed the original U-Net.[39] Amin and co-investigators presented an automated classification model aimed at distinguishing between normal and malignant areas using a dataset comprising 290 normal images and 934 OSCC images. The study employed VGG16, InceptionV3, and ResNet50 as separate models, and subsequently combined them into a concatenated model. Notably, the concatenated model demonstrated superior performance compared to the individual models, yielding the best results in the classification task with AUC = 0.997.[40] This study is limited by suboptimal image quality as the cohorts were sourced from a publicly accessible dataset composed of camera-captured images rather than WSI. In addition, there was an imbalanced training cohort and a small test dataset, which may not adequately represent the full distribution observed in the training data.

Recently, Panigrahi et al., conducted a study where they employed capsule networks (CapsNets) to classify OSCC and differentiate between cancerous and non-cancerous histology images. The study revealed that CapsNets, which are capable of processing spatial data, exhibited superior performance (accuracy 97.35% versus 96.77%) compared to CNNs when applied to the same datasets, showcasing their enhanced classification abilities.[41] The proposed model demonstrates efficiency in detecting multistage OSCC and has the potential to assist in routine clinical screening. However, further validation is required using larger multicentre datasets to confirm these findings.

2.5.3 Salivary Gland Tumours

Salivary gland tumours (SGT) are a heterogeneous group of neoplasms with morphological diversity and overlapping features. SGT can be diagnostically challenging due to a large number of entities and markedly similar features but different clinical behaviour. One study looking at different SGT subtypes and grades based on the analysis of H&E stained, digitised WSIs showed that the ML classifiers results achieved excellent performance with F1 score of 0.90, 0.92 and 0.87, for benign vs malignant, malignant subtyping and grading, respectively.[42] However, to validate these findings, it is essential to conduct further testing on larger cohorts across multiple centres.

2.5.4 Prospect of AI in Diagnosis of Odontogenic Cysts and Tumours

The use of AI in the diagnosis of odontogenic cysts and tumours is quite promising and has the potential to greatly assist in accurately identifying and classifying these lesions which are often encountered by oral pathologists in their routine clinical practice. AI can be used for analysing radiographic images, such as panoramic radiographs and cone beam computed tomography (CBCT) to detect and identify specific patterns, shapes, and densities associated with different pathologies, enabling more accurate diagnosis. Additionally, AI algorithms can be trained on large datasets of annotated histopathology images to recognise and classify different types, and assist in detecting specific cellular and architectural patterns associated with different pathologies. Moreover, AI can aid in quantitative analysis of relevant features and measurements. To date, only one study has utilized AI tools to aid in distinguishing between ameloblastoma and ameloblastic carcinoma which often pose a diagnostic challenge in routine histopathological practice due to their similarities and the limitations associated with incisional biopsies. The study utilized DL models using 30 digitized images and compared three models (ResNet50, DenseNet, and VGG16) to assess the probability of an image being classified ameloblastoma or ameloblastic carcinoma. The best performance was by ResNet50 with F1- score of 0.77.[43] The limitations of the study include an imbalanced training cohort and a small test dataset.

3. Prognostication and Predictive Potential of AI

3.1 AI in Prognostication of Oral Dysplasia and Cancer

Conventional statistical methods, like the survival analysis, Kaplan-Meier curves, and Cox regression, have been utilized to predict the survival of cancer patients. However, these methods, based on subjective categories or features, are often insufficient in handling the complexities associated with such conditions. The implementation of an AI-based predictive system becomes necessary to achieve more promising and accurate results.[44]

A prognostic model has been proposed for predicting malignant transformation in tissue sections of OED. The study included a total of 137 cases, out of which 50 cases exhibited malignant transformation, with an average time to transformation of 6.51 years (±5.35 SD). The model achieved an AUROC of 0.78 for predicting the occurrence of malignant transformation in OED using a stratified five-fold cross-validation approach. The analysis of hotspot areas revealed several significant prognostic factors for malignant transformation, including the count of peri-epithelial lymphocytes ($p < 0.05$), the count of nuclei in the epithelial layer ($p < 0.05$), and the count of nuclei in the basal layer ($p < 0.05$). These features were associated with a high risk of malignant transformation in the analysis.[45]

