



## REVIEW ARTICLE OPEN

# ChatGPT for shaping the future of dentistry: the potential of multi-modal large language model

Hanyao Huang<sup>1</sup>✉, Ou Zheng<sup>1</sup>✉, Dongdong Wang<sup>2</sup>, Jiayi Yin<sup>1</sup>, Zijin Wang<sup>2</sup>, Shengxuan Ding<sup>3</sup>, Heng Yin<sup>1</sup>, Chuan Xu<sup>4,5</sup>, Renjie Yang<sup>6</sup>, Qian Zheng<sup>1</sup> and Bing Shi<sup>1</sup>

The ChatGPT, a lite and conversational variant of Generative Pretrained Transformer 4 (GPT-4) developed by OpenAI, is one of the milestone Large Language Models (LLMs) with billions of parameters. LLMs have stirred up much interest among researchers and practitioners in their impressive skills in natural language processing tasks, which profoundly impact various fields. This paper mainly discusses the future applications of LLMs in dentistry. We introduce two primary LLM deployment methods in dentistry, including automated dental diagnosis and cross-modal dental diagnosis, and examine their potential applications. Especially, equipped with a cross-modal encoder, a single LLM can manage multi-source data and conduct advanced natural language reasoning to perform complex clinical operations. We also present cases to demonstrate the potential of a fully automatic Multi-Modal LLM AI system for dentistry clinical application. While LLMs offer significant potential benefits, the challenges, such as data privacy, data quality, and model bias, need further study. Overall, LLMs have the potential to revolutionize dental diagnosis and treatment, which indicates a promising avenue for clinical application and research in dentistry.

*International Journal of Oral Science* (2023)15:29

; <https://doi.org/10.1038/s41368-023-00239-y>

## INTRODUCTION

Artificial intelligence (AI) has promoted recent progress in digital health for many years.<sup>1,2</sup> AI-equipped applications in dentistry have been found useful in analyzing medical imaging, including diagnosing dental caries,<sup>3,4</sup> periodontitis,<sup>5</sup> and implantitis,<sup>6</sup> and assisting oral and maxillofacial surgery with surgical planning.<sup>7</sup> Besides the imaging data, audio data analyses can also benefit from deep-learning applications, as speech is one of the most important functions of the oral structure.<sup>8,9</sup> Furthermore, dental education is another emerging application.<sup>10</sup> GPT-4, released by OpenAI, embarks on a new period of AI-powered large language models (LLMs). ChatGPT, built upon GPT-4, stirred up lots of interest among millions of scientists and engineers on account of its impressive human conversational response as a chatbot.<sup>11</sup> However, its potential impact on revolutionizing a series of technologies is more significant. Unlike earlier applications, ChatGPT is cultivated conversationally upon a tremendous knowledge base, enabling informative communications for the improvement of decision knowledge. Before ChatGPT, most AI technologies focus on the system of one input and one output, which relies on the amount of training data. With the influx of new data, re-training is required to update the existing model for more accurate decision-making. ChatGPT breakthroughs this mode and incorporates conversation to dynamically capture multiple sources of existing knowledge for question answering.<sup>2,12,13</sup> This human-friendly feature facilitates the diagnosis process and causes a

significant change in the status quo, and its advancement will also shape digital health in dentistry.<sup>1,14–16</sup> The purpose of this paper is to provide an overview of the potential application of ChatGPT in dentistry.

## JOURNEY OF LLMs

Before LLMs garner significant attention, language modeling has undergone a series of revolutions in the past decade. The early natural language model is carried out with n-gram modeling,<sup>17</sup> which is probabilistic modeling yet effective for medical research.<sup>18,19</sup> The first milestone work after n-gram modeling is word embedding, which represents words in vector space to understand the natural language from a new quantitative perspective, promoting clinical research on document analysis.<sup>20,21</sup> Among a range of representation modeling, ELMO<sup>22</sup> proposed by AllenNLP changes the game to a bi-direction model pretraining. This modeling approach also influences medical language research<sup>20</sup> and is also evaluated.<sup>23</sup> Since then, bi-directional deep-learning models have been proposed like BERT<sup>24</sup> and Generative Pretrained Transformer (GPT).<sup>25</sup> Built upon these models, a range of medical language models are proposed to accelerate medical research progress, such as a family of medical BERT models,<sup>26–29</sup> and clinical researchers found that the increase in model size significantly improves a variety of medical applications.<sup>21,30–32</sup> However, they are limited to medium model

<sup>1</sup>State Key Laboratory of Oral Diseases & National Clinical Research Center for Oral Diseases & Department of Oral and Maxillofacial Surgery, West China Hospital of Stomatology, Sichuan University, Chengdu, China; <sup>2</sup>Department of Civil, Environmental & Construction Engineering, University of Central Florida, Orlando, USA; <sup>3</sup>College of Transportation Engineering, University of Central Florida, Orlando, USA; <sup>4</sup>School of Transportation and Logistics, Southwest Jiaotong University, Chengdu, China; <sup>5</sup>C2SMART Center, Tandon School of Engineering, New York University, Brooklyn, USA and <sup>6</sup>State Key Laboratory of Oral Diseases & National Clinical Research Center for Oral Diseases & Eastern Clinic, West China Hospital of Stomatology, Sichuan University, Chengdu, China

Correspondence: Hanyao Huang (huanghanyao\_cn@scu.edu.cn) or Ou Zheng (ouzheng1993@knights.ucf.edu)

These authors contributed equally: Hanyao Huang, Ou Zheng

Received: 26 March 2023 Revised: 6 July 2023 Accepted: 13 July 2023

Published online: 28 July 2023

scales due to architecture design and hardware support, although some efficient algorithms are proposed for the medical domain.<sup>20,33,34</sup> One of the most important LLMs is T5, with 11 billion parameters proposed by Google.<sup>35</sup> Another rival model is GPT-3, developed by OpenAI, which contains 175 billion parameters. These billion-parameter models embark on a new chapter of LLMs and their applications. One of the most successful application instances is ChatGPT, a variant of InstructGPT<sup>36</sup> developed upon GPT-3, optimized by conversational response training. ChatGPT is equipped with interactive training which involves human feedback reinforcement learning and exhibits powerful language skills to generate human-like texts in real-time conversation. This interactive modeling also influences medical research like education.<sup>37,38</sup> All these rely on large-scale representation pretraining, which becomes critical to complex problem-solving with data in cross-modality, even for ChatGPT.

#### Large-scale vision-language pretraining

Vision-language pretraining is an important approach to solving text-to-image or image-to-text tasks, which trains a deep neural network with large image and text datasets. One of the vital training frameworks is CLIP, proposed by OpenAI, which is further improved by Salesforce to BLIP.<sup>39</sup> For text-to-image, GAN as an image generation prototype model can be integrated with text representation to generate diverse, authentic-looking but synthetic images.<sup>40</sup> Recently, the Diffusion Model,<sup>41</sup> a rival model of GAN emerging with higher computation efficiency and image diversity, has been incorporated for vision-language representation pretraining. For example, DALL-E v2<sup>42</sup> leverages CLIP ranked representation and diffusion model to generate image understanding sentences. To address medical domain-specific problems, a series of efficient representation learning models are also developed to empower intelligent healthcare services. BERT provides an effective solution to efficient inference and analysis of disease.<sup>43</sup> Multi-modal learning is also considered to improve medical visual-questioning-answer processes.<sup>44</sup> Although various applications are proposed, the study on how to integrate this powerful model with dental diagnosis is still limited.

#### Large-scale audio-language pretraining

Compared to vision-language models, audio-language pretraining does not prevail, but the representation learning with audio and text data still exhibits impressive audio-to-text performance. Some medium-scale models like MusCALL,<sup>45</sup> CTAL,<sup>46</sup> Wav2Seq,<sup>47</sup> and LAVA<sup>48</sup> indicate the superiority of representation pretraining on speech recognition. One of the important large-scale pretraining models is Whisper,<sup>49</sup> released by OpenAI, which is trained on 680,000 h of diverse audio-text pairs from the web. Inspired by the success of these works, an improved medical speech-to-text pretraining model is developed to more effectively link vocal signals to language generation and understanding.<sup>50</sup> Since audio-language pretraining research is still under exploration, the limited study demonstrates how to employ this pretraining framework to facilitate oral treatment.

#### Multi-modal LLM

With tremendous success in cross-modal training, more research attempts to incorporate multi-modal representation learning to empower LLMs. As one of the successful attempts, GPT-4<sup>51</sup> demonstrates the competence of LLMs in a multitude of NLP applications, such as higher scores in GRE tests and other question-answering tasks. This implies a higher potential for Multi-Modal LLM in various areas, such as digital health. For example, multi-modal learning is conducted to facilitate medical services, which incorporates images, audio, and texts into training for a more comprehensive and robust model.<sup>43,45</sup> However, due to limited data availability, more research still attempts to explore the merit of multi-model LLM for medical fields, especially dental clinic research.

#### LLM as a ubiquitous solution

As LLMs become increasingly widely recognized, more representations will be embedded into the models to enhance their general problem-solving skills. The training process with a larger scale of data will yield a ubiquitous solution to problems of all kinds. For example, ChatGPT has served as a valuable tool to assist medical education for more effective instruction and analysis of teacher-student interaction.<sup>37,38</sup> Medical writing can be assisted or even accomplished by ChatGPT,<sup>32</sup> which enables efficient documentation. Language challenges in medical research or clinical processes can also be alleviated by ChatGPT.<sup>52</sup>

#### AI TECHNOLOGY FOR CLINICAL APPLICATION

AI technology has promoted clinical applications by improving patient outcomes, streamlining processes, and reducing costs. In clinical practice, AI has achieved striking success in analyzing patient data like brain-tumor segmentation,<sup>53</sup> assisting in clinical decision-making like epidemiological prediction,<sup>54</sup> and performing complex tasks such as surgery and rehabilitation, which indicates the potential to revolutionize healthcare service. In dentistry, the convolutional neural network has shown performance gain in detecting and classifying maxillofacial fractures from CT.<sup>55</sup> However, subtle details of maxillofacial fractures may not be accurately detected sometimes due to the unfavorable resolution of CT scans. Still, more advanced CT scanners can achieve higher-resolution images in future studies. Medical researchers also attempt to explore detection methods and investigate the feasibility of an automated decision-making tool for dental age estimation using deep-learning and topological approaches by analyzing the third molar maturity index (I3M) from 456 mandibular radiographs.<sup>56</sup> Another recent research proposed a more comprehensive AI system that can precisely identify individual teeth and alveolar bones from dental cone-beam CT (CBCT) images, which enables accurate and precise dental healthcare.<sup>57</sup>

The success of language modeling also promotes lots of research progress in representation learning for efficient medical services. For example, BioBERT<sup>58</sup> is developed upon BERT to achieve a large but efficient text mining model for biomedical document analysis. ClinicalBERT<sup>59,60</sup> carries out embedding training with a large volume of clinical documents to facilitate intelligent clinical diagnosis processes. SciBERT<sup>27</sup> also built a large language model for representation learning with multiple documents across various scientific research domains.

One of the milestone contributions to biomedical research is AlphaFold,<sup>61</sup> developed by DeepMind. Its success in accurate 3D protein structure prediction demonstrates the power of large-scale training to tackle significant challenges in quantitative biomedical modeling. Since then, a series of innovative large-scale frameworks have been proposed to enhance AI-powered modeling, such as OpenFold<sup>62</sup> by OpenAI and BioNeMo Megatron<sup>63</sup> by NVIDIA, etc. In addition, inspired by LLM pretraining schemes, NVIDIA developed ProT-VAE<sup>64</sup> to advance functional protein design, which indicates the potential of large-scale biomolecule-language pretraining with an LLM. As technology continues to evolve, we can expect to see even more innovative applications of AI in clinical settings, ultimately leading to more effective healthcare services tailored to the needs of patients.

#### ON EXPLORING THE CAPABILITY OF LLMs IN DENTISTRY

Automated dental diagnosis with an LLM

*Record analysis with text mining.* Contemporary medical practice widely adopts electronic health records (EHRs) for patient information documentation. Although it facilitates record generation and management, efficient analysis is still challenging since massive amounts of records are mixed with structured and

unstructured data. This challenge leads to substantial amounts of underutilized data and obstructs patient care and research improvement. Text mining is able to tackle this challenge by drawing valuable conclusions and information from textual material in a mixed structure. Some straightforward modeling frameworks have been developed to explore patterns, correlations, and trends within textual data.<sup>65-68</sup> However, the performance of these models is inadequate to process massive amounts of documents efficiently and accurately.

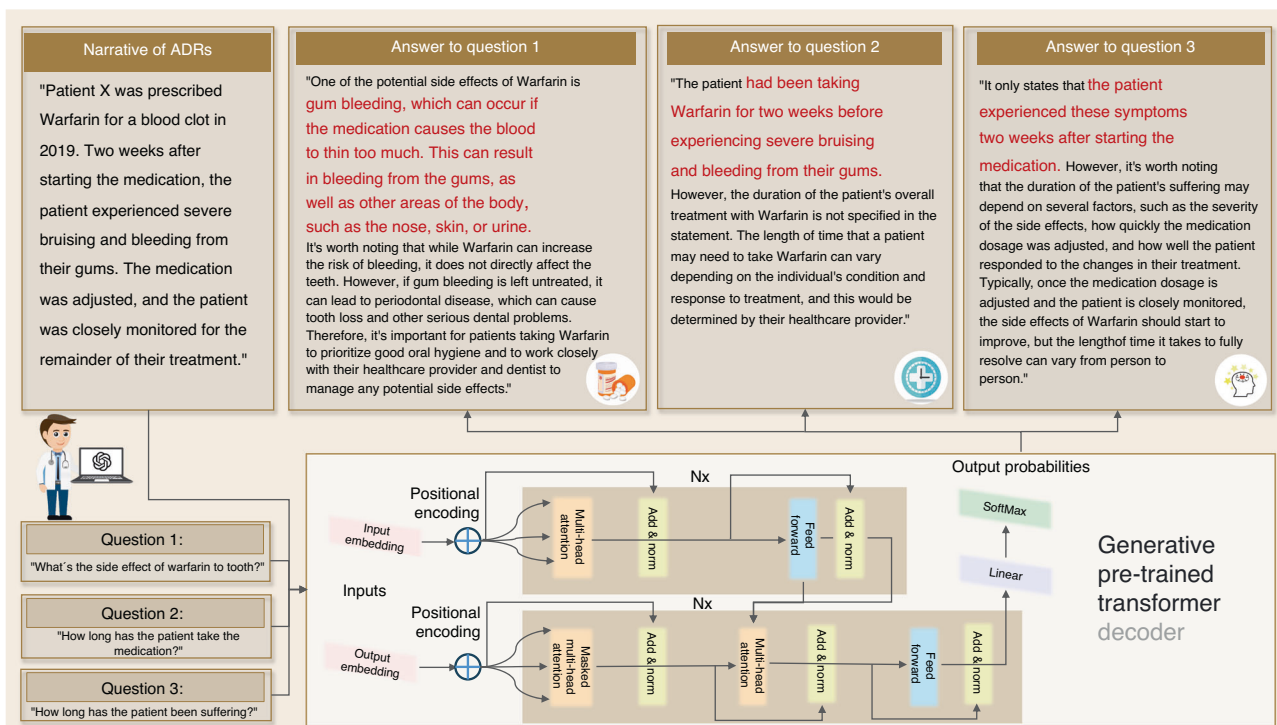
An LLM helps to find a workaround for this limitation through training on extensive documents. Given strong competence in semantic understanding, an LLM can manage documents independent from structural formats. As shown in Fig. 1, text mining can also retrieve pertinent facts from unstructured data, such as free-text notes from healthcare professionals. From this unstructured data, an LLM like ChatGPT can be used to swiftly extract pertinent information, like a patient's unique illnesses or adverse effects.

**Treatment planning with natural language reasoning.** As aforementioned, medical service experiences the influx of a large volume of digital information. In addition to straightforward document analysis, these data can assist healthcare providers in customizing treatment plans.<sup>69</sup> Although this data analytics is stimulating, the work is taxing since more labor is involved in document analysis. LLMs can easily automate document comprehension and make treatment plan analysis feasible, which reaps the benefit of large-scale pretraining. Furthermore, billions of documents help LLMs cultivate the capability of natural language reasoning (NLR) to perceive contexts. This capability of NLR can assist dental practitioners more efficiently in establishing treatment plans tailored to patients' backgrounds.<sup>70</sup> For instance, NLR algorithms can examine adverse drug reaction (ADR) patterns linked to various dental procedures and drugs.<sup>71</sup> Sometimes, drug administration can lead to gum bleeding and severer diseases like bisphosphonate-related osteonecrosis. Dentists can modify their treatment plans to lower the likelihood of side effects by

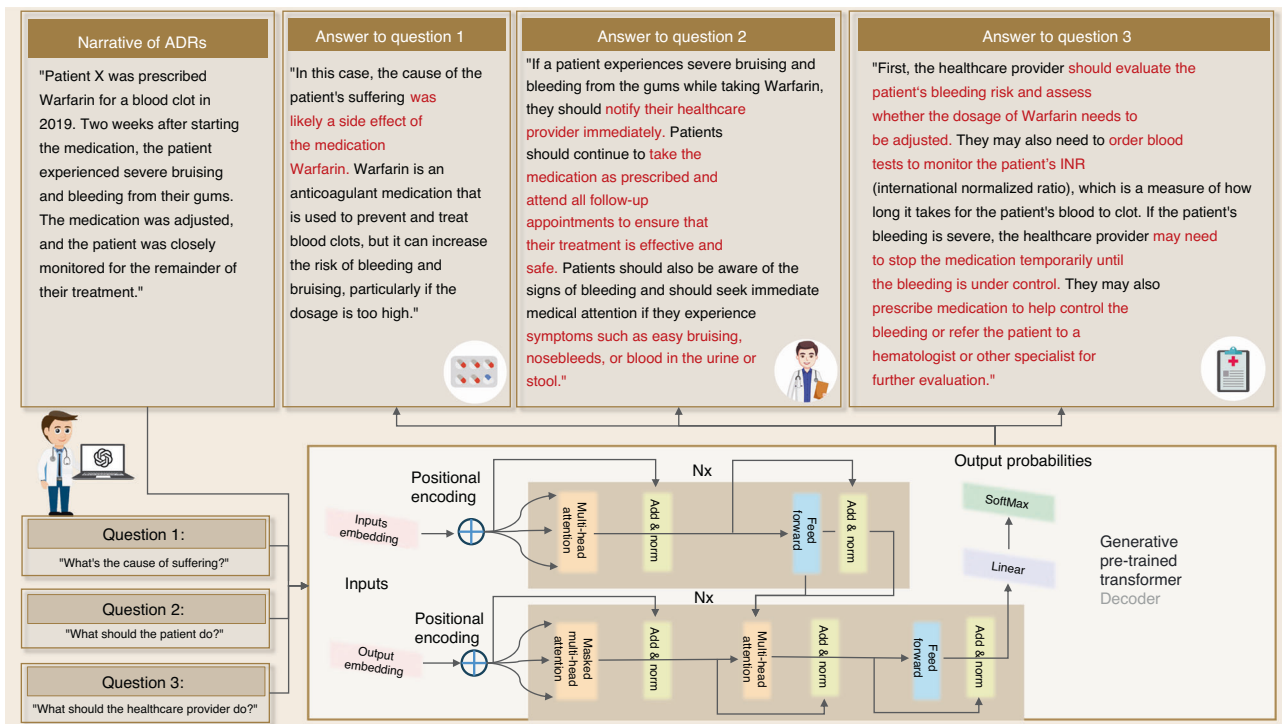
understanding the most typical ADRs linked to particular medications. We maintain that an LLM can facilitate this process and provide a case of this application in Fig. 2. It has been found that NLR may be used to identify comorbidities by analyzing patient records for common risk factors and symptoms, identification of ADRs,<sup>71</sup> drug safety surveillance,<sup>72</sup> and patient education.<sup>73</sup>

**Medical documentation with natural language generation.** In dental clinical practice, a synthetic yet faithfully representative EHR is essential to efficient medical information conveyance between healthcare providers and other medical professionals. Traditionally, this document preparation process is completed manually. Given keywords, practitioners organize the context following medical record documentation standards. It can be quickly assisted with an LLM. Natural Language Generation (NLG) is one of the important tasks of LLM. Generally, NLG accomplishes text generation given the understanding of natural language input, like structured texts or separate keywords. Since a well-trained LLM is highly skilled in this task, this merit can be leveraged to automate a series of common documentation tasks, such as generating reports on medical history, dental procedures, and treatment plans. For example, Fig. 3 shows an example of medical report narrative generation with only several keywords by ChatGPT. The generated texts are well formatted in a professional structure and encompass all important information in eloquent expression. The validity of the generated content has been validated.<sup>74,75</sup>