In another investigation, a CNN-based oral mucosa risk stratification model (OMRS) was utilized to categorize a group of non-dysplastic oral mucosa samples (n=31) and a collection of OSCC (n=31) H&E-stained slides. The results indicated that low-risk patients had a 5-year OSCC development probability of 21.3%, while high-risk patients had a probability of 52.5%. This demonstrates the efficacy of the OMRS model in identifying oral leukoplakia (OL) patients at a heightened risk of developing OSCC, thus potentially contributing to improved early diagnosis and prevention strategies for OSCC.[46] The study's limitations include a small cohort and a prognostic analysis limited to patients with OL without concurrent OSCC. It would be advisable to broaden the study to include other subsets of OPMD and explore patients both with and without concurrent OSCC to obtain a more comprehensive understanding of the utility of the model.

Lu et al. developed a classifier utilizing image analysis techniques to assess nuclear shape, size, and texture diversity in cell clusters from 2 mm OSCC microarray tumour sections images. The findings indicate that the quantitative histomorphometric features of local nuclear architecture have the potential to serve as independent predictors of patient survival.[5] However, this study has several issues as its limitations. The image analysis was restricted to tissue microarrays, which only represent a small fraction of the complete tumour. Furthermore, the sample size in the study was relatively small, and certain well-established histological prognostic features, such as depth of invasion and nodal extracapsular extension, were not considered.

In 2019, Shaban et al. introduced an automated method that utilizes DL-based tissue segmentation to quantify the abundance of tumour-infiltrating lymphocytes (TILs) in histological images of OSCC. The digital TILs score was calculated to investigate its potential as a prognostic marker. The proposed approach achieved an accuracy of 96.31%. Notably, the automatically generated TILs score exhibited superior predictive value compared to the manually determined TILs score.[9] In a subsequent investigation, researchers employed a DL-driven automated approach to examine the predictive importance of tumour-associated stroma infiltrating lymphocytes (TASILs) in OSCC. The results indicated that the TASIL-score demonstrated superior discrimination between low-risk and high-risk patients in terms of both disease-specific and disease-free survival when compared to the conventional manual scoring of TILs by pathologists.[47]

Overall, the available research evidence on successful deployment of AI in oral cancer is insufficient. This is primarily attributable to the utilization of small, single-centre datasets and a notable risk of bias that may have led to an overestimation of the accuracy rate of the models.[13] To validate the effectiveness of the classifiers on WSI, and control for all established clinical and pathological features, a larger, statistically-powered, patient cohort, should be analysed.[5]

4. Large language models (LLMs) and Pathology

LLMs are powerful AI tools that learn to understand and generate human-like text after being trained on massive amounts of text data. Examples of LMMs include Open AI GPT series used in Chat GPT and Microsoft Copilot, Gemini by Google, LLaMA by Meta, Granite series by IBM, and most recently DeepSeek. The effectiveness of AI in handling pathology queries depends on the complexity of the question and the scope of the AI model's training data. While a chatbot might excel at answering basic questions about medical sciences, complex questions that require deep medical knowledge, nuanced pathology interpretation, critical thinking, complex reasoning and subtle judgement might still be best handled by human experts. [48] Currently, AI is driving advancements in the integration of visual and textual analysis of pathology data. It can seamlessly connect specific areas of interest within a pathology image (such as a tumour) to their associated descriptions in the accompanying pathology report.[49]

For example, chatbot generative pre-trained transformer (ChatGPT) offers potential benefits in research, education, and patient communication. However, significant challenges regarding integration into clinical practice abound. These include limitations in handling complex pathological details, the lack of real-time interaction and visual data analysis, concerns about how the model makes decisions, data privacy, the need for constant updates, and seamless integration into existing workflows. Few studies have explored the reliability of ChatGPT as a tool for pathologists in their routine work.[50] While tools such as ChatGPT hold promise for aiding pathologists with scientific data in routine diagnosis, limitations like its training data, data availability, and the 'hallucination' effect (generation of false or misleading information or conclusions) must be addressed.[51]

Recently, retrieval-augmented Generation (RAG) has been integrated into histopathology to enhance diagnostic accuracy and efficiency. RAG combines content-based image retrieval (CBIR) with generative models, enabling pathologists to retrieve visually similar histopathological images from extensive databases. This approach assists in identifying disease patterns by comparing query images with archived cases, thereby supporting diagnosis, treatment planning, and education.[52]