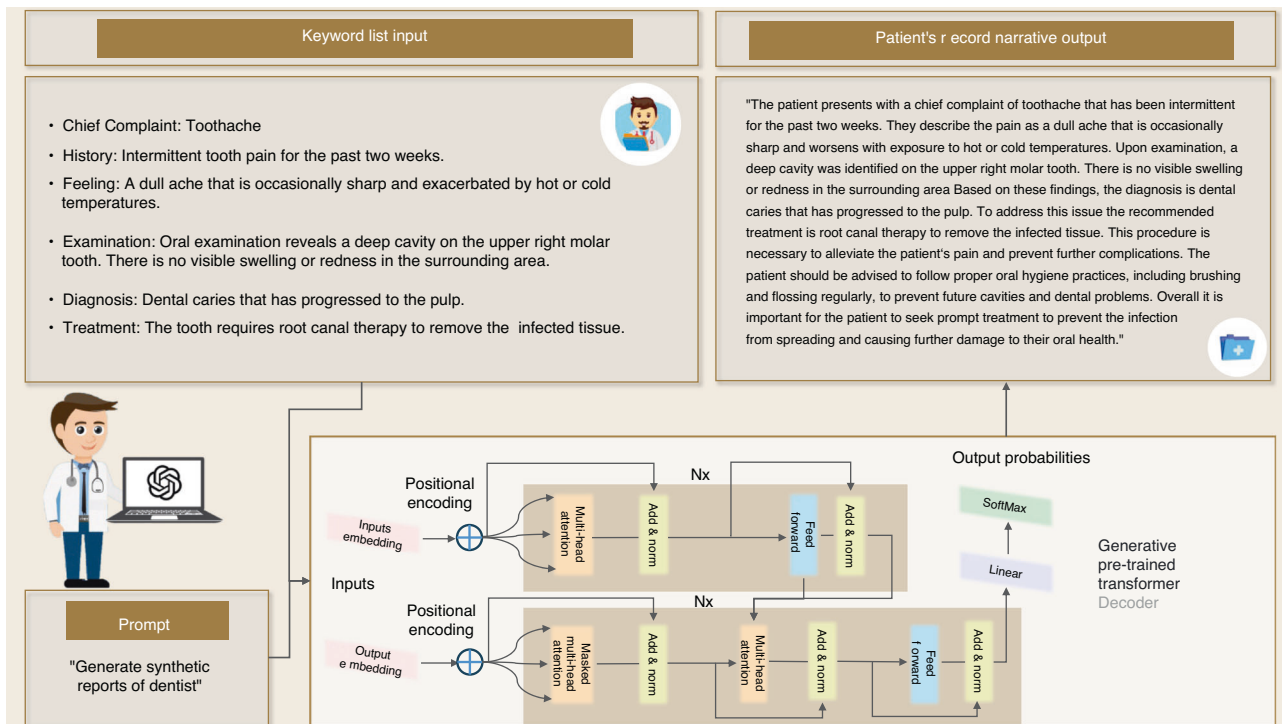
Compared to other application fields, the influence of AI in dentistry has unquestionably been slower and more constrained. This is mostly because patient privacy concerns have prevented patient data from being made widely available to the broader AI research community. Meanwhile, training data plays a crucial role in the advancement of AI methods. To address this dilemma and promote methodological advances in dentistry, using high-quality synthetic quasi-EHRs data is a practical approach since it facilitates data sharing between healthcare providers and external investigators. We employ ChatGPT 3.5, configured with specific



**Fig. 1** Examples of a text mining application for extracting a patient's historical record with an LLM like ChatGPT. Keywords are highlighted in red



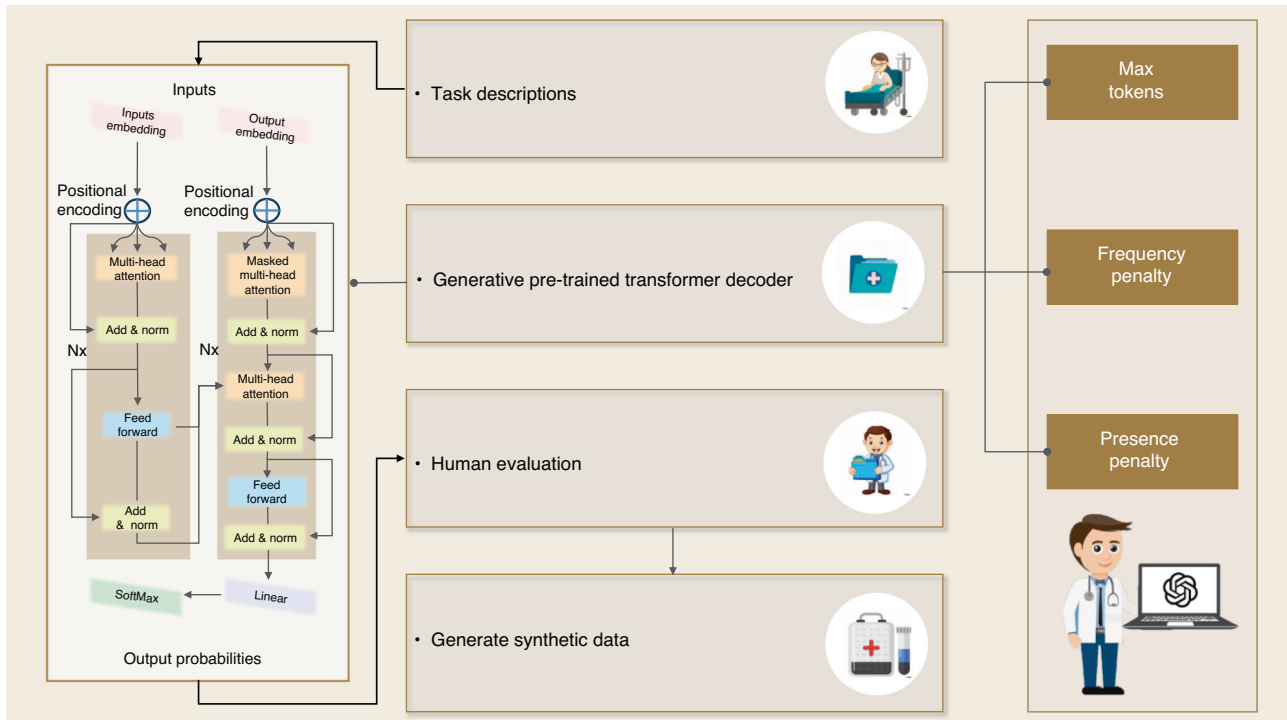
**Fig. 2** Examples of natural language reasoning (NLR) application of medication suggestions derived from a patient's record. Keywords are highlighted in red



**Fig. 3** Example of a narrative output of the Patient's record generated from keywords with NLG

parameters, to generate synthetic data, shown in Fig. 4. The parameters include max tokens, frequency penalty, and presence penalty, which were set to enhance diversity in the generated text. The frequency penalty reduces the likelihood of selecting words based on their frequency of occurrence, while the presence penalty imposes a fixed cost on each word in the text. These penalties encourage the model to generate text with higher

perplexity rather than relying solely on the most probable word choices. Additionally, temperature scaling is used to adjust the distribution of probabilities for the next tokens, and a top-p value of 1 ensures consideration of all available tokens. Post-processing is applied to refine the generated data to eliminate any artifacts introduced during the generation process. These post-processing rules are determined through manual examination. These data



**Fig. 4** Framework of generating synthetic quasi-EHRs data by LLMs

can be quickly generated and harvested with the assistance of LLMs. Synthetic EHRs can be more realistic by introducing variability in the generated data. LLMs can be guided to generate different patient profiles, medical histories, treatment plans, and outcomes. This helps mimic the diversity and complexity seen in real EHRs. Thus, it implies that an LLM has competence in efficiently preparing medical information and protecting the privacy of patients.

#### Cross-modal dental diagnosis with LLMs

##### *Vision-language deployment*

**Visual grounding:** Traditionally caries-related diagnosis is administered by dentists through visual and tactile examination. Before any treatment plan, a quick but comprehensive examination of oral health conditions is imperative. It sometimes takes experts' effort and time to diagnose tooth conditions, possibly through X-ray images and CBCT, and reach a reliable conclusion. Some existing research has explored the potential of AI-assistant models in assisting diagnosis for caries,<sup>3</sup> periodontitis,<sup>5</sup> medication-related osteonecrosis,<sup>76</sup> maxillofacial bone fracture,<sup>55</sup> oral squamous cell carcinoma,<sup>77</sup> and temporomandibular disorders.<sup>78</sup> These diseases can be diagnosed based on medical imaging. Also, AI-assistant models for imaging analyses show the potential in assisting dental treatment, including orthodontics,<sup>79</sup> restorative dentistry,<sup>80</sup> oral implantology,<sup>6</sup> and oral and maxillofacial surgery.<sup>7</sup>

However, limited data representation hinders accurate diagnosis and treatment planning when the disease is intricate. The majority of study has been confined to image-only approaches, which restrain the effective conveyance of information and explore the untapped potential of AI models in dentistry. LLMs open the possibility of data-fused diagnosis by leveraging cross-modal perception. An LLM is highly skilled in aligning textual and visual representations for image-text analysis, which can facilitate the diagnosis of tooth problems by X-ray image interpretation.

Specifically, the inference by an LLM can be blended with specific visualization techniques to identify caries regions. For example, the practitioners can provide some keywords to query the model of ALBEF (A Lite BERT for Adaptive Embedding Factorization), which is

specifically designed for image-to-text tasks and is integrated with Grad-CAM (Gradient-weighted Class Activation Mapping) to visualize the critical region for decision-making from the ALBEF model. The warmer color indicates the plausible regions corresponding to the described words. As shown in Fig. 5, root canal therapy is plausibly required in the region with a warmer color. Another tangible benefit of an LLM is training cost reduction. Without fine-tuning a large set of image data, an LLM can provide plausible affected teeth and likely locations of dental problems.

**Visual question answering:** In addition to visual examination of medical imaging data, diagnosis documentation is more critical to patient-centered care. From the interpretation of X-ray images, healthcare professionals will write down the observation, analysis, and medication suggestions to patients. These documents are also essential to healthcare big data analytics, while the document summarization on the medical transcripts takes much time. An LLM is able to reduce the processing labor significantly through specific tasks, like visual question answering (VQA). Commonly, a VQA model can convert the encoded image representations to word embedding for dental diagnosis questions. With the diagnosis questions, the answers are generated to facilitate diagnostic report generation. This VQA-assisted diagnosis can be performed to assess potential dental health issues.<sup>81</sup> As shown in Fig. 6, the X-ray image of a patient's teeth is fed into an image encoder, like BLIP-2, generating a natural language representation, i.e., embedding based upon image understanding. Meanwhile, different questions are fed into LLM to generate another set of question embeddings. Both question and image embeddings are mathematically combined to generate the interpreted answers to the questions about images.

However, sometimes the raw image contains too much noise, or the resolution is not acceptable by the BLIP model; therefore, it is difficult to extract the desired information through the VQA model. To address this, training a semantic segmentation model that divides the areas of the image with different properties into different classes is a potential solution because it allows an LLM to learn each element separately.<sup>82</sup> For example, as Fig. 7 shows, the