Chatbots, such as ChatGPT can serve as an adjunctive tool in diagnosing oral diseases by assisting in symptom analysis, patient education, and clinical decision support. While AI technologies enhance early detection of oral pathologies like OSCC and leukoplakia, their effectiveness depends on accurate training data and clinician oversight. However, ethical concerns such as misdiagnosis risk, data privacy, and regulatory compliance necessitate that AI chatbots be used as supplementary tools rather than standalone diagnostic systems.[48]

CPath has progressed with task-specific predictive models and self-supervised vision encoders, yet there has been limited research on multimodal AI assistants designed specifically for pathology. For instance, PathChat, a vision-language AI assistant, was created by integrating a foundational vision encoder with a pretrained large language model and fine-tuning it on over 456,000 visual language instructions. PathChat surpassed other multimodal AI assistants, including GPT4V, in diagnostic accuracy and responses preferred by pathologists, indicating its potential use in pathology education, research, and clinical decision-making.[53] Future versions of generative AI tools with extensive pretraining could further expand their utility in CPath applications.[54]

5. AI opportunities and challenges in diagnostic sciences

5.1 Opportunities

AI has the potential to reduce workload, enhance clinical practice and improve patient care by minimizing medical errors and bias.[55] It can also increase operational efficiency by reducing turnaround time, automating repetitive tasks, and mitigating errors that may arise from manual labour. In the field of pathology, AI algorithms that utilize handcrafted features, such as nuclear size, have demonstrated the ability to achieve high levels of accuracy in identifying various conditions with significantly less effort. Additionally, AI software solutions have the capacity to handle laborious tasks like mitosis counting and streamline intricate processes like triaging urgent cases.[56] Studies have shown that ML/DL technologies have achieved impressive levels of accuracy in tasks such as identifying, segmenting, classifying, and grading different types of cancer, while also providing valuable prognostic information. Moreover, certain complex cases require additional molecular testing to confirm diagnoses or determine specific genetic mutations. This process typically involves extra tissue sectioning, additional laboratory procedures, and increased time and costs. CPath also holds promise in predicting molecular changes on H&E-stained WSI, enabling rapid diagnoses and ultimately enhancing patient care.[56]

5.2 Challenges

Despite significant advancements made in the past decade, the widespread integration of AI into routine pathology practice is still a distant goal. The challenges of achieving this goal are related to issues inherent in the pathologists themselves, the resources needed, the quality and the variability of the data available, ethical concerns and the consequence of possible misdiagnosis by AI.

The pathologists are often limited in their interaction with AI researchers or face time constraints that hinder their ability to learn new technologies despite their interest in adopting it.[57] Generally, there may be lack of familiarity with AI technology among pathologists due to insufficient training.[58] DL models involve intricate decision-making and would require enhancing their interpretability and addressing any associated ambiguity to the pathologist. Additionally, information needs to be offered to the pathologists about the characteristics of the algorithm to make the model more understandable and transparent.[59]

The progress of AI technology in relation to pathology requires high-performance computing resources, good bandwidth, efficient and modern laboratory management information, flexibility and storage in server or cloud configurations and excellent cybersecurity. These are often lacking in many histopathology laboratories due to the associated financial implications.[56,59]

The quality of the data underlying an AI algorithm is extremely important. In general, these algorithms require large amounts of data to be annotated by pathologists which can be monotonous and present additional challenges in cases of low-resolution images or unclear images and ambiguous features.[56,58] DL algorithms, in particular, are data-dependent and require large amounts of data to automatically recognise important features unlike traditional image analysis methods that rely on manual selection of features.[59] Most DL methods in pathology focus on analysing small image patches rather than the entire WSI thereby limiting its prediction capacity due to restricted view and lack of contextual information about the surrounding structure.[56] Data variability is also an important issue as the best AI CPath solution will be to create an optimal algorithm from a wide range of sources to effectively manage variations between different datasets.[59] For instance, variations in stain colour can occur due to several factors, including variances in slide thickness, tissue thickness, fixation methods, tissue processing schedules, differences in staining techniques, and variations between laboratories. Therefore, it is necessary to implement consistent pre-imaging procedures such as colour normalisation, manual or automated image quality control, and the utilization of larger training sets.[59]