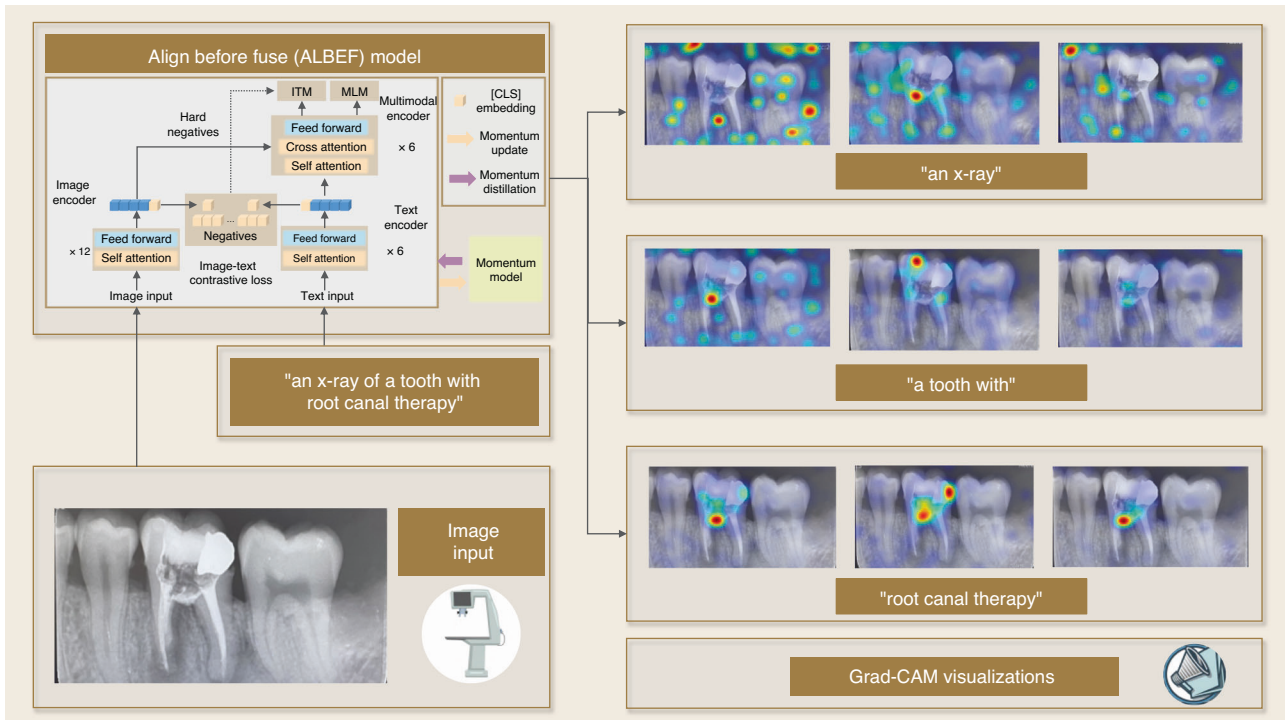


Fig. 5 Schematic of dental condition diagnosis with a vision-language model of ALBEF

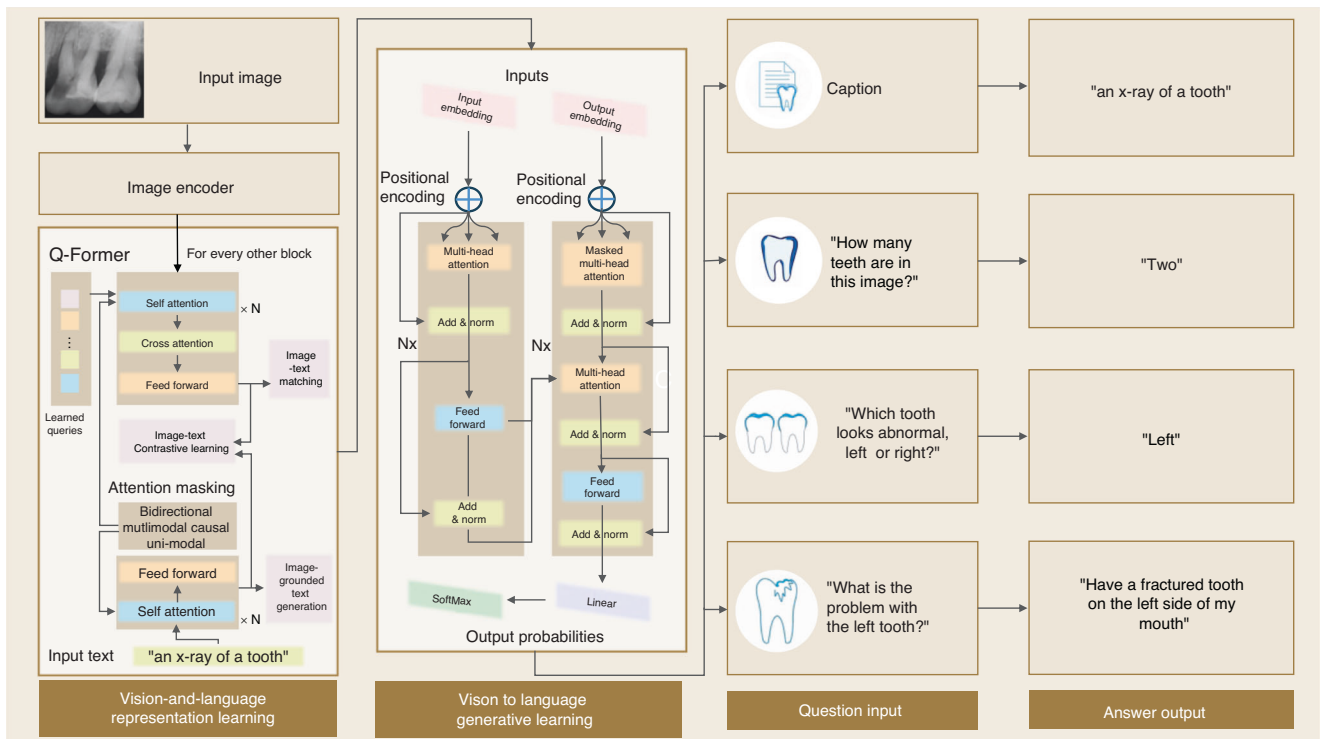
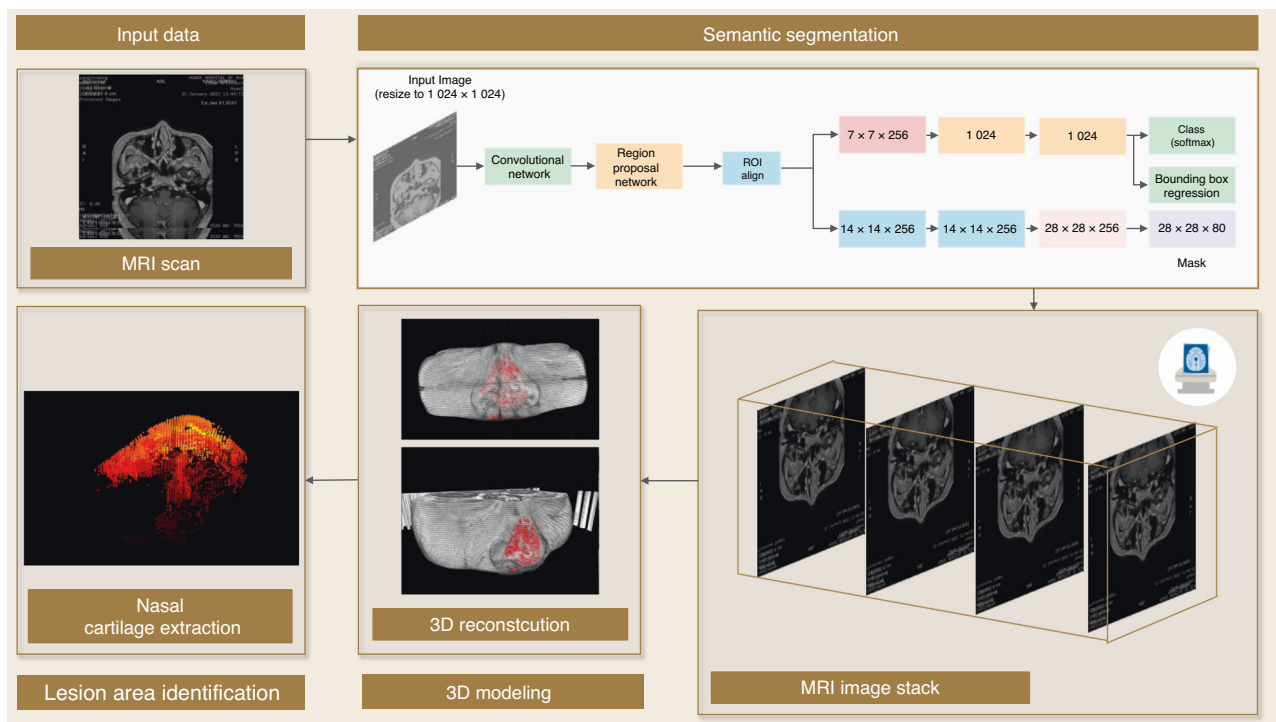


Fig. 6 A VQA example framework with the assistance of BLIP-2

soft tissue envelope and nasal septum/concha are classified into orange and blue segments, respectively, and it is expected to improve the model performance and enhance image understanding to extract the morphology information of nasal cartilages, as the cartilages are small and embedded by the soft tissue.<sup>83</sup> Figure 7 also demonstrates the potential of training

machine learning models to help reconstruct the nasal cartilage based on MRI for patients with orofacial clefts, who can suffer significant nasal deformity.<sup>84–87</sup> Due to the limitations of current imaging software, the differences between the cartilage and soft tissue cannot be easily defined, but with training machine learning models, the purpose should be achieved in the future.



**Fig. 7** 2D semantic segmentation with 3D reconstruction for lesion identification

Visual data generation: DALL-E 2, empowered by integrating GPT-3.5 as encoder and diffusion model as decoder, can be utilized in the medical field to generate synthetic medical images. Once the LLM has been trained on real EHR data, it can be used to generate synthetic quasi-EHRs. By providing prompts or specific instructions to the model, such as patient characteristics or medical conditions, the LLM can generate realistic synthetic records that resemble real-world EHRs. For example, by describing a patient's CBCT scan in text, including details of any abnormalities, like an odontogenic cyst or alveolar cleft, which can observe obvious lesions on the alveolar bone structure, and feeding it into DALL-E 2, synthetic medical images that match the description can be efficiently produced in large quantities to improve the performance of deep-learning models by use as the training dataset. Figure 8 shows that synthetic medical images can be generated with varying levels of noise, contrast, or resolution to create images with specific properties or characteristics that are relevant to the medical condition being studied. Moreover, patient privacy can be protected since synthetic medical images are generated from textual descriptions rather than real medical data. This technique is valuable for medical research based on any medical 3D imaging techniques, including CBCT, CT, MRI, etc., and improving patient care by training machine learning models while maintaining patient privacy.

Alternatively, LLMs can also be employed to generate medical illustrations or diagrams based on textual descriptions. For example, a description of a surgical procedure can be fed into illustration software to create an accurate and detailed illustration of the procedure.

**Audio-language deployment.** Besides imaging and dialogs, a patient's voice is also critical to medical diagnosis. A person's voice can potentially reveal important clues about their speech function, as certain vocal characteristics may be indicative of the function of teeth, tongue, pharyngeal structure, and muscles. Analyzing these vocal attributes can assist healthcare professionals in identifying potential health concerns. One of

the common medical diagnosis applications is waveform-spectrogram analysis on patients' audio recordings, which are collected by requesting the patients to read certain words or paragraphs. The waveform is a curve-based representation of an audio signal, the shape of which enables acoustic analysis. The spectrogram is an alternative representation of sounds in the frequency domain, which facilitates signal processing and analysis.

Velopharyngeal insufficiency related to cleft palate, oronasal fistula, and so forth, that affects the contact between the soft palate and posterior pharyngeal wall, or changes the needed separation between the oral cavity and nasal cavity,<sup>87–89</sup> exhibits some typical marks on voice waveforms and spectrograms. In velopharyngeal insufficiency, for example, nasal emission can lead to distinct variation in speech.<sup>90</sup> A person with velopharyngeal insufficiency may exhibit a waveform that shows the reduced intensity of the sound waves during certain frequencies or periods, leading to altered speech patterns because of the abnormal airflow in oral and nasal cavities.<sup>91–94</sup> Figure 9 shows an example of a comparison of the waveforms and spectrograms between normal people and patients with velopharyngeal insufficiency. It can be observed that normal people have a more intense waveform and continuous spectrogram, while the patients' samples are more dispersive and broken.

Furthermore, the waveforms and spectrograms of different patients can be fed into the pretrained LLMs such as GPT-4 for potential disease and severity diagnosis. As shown in Fig. 10, a pair of graphs are inputted into GPT-4 while asking for disease deducing, and the model provides several answers for reference. Although the final answers show little about the velopharyngeal insufficiency, the output mentions muscle dysfunction. Also, a more precise output can be achieved by further fine-tuning with more labeled patients' audio data. In addition, NLG can also be used in conjunction with speech recognition software to convert voice commands into written text, such as when dictating clinical notes or treatment plans.

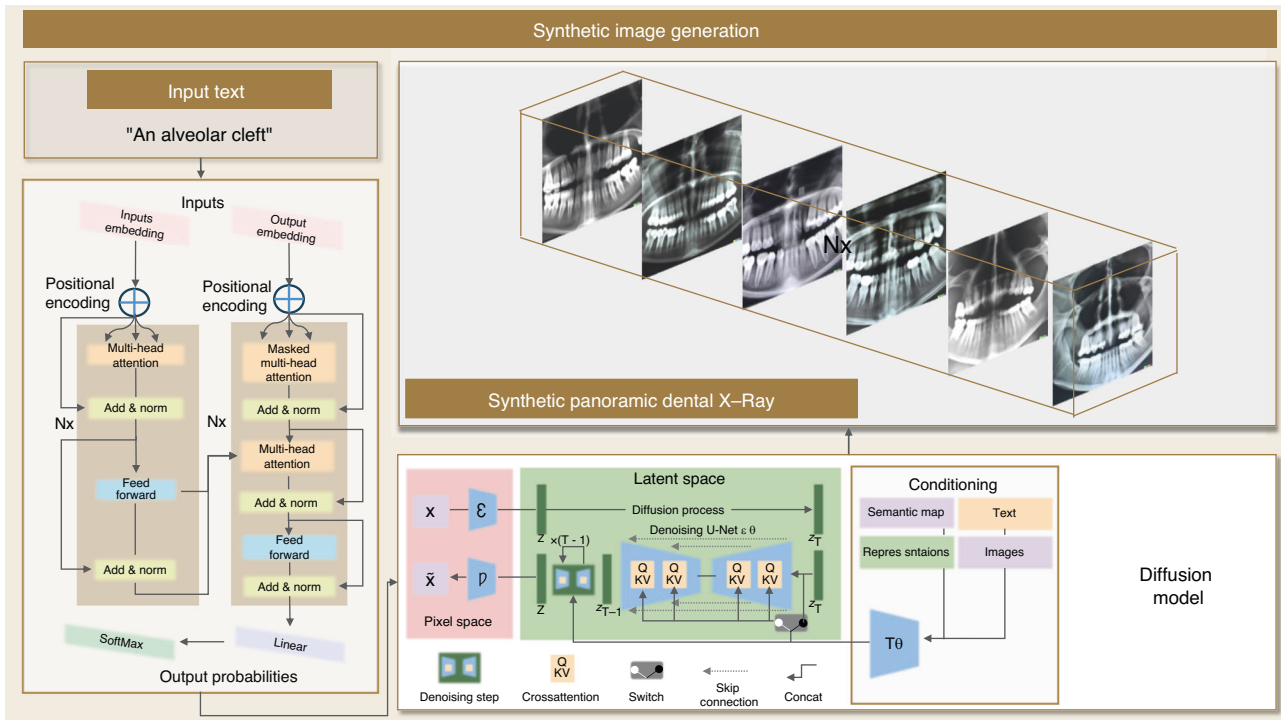


Fig. 8 Example of visual data generation

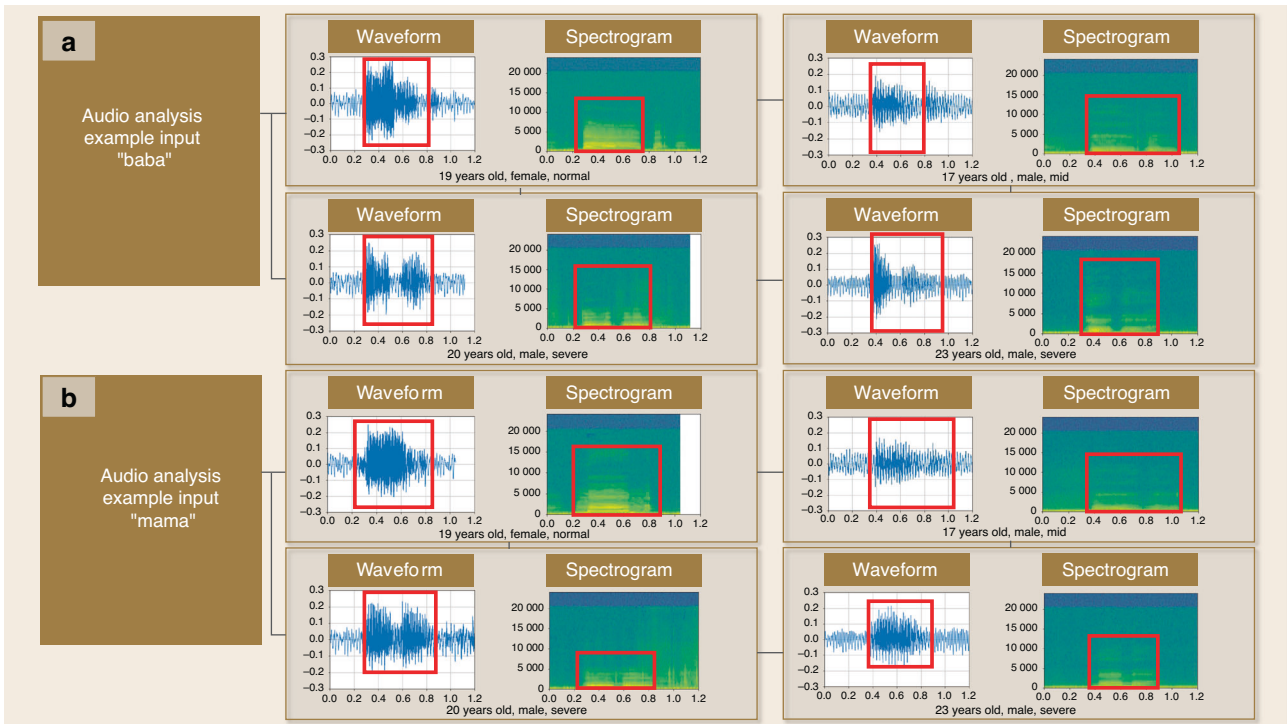


Fig. 9 Example of the audio waveform and spectrogram analysis in speaking a "baba" and b "mama" with TorchAudio of normal people and patients with velopharyngeal insufficiency

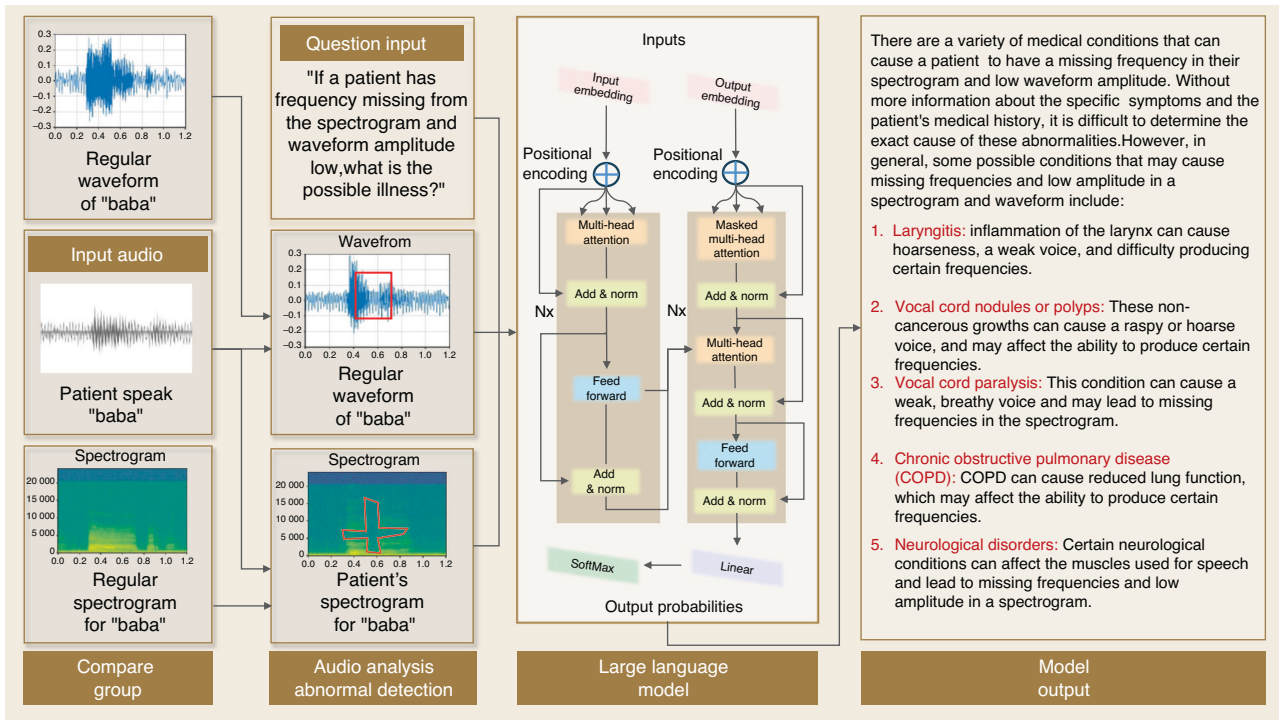
**OTHER POTENTIAL CROSS-MODAL DEPLOYMENTS**

**Biopsy**

A biopsy is a medical procedure in which a small sample of tissue is removed from a person's body to be examined under a microscope. This is typically done to diagnose or investigate

various conditions, such as cancer, infections, or inflammation. Visualization techniques can be applied to biopsy to understand the tissue structure and cell morphology better and identify any abnormalities. Common visualization approaches include light microscopy, immunohistochemistry (IHC), immunofluorescence





**Fig. 10** Schematic of audio-language assisted diagnosis based upon audio waveform and spectrogram analysis with TorchAudio

(IF), confocal microscopy, etc. With the embedding generated by a vision transformer, the image input can be projected into the language space and used for understanding the characteristics of disease identification. For instance, prostate cancer in biopsies<sup>95</sup> and pre-implantation kidney biopsy pathology practice<sup>96</sup> has been regarded as potential application fields, which also shows the potential for biopsy and histological analyses in dentistry, and oral and maxillofacial surgery.<sup>97,98</sup>

#### Blood test

LLMs can help users understand the results of their blood tests by providing information on the normal ranges for different biomarkers and explaining the potential implications of high/low values. The changes in the test indicators also provide rich information about the human body condition, as well as track the recovery or disease deterioration process. These changed conditions can affect the patient's treatment planning and treatment for dental problems.<sup>99</sup> For instance, anemia may present with low hemoglobin, hematocrit, and red blood cell count, while liver disease may present with elevated liver enzymes (ALT, AST, and ALP). These abnormal parameters may postpone the treatments like oral and maxillofacial surgeries as most of these surgeries are elective. The internal relationships and connections between these indicators can be well captured by LLMs, and thus the potential diseases can be linked with the inputted information.