Ethical issues also need to be addressed to ensure safety, accuracy and effectiveness of AI models. The release of public data may give rise to ethical concerns regarding privacy violations, which can lead to the implementation of restrictive governance policies.[59] Specific regulatory approvals and permissions need to be in place before algorithms can be utilized in clinical practice, aiming to promote the use of secure and reliable models. However, in the field of CPath, there is a limited number of accessible datasets due to confidentiality, copyright, and financial considerations.[58] Also, there would be a need to avoid creating disparities in health outcomes by eliminating using AI solutions that disproportionately exclude or include individuals based on factors such as socioeconomic status or location.[60]

The question of when AI makes a misdiagnosis should also be addressed. Such errors that can impact patients should be mitigated by overall supervision of quality and effectiveness from the human clinician.[60]

6. Conclusion

The influence of AI has grown significantly in the last decade, and encompasses all aspects of healthcare. For oral pathology, the growth spurt especially covers diagnosis and prognosis of oral lesions. While it will not replace but complement the pathologist, it is envisaged that AI utilization will facilitate excellent diagnosis, better and more efficient workload and generally enhance all aspects of oral pathology practice. However, the limitations associated with current evidence in support of the utility of AI for these purposes need to be addressed.

References

1. Hunter KD, Niklander S. Pitfalls in odontogenic lesions and tumours: a practical guide. *Diagnostic Histopathology.* 2020;26(4):173-180.
2. Das N, Hussain E, Mahanta LB. Automated classification of cells into multiple classes in epithelial tissue of oral squamous cell carcinoma using transfer learning and convolutional neural network. *Neural Networks.* 2020;128:47-60.
3. Das DK, Bose S, Maiti AK, Mitra B, Mukherjee G, Dutta PK. Automatic identification of clinically relevant regions from oral tissue histological images for oral squamous cell carcinoma diagnosis. *Tissue and Cell.* 2018;53:111-119.
4. Ehteshami Bejnordi B, Mullooly M, Pfeiffer RM, et al. Using deep convolutional neural networks to identify and classify tumor-associated stroma in diagnostic breast biopsies. *Modern Pathology.* 2018;31(10):1502-1512.
5. Lu C, Lewis Jr JS, Dupont WD, Plummer Jr WD, Janowczyk A, Madabhushi A. An oral cavity squamous cell carcinoma quantitative histomorphometric-based image classifier of nuclear morphology can risk stratify patients for disease-specific survival. *Modern Pathology.* 2017;30(12):1655-1665.
6. Ramani RS, Tan I, Bussau L, et al. Convolutional neural networks for accurate real-time diagnosis of oral epithelial dysplasia and oral squamous cell carcinoma using high-resolution in vivo confocal microscopy. *Scientific Reports.* 2025;15(1):2555.
7. Bejnordi BE, Veta M, Van Diest PJ, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *Jama.* 2017;318(22):2199-2210.
8. Lipkova J, Chen RJ, Chen B, et al. Artificial intelligence for multimodal data integration in oncology. *Cancer cell.* 2022;40(10):1095-1110.

9. Shaban M, Khurram SA, Fraz MM, et al. A novel digital score for abundance of tumour infiltrating lymphocytes predicts disease free survival in oral squamous cell carcinoma. *Scientific reports.* 2019;9(1):1-13.

10. Williams BJ, Bottoms D, Treanor D. Future-proofing pathology: the case for clinical adoption of digital pathology. *Journal of clinical pathology.* 2017;70(12):1010-1018.

11. Hanna MG, Parwani A, Sirintrapun SJ. Whole slide imaging: technology and applications. *Advances in Anatomic Pathology.* 2020;27(4):251-259.

12. Hanna MG, Ardon O, Reuter VE, et al. Integrating digital pathology into clinical practice. *Modern Pathology.* 2022;35(2):152-164.

13. Mahmood H, Shaban M, Indave B, Santos-Silva A, Rajpoot N, Khurram S. Use of artificial intelligence in diagnosis of head and neck precancerous and cancerous lesions: A systematic review. *Oral Oncology.* 2020;110:104885.

14. Zarella MD, Bowman D, Aeffner F, et al. A practical guide to whole slide imaging: a white paper from the digital pathology association. *Archives of pathology & laboratory medicine.* 2019;143(2):222-234.