#### Gene detection

Gene detection is the process of identifying and analyzing specific genes or genetic sequences. Classic or more recent approaches can help obtain the gene sequence, including RNA sequencing, DNA sequencing, single-cell sequencing, etc. The genes or genetic sequence can be projected into language embeddings with the corresponding encoder and then input to LLMs. As the sequence contains the underlying logic of the gene's property, the LLMs can help to understand these logics by learning from large gene samples after training.<sup>100,101</sup> Potential applications may include understanding gene function, identifying genetic variations or mutations, and studying the relationships between genes and various biological

processes or diseases, which can further influence the development of dental problems and treatments related to genetic disorders.

### AI SYSTEM FOR DENTISTRY APPLICATION WITH A FULLY AUTOMATIC MULTI-MODAL LLM

To demonstrate the effectiveness and potential of LLMs' application in dentistry, we present a framework of a fully automatic diagnosis system based on Multi-Modal LLMs. The system consists of three input modules from different models: vision input, audio input, and language input.

The image input could be dental X-ray, cone-beam computed tomography, and other medical imaging. For semantic classification, we focus on optimizing the capture of the critical elements. By applying vision-language models, the condition of the tooth is evaluated, potential anomaly or disease is detected, and specific diagnosis and corresponding suggestions can be given.

In this case, audio sources have two usages: voice anomaly detection and patients' narrative understanding. For the first usage, the system receives the voice input from patients, plots waveform and spectrogram, and then performs amplitude and frequency analysis. For the second usage, the patients' narratives are collected and converted into texts using speech recognition techniques. Afterward, the key elements, like the symptoms that patients stated, can be extracted and summarized to form reports or bullet points for doctors' reference.

Targeting automatic diagnosis for dentistry, the AI system can be embedded into the dental clinics' internal communication systems. Thus, a fully developed automatic application can encompass patients' information from multiple sources and provide a professional medical diagnosis, as shown in Fig. 11.

### A SPECIFIC CASE FOR THE MULTI-MODAL LLM AI SYSTEM FOR DENTISTRY CLINICAL APPLICATION

To demonstrate the application of the multi-modal LLM AI system in dentistry, we use a sample with dental caries to explain how it works by LLM, including vision-language modeling and treatment planning

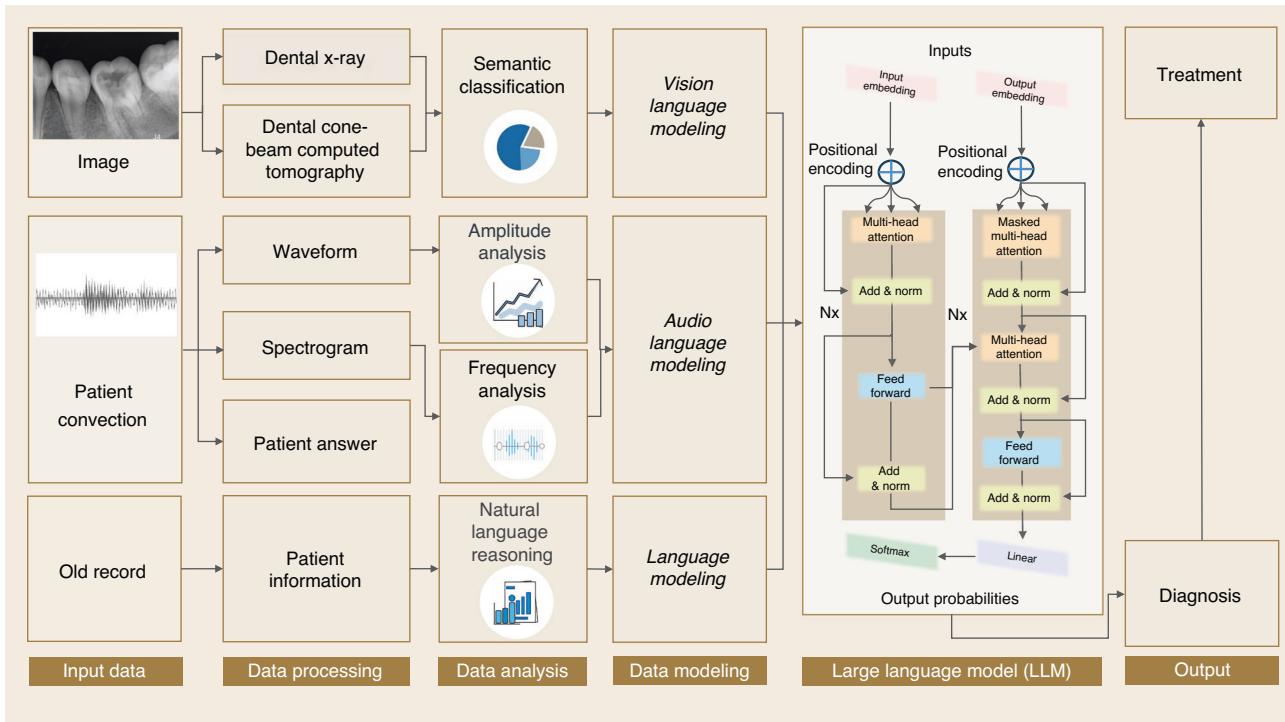


Fig. 11 Concept of automatic multi-modal LLM AI system for dentistry application

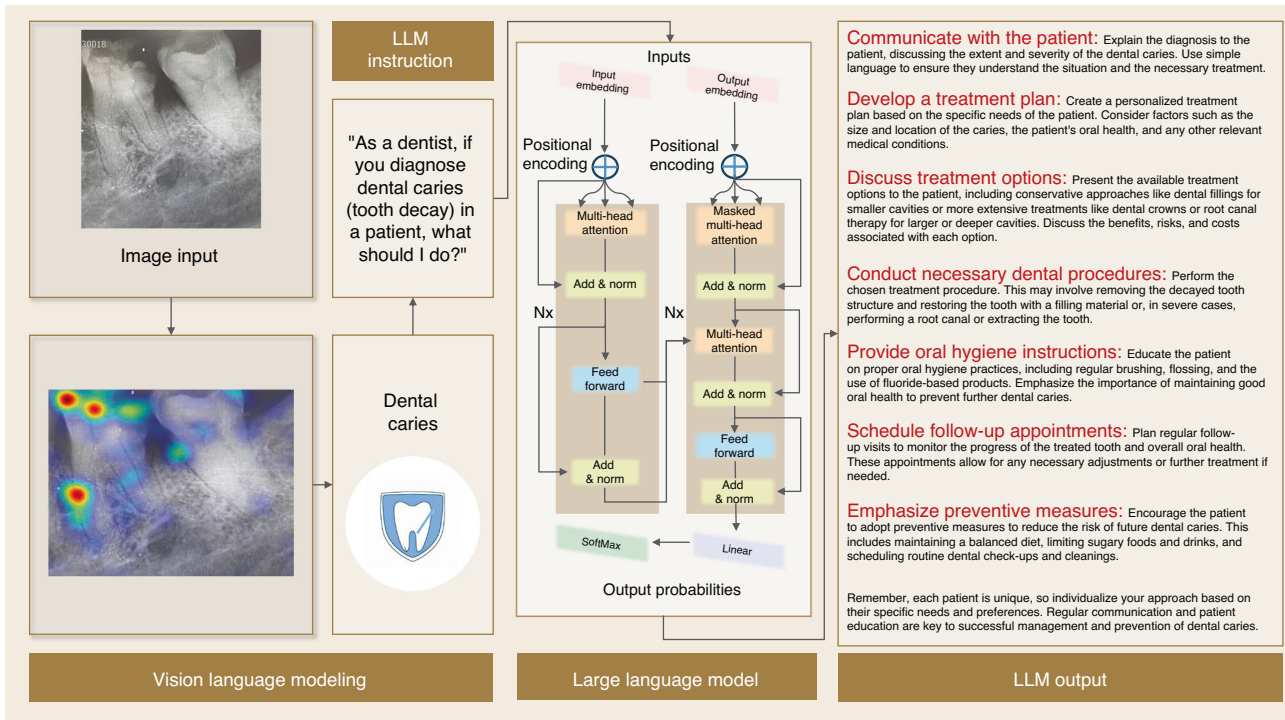


Fig. 12 Application of the multi-modal LLM AI system in dental caries

with natural language reasoning. As shown in Fig. 12, an X-ray of the tooth is inputted into the system, and the abnormal morphology, like decay on the tooth, can be located on the X-ray by vision-language modeling, and then the first question can be answered that dental caries can be found on the tooth. Then the next question becomes what the treatment planning for this problem is, and using the LLM again to output seven steps, including communicating with the

patient, developing a treatment plan, discussing treatment options, conducting necessary dental procedures, providing oral hygiene instructions, scheduling follow-up appointments, and emphasizing preventive measures. However, from the X-ray, we can also observe potential bone loss near the distal root, which is not detected by the system according to this pilot study. Thus, further study should be done to improve the system.

## ISSUES AND LIMITATIONS

While there is much excitement around the potential applications of LLMs in the field of dentistry, some issues and limitations must be addressed before these models can be widely adopted.

### Data quality

Despite rigorous efforts to sanitize and filter the vast amount of training data, it is challenging to eliminate all harmful or inappropriate content, which may inadvertently propagate through the responses generated by LLMs.<sup>102</sup> Inherently, these LLMs operate as sophisticated pattern-matching machines without a genuine understanding of the data they are trained on, which occasionally leads to nonsensical or inappropriate responses.<sup>103</sup> Compounding these issues, they lack the capacity to validate the information they generate, remaining incapable of accessing real-time data or verifying the current status of events post-training. Moreover, the knowledge base of an LLM is static, established at the time of training, thus unable to update its knowledge or assimilate new developments in the evolving data landscape. One possible solution could be to develop "Human-in-the-loop system". Pairing LLM systems with human supervisors can safeguard important decisions, helping to catch and correct mistakes that the LLM might make<sup>104</sup>.

### Model bias

Biased clinicopathologic analysis results are the first noteworthy issue. Because LLMs are data-driven-only models which learn the features and patterns in the training data, the correctness of the LLM is highly dependent on the quality and adequacy of the data.<sup>105</sup> Although the LLMs are evolving iteratively, even in the era of GPT-4, we can't fully trust the AI-generated clinicopathologic analysis results, and human-in-the-loop validating work is still necessary. In the future, neural-symbolic models, which can combine two approaches (neural networks and symbolic reasoning) by using neural networks to learn the statistical patterns in large datasets and then using symbolic reasoning to perform logical operations on the learned representations, can be a potential research direction.<sup>106</sup>

### Data privacy

The patient data breach is another big issue, especially in today's privacy-sensitive world. Fine-tuning the LLMs in the dentistry domain is expected, and a huge amount of patient data is necessary. A data breach is likely to happen during this process if the healthcare providers and developers don't take appropriate measures to safeguard patient data. It's crucial to implement strict data handling protocols and use secure communication channels for transmitting and storing patient data. In addition to these security measures, it is equally essential to inform patients about using their data in advance and obtain their consent.<sup>107</sup> Another possible point of data leakage is dental diagnosis. Inputting patient data is necessary for these diagnosing applications, where there is a risk of violating patient privacy and confidentiality. One possible solution to address this concern is to use offline LLM such as META LLaMA, where the LLM is run locally on the device or edge server rather than on a centralized server API Call.<sup>108</sup>

### Computational cost

Computational resources can also limit the application of LLMs in dentistry. It is reasonable to expect that the LLMs in the dentistry domain will be operated locally due to data sensitivity. Firstly, fine-tuning the LLMs in the dentistry domain using local computational resources can be challenging. Then, running a full LLM to support the application in the dentistry domain is a waste of computational resources and unnecessary. A sparse expert model, a type of LLM that incorporates a set of specialized expert models, can be a future solution. It can reduce the computational resources required to train

and run LLMs while handling specific tasks or domains more efficiently than the main LLM.

## CONCLUSIONS

The utilization of language models like ChatGPT holds significant potential for advancing clinical applications and research in dentistry. By employing these models in a rational manner, a paradigm shift can be achieved in dental diagnosis and treatment planning. Further exploration based on diverse medical examination data will facilitate the realization of precision medicine and personalized healthcare in dentistry. A crucial future endeavor of practical deployment involves fine-tuning language models with dentistry domain-specific knowledge. This entails training the models with dentistry teaching materials, patient records, and other relevant domain information, resulting in enhanced accuracy by capturing pertinent patterns, terminology, and context. Consequently, the models acquire a profound comprehension of dentistry concepts, enabling them to generate contextually relevant and insightful responses. Customizing the outputs in alignment with domain requirements and preferences enhances efficiency, saving valuable time and resources. These benefits substantially contribute to improved performance and usability, rendering fine-tuned language models invaluable tools for research paper composition. Concurrently, the adoption of LLMs will further reduce medical costs and enhance medical efficiency.

## AUTHOR CONTRIBUTIONS

H.H. and O.Z. contributed to the conception and design of the work. H.H., O.Z., D.W., J.Y., Z.W., S.D., H.Y., C.X., and R.Y. performed data acquisition, analysis, and interpretation. O.Z. and D.W. performed data acquisition. H.H., O.Z., and D.W. performed analysis and interpretation. H.H., O.Z., D.W., J.Y., Z.W., S.D., H.Y., C.X., R.Y., Q.Z., and B.S. drafted and critically revised the paper. All authors gave final approval and agreed to be accountable for all aspects of the work.

## FUNDING

This work was supported by the Research and Development Program, West China Hospital of Stomatology, Sichuan University (RD-02-202107), Sichuan Province Science and Technology Support Program (2022NSFC0743), and Sichuan Post-doctoral Science Foundation (TB2022005) grant to H. Huang.

## ADDITIONAL INFORMATION

**Competing interests:** The authors declare no competing interests.

## REFERENCES

1. Kurian, N., Cherian, J. M., Sudharson, N. A., Varghese, K. G. & Wadhwa, S. AI is now everywhere. *Br. Dent. J.* **234**, 72–72 (2023).
2. Johnson, S. B. et al. Using ChatGPT to evaluate cancer myths and misconceptions: artificial intelligence and cancer information. *JNCI Cancer Spectr.* **7**, <https://doi.org/10.1093/jncics/pkad015> (2023).
3. Mohammad-Rahimi, H. et al. Deep learning for caries detection: a systematic review. *J. Dent.* **122**, 104115 (2022).
4. Urban, R. et al. AI-assisted CBCT data management in modern dental practice: benefits, limitations and innovations. *Electronics* **12**, 1710 (2023).
5. Revilla-León, M. et al. Artificial intelligence models for diagnosing gingivitis and periodontal disease: a systematic review. *J. Prosthet. Dent.* <https://doi.org/10.1016/j.prosdent.2022.01.026> (2022).
6. Mohammad-Rahimi, H. et al. Deep learning in periodontology and oral implantology: a scoping review. *J. Periodontol. Res.* **57**, 942–951 (2022).
7. Minnema, J. et al. A review on the application of deep learning for CT reconstruction, bone segmentation and surgical planning in oral and maxillofacial surgery. *Dentomaxillofac. Radiol.* **51**, 20210437 (2022).
8. He, L. et al. Automatic initial and final segmentation in cleft palate speech of Mandarin speakers. *PLoS ONE* **12**, e0184267 (2017).
9. Fu, J., He, F., Yin, H. & He, L. Automatic detection of pharyngeal fricatives in cleft palate speech using acoustic features based on the vocal tract area spectrum. *Comput. Speech Lang.* **68**, 101203 (2021).

10. Thurzo, A., Strunga, M., Urban, R., Surovková, J. & Afrashtehfar, K. I. Impact of artificial intelligence on dental education: a review and guide for curriculum update. *Educ. Sci.* **13**, 150 (2023).
11. Zheng, O., Abdel-Aty, M., Wang, D., Wang, Z. & Ding, S. ChatGPT is on the horizon: could a large language model be all we need for Intelligent Transportation? Preprint at arXiv:2303.05382 (2023).
12. Schuppe, K. et al. Atypical Nelson syndrome following right partial and left total nephrectomy with incidental bilateral total adrenalectomy of renal cell carcinoma: a chat generative pre-trained transformer (ChatGPT)-assisted case report and literature review. *Cureus* **15**, e36042 (2023).
13. Şendur, H. N., Şendur, A. B. & Cerit, M. N. ChatGPT from radiologists' perspective. *Br. J. Radiol.* <https://doi.org/10.1259/bjr.20230203> (2023).
14. Alhaidry, H., Fatani, B., Alrayes, J., Almana, A. & Alfhaed, N. ChatGPT in dentistry: a comprehensive review. *Cureus* <https://doi.org/10.7759/cureus.38317> (2023).
15. Eggmann, F., Weiger, R., Zitzmann, N. U. & Blatz, M. B. Implications of large language models such as ChatGPT for dental medicine. *J. Esthet. Restor. Dent.* <https://doi.org/10.1111/jerd.13046> (2023).
16. Fatani, B. ChatGPT for future medical and dental research. *Cureus* **15**, e37285 (2023).
17. Damashek, M. Gauging similarity with n-grams: language-independent categorization of text. *Science* **267**, 843–848 (1995).
18. Eichstaedt, J. C. et al. Facebook language predicts depression in medical records. *Proc. Natl Acad. Sci.* **115**, 11203–11208 (2018).
19. Marafino, B. J., Davies, J. M., Bardach, N. S., Dean, M. L. & Dudley, R. A. N-gram support vector machines for scalable procedure and diagnosis classification, with applications to clinical free text data from the intensive care unit. *J. Am. Med. Inf. Assoc.* **21**, 871–875 (2014).
20. Romanov, A. & Shivade, C. Lessons from natural language inference in the clinical domain. Preprint at arXiv:1808.06752 (2018).
21. Choi, E., Xiao, C., Stewart, W. & Sun, J. MiME: multilevel medical embedding of electronic health records for predictive healthcare. *Adv. Neural Inf. Process. Syst.* **31**, 19 (2018).
22. Sarzynska-Wawer, J. et al. Detecting formal thought disorder by deep contextualized word representations. *Psychiatry Res.* **304**, 114135 (2021).
23. Peng, Y., Yan, S. & Lu, Z. Transfer learning in biomedical natural language processing: an evaluation of BERT and ELMo on ten benchmarking datasets. Preprint at arXiv:1906.05474 (2019).
24. Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. Bert: pre-training of deep bidirectional transformers for language understanding. Preprint at arXiv:1810.04805 (2018).
25. Radford, A., Narasimhan, K., Salimans, T. & Sutskever, I. Improving language understanding by generative pre-training. (2018).
26. Michalopoulos, G., Wang, Y., Kaka, H., Chen, H. & Wong, A. Umlsbert: Clinical domain knowledge augmentation of contextual embeddings using the unified medical language system metathesaurus. Preprint at arXiv:2010.10391 (2020).
27. Beltagy, I., Lo, K. & Cohan, A. SciBERT: a pretrained language model for scientific text. Preprint at arXiv:1903.10676 (2019).
28. Rasmay, L., Xiang, Y., Xie, Z., Tao, C. & Zhi, D. Med-BERT: pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction. *NPJ Digit. Med.* **4**, 86 (2021).
29. Liu, W., et al. K-bert: Enabling language representation with knowledge graph. *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 34. No. 03 (2020).
30. Gu, Y. et al. Domain-specific language model pretraining for biomedical natural language processing. *ACM Trans. Comput. Healthc.* **3**, 1–23 (2021).
31. Devaraj, A., Marshall, I., Wallace, B. C. & Li, J. J. Paragraph-level simplification of medical texts. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 4972–4984 (2021).
32. Roy, A., & Pan, S. Incorporating medical knowledge in BERT for clinical relation extraction. In *Proceedings of the 2021 conference on empirical methods in natural language processing*, 5357–5366 (2021).
33. Neumann, M., King, D., Beltagy, I. & Ammar, W. ScispaCy: fast and robust models for biomedical natural language processing. Preprint at arXiv:1902.07669 (2019).
34. Rae, J. W. et al. Scaling language models: methods, analysis & insights from training gopher. Preprint at arXiv:2112.11446 (2021).
35. Raffel, C. et al. Exploring the limits of transfer learning with a unified text-to-text transformer. *J. Mach. Learn. Res.* **21**, 5485–5551 (2020).
36. Ouyang, L. et al. Training language models to follow instructions with human feedback. Preprint at arXiv:2203.02155 (2022).
37. Kung, T. H. et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLoS Digit. Health* **2**, e0000198 (2023).
38. Gilson, A. et al. How does CHATGPT perform on the United States Medical Licensing Examination? The implications of large language models for medical education and knowledge assessment. *JMIR Med. Educ.* **9**, e45312 (2023).
39. Li, J., Li, D., Savarese, S., & Hoi, S. Blip-2: bootstrapping language-image pre-training with frozen image encoders and large language models. Preprint at arXiv:2301.12597 (2023).
40. Saharia, C. et al. Photorealistic text-to-image diffusion models with deep language understanding. Preprint at arXiv:2205.11487 (2022).
41. Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A. & Sutskever, I. Zero-shot text-to-image generation. In *International Conference on Machine Learning*, pp. 8821–8831 (PMLR, 2021).
42. Rombach, R., Blattmann, A., Lorenz, D., Esser, P. & Ommer, B. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 10684–10695 (2022).
43. He, Y., Zhu, Z., Zhang, Y., Chen, Q. & Caverlee, J. Infusing disease knowledge into BERT for health question answering, medical inference and disease name recognition. Preprint at arXiv:2010.03746 (2020).
44. Khare, Y., Bagal, V., Mathew, M., Devi, A., Priyakumar, U. D. & Jawahar, C. V. Mmbert: Multimodal bert pretraining for improved medical vqa. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, 1033–1036 (IEEE, 2021).
45. Manco, I., Benetos, E., Quinton, E. & Fazekas, G. Contrastive audio-language learning for music. Preprint at arXiv:2208.12208 (2022).
46. Li, H., Kang, Y., Liu, T., Ding, W. & Liu, Z. CTAL: Pre-training cross-modal transformer for audio-and-language representations. Preprint at arXiv:2109.00181 (2021).
47. Wu, F. et al. Wav2Seq: Pre-training speech-to-text encoder-decoder models using pseudo languages, *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1–5, Rhodes Island, Greece, 2023, <https://doi.org/10.1109/ICASSP49357.2023.10096988> (2022).
48. Gurram, S., Chan, D., Fang, A., & Canny, J. LAVA: Language Audio Vision Alignment for Data-Efficient Video Pre-Training. In *First Workshop on Pre-training: Perspectives, Pitfalls, and Paths Forward at ICML* (2023).
49. Radford, Alec, et al. Robust speech recognition via large-scale weak supervision. *International Conference on Machine Learning* (PMLR, 2023).
50. Huh, J., Park, S., Lee, J. E. & Ye, J. C. Improving medical speech-to-text accuracy with vision-language pre-training model. Preprint at arXiv:2303.00091 (2023).
51. OpenAI. GPT-4 Technical Report. Preprint at arXiv:2303.08774 (2023).
52. Baumgartner, C. The potential impact of ChatGPT in clinical and translational medicine. *Clin. Transl. Med.* **13**, e1206 (2023).
53. Wang, D., Zhang, S. & Wang, L. Deep epidemiological modeling by black-box knowledge distillation: an accurate deep learning model for COVID-19. *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 35. No. 17 (2021).
54. Wang, D., Gong, B. & Wang, L. On calibrating semantic segmentation models: analyses and an algorithm. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2023).
55. Warin, K. et al. Maxillofacial fracture detection and classification in computed tomography images using convolutional neural network-based models. *Sci. Rep.* **13**, 3434 (2023).
56. Bui, R. et al. Artificial intelligence as a decision-making tool in forensic dentistry: a pilot study with I3M. *Int. J. Environ. Res. Public Health* **20**, 4620 (2023).
57. Cui, Z. et al. A fully automatic AI system for tooth and alveolar bone segmentation from cone-beam CT images. *Nat. Commun.* **13**, 2096 (2022).
58. Lee, J. et al. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics* **36**, 1234–1240 (2020).
59. Alsentzer, E. et al. Publicly available clinical BERT embeddings. Preprint at arXiv:1904.03323 (2019).
60. Huang, K., Altosaar, J. & Ranganath, R. Clinicalbert: modeling clinical notes and predicting hospital readmission. Preprint at arXiv:1904.05342 (2019).
61. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
62. Ahdrizt, G. et al. OpenFold: retraining AlphaFold2 yields new insights into its learning mechanisms and capacity for generalization. Preprint at *bioRxiv* <https://doi.org/10.1101/2022.11.20.517210> (2022).
63. Carbajosa, G., Trigo, A., Valencia, A. & Cases, I. Bionemo: molecular information on biodegradation metabolism. *Nucleic Acids Res.* **37**, D598–D602 (2009).
64. Sevgen, E. et al. ProT-VAE: protein transformer variational autoencoder for functional protein design. Preprint at *bioRxiv* <https://doi.org/10.1101/2023.01.23.525232> (2023).
65. Jensen, P. B., Jensen, L. J. & Brunak, S. Mining electronic health records: towards better research applications and clinical care. *Nat. Rev. Genet.* **13**, 395–405 (2012).
66. Kocbek, S. et al. Text mining electronic hospital records to automatically classify admissions against disease: measuring the impact of linking data sources. *J. Biomed. Inform.* **64**, 158–167 (2016).