15. Cohen S. *Artificial intelligence and deep learning in pathology.* Elsevier Health Sciences; 2020.

16. de Souza LL, Fonseca FP, Araujo ALD, et al. Machine learning for detection and classification of oral potentially malignant disorders: A conceptual review. *Journal of Oral Pathology & Medicine.* 2023;52(3):197-205.

17. Levner I, Zhang H. Classification-driven watershed segmentation. *IEEE Transactions on Image Processing.* 2007;16(5):1437-1445.

18. Otsu N. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics.* 1979;9(1):62-66.

19. Begg RK, Palaniswami M, Owen B. Support vector machines for automated gait classification. *IEEE transactions on Biomedical Engineering.* 2005;52(5):828-838.

20. Breiman L. Random forests. *Machine learning.* 2001;45(1):5-32.

21. Quinlan JR. Induction of decision trees. *Machine learning.* 1986;1(1):81-106.

22. Rish I. An empirical study of the naive Bayes classifier. Paper presented at: IJCAI 2001 workshop on empirical methods in artificial intelligence2001.

23. Yu X, Wang B, Batbayar B, Wang L, Man Z. An improved training algorithm for feedforward neural network learning based on terminal attractors. *Journal of Global Optimization.* 2011;51(2):271-284.

24. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology—new tools for diagnosis and precision oncology. *Nature reviews Clinical oncology.* 2019;16(11):703-715.

25. Hong R, Fenyö D. Deep Learning and Its Applications in Computational Pathology. *BioMedInformatics.* 2022;2(1):159-168.

26. Ferri C, Hernández-Orallo J, Modroiu R. An experimental comparison of performance measures for classification. *Pattern recognition letters.* 2009;30(1):27-38.

27. Wang H, Cruz-Roa A, Basavanhally A, et al. Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features. *Journal of Medical Imaging.* 2014;1(3):034003-034003.

28. Milletari F, Navab N, Ahmadi S-A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. Paper presented at: 2016 fourth international conference on 3D vision (3DV)2016.

29. Hameed Z, Garcia-Zapirain B, Aguirre JJ, Isaza-Ruget MA. Multiclass classification of breast cancer histopathology images using multilevel features of deep convolutional neural network. *Scientific Reports.* 2022;12(1):15600.

30. Speight PM, Khurram SA, Kujan O. Oral potentially malignant disorders: risk of progression to malignancy. *Oral surgery, oral medicine, oral pathology and oral radiology.* 2018;125(6):612-627.

31. Muthu Rama Krishnan M, Shah P, Chakraborty C, Ray AK. Statistical analysis of textural features for improved classification of oral histopathological images. *Journal of medical systems.* 2012;36(2):865-881.

32. Tobias MA, Nogueira BP, Santana MC, Pires RG, Papa JP, Santos PS. Artificial intelligence for oral cancer diagnosis: What are the possibilities? *Oral Oncology.* 2022;134:106117.

33. Bashir RS, Mahmood H, Shaban M, et al. Automated grade classification of oral epithelial dysplasia using morphometric analysis of histology images. Paper presented at: Medical Imaging 2020: Digital Pathology2020.

34. Chamoli A, Gosavi AS, Shirwadkar UP, et al. Overview of oral cavity squamous cell carcinoma: Risk factors, mechanisms, and diagnostics. *Oral oncology.* 2021;121:105451.

35. Gupta B, Johnson NW, Kumar N. Global epidemiology of head and neck cancers: a continuing challenge. *Oncology.* 2016;91(1):13-23.

36. Das DK, Chakraborty C, Sawaimoon S, Maiti AK, Chatterjee S. Automated identification of keratinization and keratin pearl area from in situ oral histological images. *Tissue and Cell.* 2015;47(4):349-358.

37. Das DK, Koley S, Bose S, et al. Computer aided tool for automatic detection and delineation of nucleus from oral histopathology images for OSCC screening. *Applied Soft Computing.* 2019;83:105642.

38. Rahman T, Mahanta L, Chakraborty C, Das A, Sarma J. Textural pattern classification for oral squamous cell carcinoma. *Journal of microscopy.* 2018;269(1):85-93.

39. Martino F, Bloisi DD, Pennisi A, et al. Deep learning-based pixel-wise lesion segmentation on oral squamous cell carcinoma images. *Applied Sciences.* 2020;10(22):8285.

40. Amin I, Zamir H, Khan FF. Histopathological image analysis for oral squamous cell carcinoma classification using concatenated deep learning models. *medRxiv.* 2021:2021.2005. 2006.21256741.

41. Panigrahi S, Das J, Swarnkar T. Capsule network based analysis of histopathological images of oral squamous cell carcinoma. *Journal of King Saud University-Computer and Information Sciences.* 2022;34(7):4546-4553.

42. Alsanie I, Shephard A, Azarmehr N, Rajpoot N, Khurram SA. Using Artificial Intelligence for Analysis of Histological and Morphological Diversity in Salivary Gland Tumors. *researchsquarecom* 2022.

43. Giraldo-Roldan D, Ribeiro ECC, Araújo ALD, et al. Deep learning applied to the histopathological diagnosis of ameloblastomas and ameloblastic carcinomas. *Journal of Oral Pathology & Medicine.* 2023;52(10):988-995.

44. Khanagar SB, Naik S, Al Kheraif AA, et al. Application and performance of artificial intelligence technology in oral cancer diagnosis and prediction of prognosis: a systematic review. *Diagnostics.* 2021;11(6):1004.

45. Bashir RMS, Shephard AJ, Mahmood H, et al. A digital score of peri-epithelial lymphocytic activity predicts malignant transformation in oral epithelial dysplasia. *The Journal of Pathology.* 2023.

46. Zhang X, Gleber-Netto FO, Wang S, et al. Deep learning-based pathology image analysis predicts cancer progression risk in patients with oral leukoplakia. *Cancer medicine.* 2023.

47. Shaban M, Raza SEA, Hassan M, et al. A digital score of tumour-associated stroma infiltrating lymphocytes predicts survival in head and neck squamous cell carcinoma. *The Journal of Pathology.* 2022;256(2):174-185.

48. Panwar P, Gupta S. A review: Exploring the role of ChatGPT in the diagnosis and treatment of oral pathologies. *Oral Oncology Reports.* 2024:100225.

49. Muftić F, Kadunić M, Mušinbegović A, Abd Almisreb A. Exploring Medical breakthroughs: a systematic review of ChatGPT Applications in Healthcare. *Southeast Europe Journal of Soft Computing.* 2023;12(1):13-41.

50. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Paper presented at: Healthcare2023.

51. Cazzato G, Capuzzolo M, Parente P, et al. Chat GPT in Diagnostic Human Pathology: Will It Be Useful to Pathologists? A Preliminary Review with 'Query Session' and Future Perspectives. *AI.* 2023;4(4):1010-1022.

52. Tizhoosh HR, Pantanowitz L. On image search in histopathology. *Journal of Pathology Informatics.* 2024:100375.

53. Lu MY, Chen B, Williamson DF, et al. A Multimodal Generative AI Copilot for Human Pathology. *Nature.* 2024:1-3.

54. Ullah E, Parwani A, Baig MM, Singh R. Challenges and barriers of using large language models (LLM) such as ChatGPT for diagnostic medicine with a focus on digital pathology–a recent scoping review. *Diagnostic Pathology.* 2024;19(1):1-9.

55. Aung YY, Wong DC, Ting DS. The promise of artificial intelligence: a review of the opportunities and challenges of artificial intelligence in healthcare. *British medical bulletin.* 2021;139(1):4-15.

56. Tizhoosh HR, Pantanowitz L. Artificial intelligence and digital pathology: challenges and opportunities. *Journal of pathology informatics.* 2018;9(1):38.

57. Regitnig P, Mueller H, Holzinger A. Expectations of artificial intelligence for pathology. In: *Artificial Intelligence and Machine Learning for Digital Pathology.* Springer; 2020:1-15.

58. Cheng JY, Abel JT, Balis UG, McClintock DS, Pantanowitz L. Challenges in the development, deployment, and regulation of artificial intelligence in anatomic pathology. *The American Journal of Pathology.* 2021;191(10):1684-1692.

59. Abels E, Pantanowitz L, Aeffner F, et al. Computational pathology definitions, best practices, and recommendations for regulatory guidance: a white paper from the Digital Pathology Association. *The Journal of pathology.* 2019;249(3):286-294.

60. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nature medicine.* 2019;25(1):44-56.

61. Pakdemirli E. A preliminary glossary of artificial intelligence in radiology. *Acta Radiologica Open.* 2019;8(7):2058460119863379.

Figure legends

Figure 1: Deep learning models' framework, including input, output and hidden neural network layers. (Modified from Alsanie, Ibrahim (2023) Using Artificial Intelligence for Analysis of Histological and Morphological Diversity in Salivary Gland Tumours. White Rose eTheses Online).

Figure 2: Overview of the workflow of ML/DL models in computational pathology. (Modified from Alsanie, Ibrahim (2023) Using Artificial Intelligence for Analysis of Histological and Morphological Diversity in Salivary Gland Tumours. White Rose eTheses Online).

Table legends

Table 1: Components of AI relevant to oral pathology.

Table 2: Training methods in machine learning and deep learning.

Table 1: Components of AI relevant to oral pathology

Term	Definition
Artificial intelligence (AI)	Intelligence achieved by machines in performing tasks traditionally carried out by humans, such as visual perception, translation, image interpretation, speech recognition and decision-making. [61]
Machine learning (ML)	The utilization of algorithms that identify patterns in data without explicit instructions.[13,61]
Deep learning (DL)	A branch of machine learning (ML) that focuses on algorithms acquiring knowledge from input data through examples.[13]
Neural network	A structured collection of algorithms that emulates the neural network system of the human brain. It uses layers of interconnected nodes to learn patterns and make predictions.[13,23]
Black box	Deep learning systems process inputs and produce outputs, but the intermediate computations are hard for humans to interpret.[61]

Algorithm	A computational procedure that provides a systematic approach to solving a specific class of problems. It enables calculations, data processing, and automated reasoning.[61]
Precision (p)	The ratio of correctly identified positive results to the total positive results predicted by the classifier.[13]
Recall (r)	The proportion of correctly identified positive results out of all relevant samples.[13]
F1 score	The statistical analysis of binary classification involves evaluating the accuracy of a test by calculating an overall score that takes into account the weighted average of precision (p) and recall (r).[13]
The Dice Similarity Coefficient (DSC)	A statistical measure used to assess the similarity between two sets, often applied in image segmentation tasks.[15]
Intersection over Union (IoU)	A metric used to measure the overlap between two sets, commonly used in the segmentation of images.[15]
Area Under the Receiver Operating Characteristic	A metric used for classification models, especially in binary classification. It measures the model's ability to distinguish between classes by plotting the True Positive Rate (TPR) against the False Positive Rate (FPR) at various threshold settings. AUC ranges from

Curve (AUC-ROC)	0 to 1, where 1.0 signifies a perfectly accurate classifier, while 0.0 represents a totally misclassified model.[15]
Data augmentation	A technique employed during model training where images undergo slight modifications. The objective is to emphasize the learning of essential segmentation features by the model, rather than relying on image-specific attributes.[58]
Transfer learning	A machine learning technique where an algorithm acquires knowledge from performing one task and leverages that knowledge when learning a different yet related task.[56,58]

Table 2: Training methods in machine learning and deep learning

Training methods	Training methods types	
Machine learning	**Unsupervised learning**	**Clustering:** A technique that groups similar data points based on features like intensity, colour, or texture.[16]
		Watershed: A region-based algorithm treating an image as a topographic surface, where ridges act as boundaries.[17]
		Otsu: A thresholding technique that finds the optimal threshold by minimizing intra-class variance in a histogram.[18]
	Supervised learning	**Decision Trees:** A flowchart-like model that splits data based on feature conditions to make decisions.[15,21]
		Random Forests (RF): An ensemble of multiple decision trees that improves accuracy and reduces overfitting.[20]
		K-Nearest Neighbour (KNN): A non-parametric method that classifies data based on the majority vote of its closest neighbours.[15]
		Support Vector Model (SVM): A model that finds the optimal hyperplane to separate data into different classes.[19]

		Bayesian classifiers: Probabilistic models that use Bayes' theorem to predict class membership based on prior probabilities.[22]
		Neural Networks: Defined earlier in Table 1.
Deep learning	**Artificial neural networks (ANNs):** A multi-layered networks of neurons that learns patterns from data.[25]	
	Convolutional neural networks (CNNs): A model that use convolutional layers for image and spatial data processing.[25]	
	Recurrent Neural Networks (RNNs): A neural network for sequential data that retains past information for context.[15]	
	Generative Adversarial Networks (GANs): A model with two competing networks that generate realistic synthetic data.[25]	