67. Sun, W. et al. Data processing and text mining technologies on electronic medical records: a review. *J. Healthc. Eng.* **2018**, 4302425 (2018).
68. Rajkomar, A. et al. Scalable and accurate deep learning with electronic health records. *NPJ Digit. Med.* **1**, 18 (2018).
69. Huber, M. T., Highland, J. D., Krishnamoorthi, V. R. & Tang, J. W.-Y. Utilizing the electronic health record to improve advance care planning: a systematic review. *Am. J. Hosp. Palliat. Med.* **35**, 532–541 (2018).
70. Luo, Y. et al. Natural language processing for EHR-based pharmacovigilance: a structured review. *Drug Saf.* **40**, 1075–1089 (2017).
71. Hirschberg, J. & Manning, C. D. Advances in natural language processing. *Science* **349**, 261–266 (2015).
72. Natsiavas, P., Maglaveras, N., & Koutkias, V. A public health surveillance platform exploiting free-text sources via natural language processing and linked data: application in adverse drug reaction signal detection using PubMed and Twitter. In *Knowledge Representation for Health Care: HEC 2016 International Joint Workshop, KR4HC/ProHealth 2016*, 51–67, Munich, Germany (Springer International Publishing, 2017).
73. Geng, W. et al. Model-based reasoning of clinical diagnosis in integrative medicine: real-world methodological study of electronic medical records and natural language processing methods. *JMIR Med. Inform.* **8**, e23082 (2020).
74. Lee, S. H. Natural language generation for electronic health records. *NPJ Digit. Med.* **1**, 63 (2018).
75. Hoogi, A., Mishra, A., Gimenez, F., Dong, J. & Rubin, D. Natural language generation model for mammography reports simulation. *IEEE J. Biomed. Health Inform.* **24**, 2711–2717 (2020).
76. Wongratwanich, P. et al. Do various imaging modalities provide potential early detection and diagnosis of medication-related osteonecrosis of the jaw? A review. *Dentomaxillofac. Radiol.* **50**, 20200417 (2021).
77. Alabi, R. O. et al. Machine learning in oral squamous cell carcinoma: current status, clinical concerns and prospects for future—a systematic review. *Artif. Intell. Med.* **115**, 102060 (2021).
78. Jha, N., Lee, K. S. & Kim, Y. J. Diagnosis of temporomandibular disorders using artificial intelligence technologies: a systematic review and meta-analysis. *PLoS ONE* **17**, e0272715 (2022).
79. Monill-González, A., Rovira-Calatayud, L., d'Oliveira, N. G. & Ustrell-Torrent, J. M. Artificial intelligence in orthodontics: where are we now? a scoping review. *Orthod. Craniofac. Res.* **24**, 6–15 (2021).
80. Revilla-León, M. et al. Artificial intelligence applications in restorative dentistry: a systematic review. *J. Prosthet. Dent.* **128**, 867–875 (2022).
81. Schwendicke, F. A., Samek, W. & Krois, J. Artificial intelligence in dentistry: chances and challenges. *J. Dent. Res.* **99**, 769–774 (2020).
82. Amer, Y. Y. & Aqel, M. J. An efficient segmentation algorithm for panoramic dental images. *Procedia Comput. Sci.* **65**, 718–725 (2015).
83. Shi, B. & Huang, H. Computational technology for nasal cartilage-related clinical research and application. *Int. J. Oral. Sci.* **12**, 21 (2020).
84. Huang, H., Cheng, X., Luo, X., Shi, B. & Li, J. Biomechanical analyses of common suspension sutures in primary cleft lip rhinoplasty. *Head Face Med.* **15**, 20 (2019).
85. Huang, H. et al. Mechanical analyses of critical surgical maneuvers in the correction of cleft lip nasal deformity. *PLoS ONE* **13**, e0195583 (2018).
86. Huang, H., Luo, X., Cheng, X., Shi, B. & Li, J. Biomechanical simulation of correcting primary unilateral cleft lip nasal deformity. *PLoS ONE* **13**, e0199964 (2018).
87. Huang, H. et al. Recapitulation of unilateral cleft lip nasal deformity on normal nasal structure: a finite element model analysis. *J. Craniofac. Surg.* **29**(8), 2220–2225 (2018).
88. Sakran, K. A. et al. Early cleft palate repair by a modified technique without relaxing incisions. *Cleft Palate Craniofac. J.* <https://doi.org/10.1177/10556656221135288> (2022).
89. Sakran, K. A. et al. Evaluation of late cleft palate repair by a modified technique without relaxing incisions. *J. Stomatol. Oral Maxillofac. Surg.* **124**, 101403 (2023).
90. Huang, H. et al. Validation of the Chinese velopharyngeal insufficiency effects on life outcomes instrument. *Laryngoscope* **129**, E395–E401 (2019).
91. Huang, H. et al. Analysis of velopharyngeal functions using computational fluid dynamics simulations. *Ann. Otol. Rhinol. Laryngol.* **128**, 742–748 (2019).
92. Huang, H. et al. Computational fluid dynamic analysis of different velopharyngeal closure patterns. *Ann. Otol. Rhinol. Laryngol.* **129**, 157–163 (2019).
93. Huang, H. et al. Airflow of the two-port velopharyngeal closure: study using computational fluid dynamics. *J. Craniofac. Surg.* **31**, 2188–2192 (2020).
94. Yang, C. et al. Inspiration after posterior pharyngeal flap palatoplasty: a preliminary study using computational fluid dynamic analysis. *Front. Pediatr.* **10**, 823777 (2022).
95. Ström, P. et al. Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. *Lancet Oncol.* **21**, 222–232 (2020).
96. Girolami, I. et al. Artificial intelligence applications for pre-implantation kidney biopsy pathology practice: a systematic review. *J. Nephrol.* **35**, 1801–1808 (2022).
97. Wan, A. & Savage, N. Biopsy and diagnostic histopathology in dental practice in Brisbane: usage patterns and perceptions of usefulness. *Aust. Dent. J.* **55**, 162–169 (2010).
98. Ilhan, B., Lin, K., Guneri, P. & Wilder-Smith, P. Improving oral cancer outcomes with imaging and artificial intelligence. *J. Dent. Res.* **99**, 241–248 (2020).
99. Miller, C. S. & Westgate, P. M. Implications of medical screenings of patients arriving for dental treatment: the results of a comprehensive laboratory screening. *J. Am. Dent. Assoc.* **145**, 1027–1035 (2014).
100. Yu, X. et al. GPT paternity test: GPT generated text detection with GPT genetic inheritance. Preprint at arXiv:2305.12519 (2023).
101. Zhang, N. et al. Ontoprotein: protein pretraining with gene ontology embedding. Preprint at arXiv:2201.11147 (2022).
102. Bubeck, S. et al. Sparks of artificial general intelligence: early experiments with GPT-4. Preprint at arXiv:2303.12712 (2023).
103. Caufield, J. H. et al. Structured prompt interrogation and recursive extraction of semantics (SPIRES): a method for populating knowledge bases using zero-shot learning. Preprint at arXiv:2304.02711 (2023).
104. Sakran, K. A. et al. Evaluation of postoperative outcomes in two cleft palate repair techniques without relaxing incisions. *Plast. Reconstr. Surg.* **152**, 145–154 (2023).
105. McNichols, H., Zhang, M. & Lan, A. Algebra error classification with large language models. Preprint at arXiv:2305.06163 (2023).
106. Lamb, L. C. et al. Graph neural networks meet neural-symbolic computing: a survey and perspective. Preprint at arXiv:2003.00330 (2020).
107. Al Omar, A., Bhuiyan, M. Z. A., Basu, A., Kiyomoto, S. & Rahman, M. S. Privacy-friendly platform for healthcare data in cloud based on blockchain environment. *Future Gener. Comput. Syst.* **95**, 511–521 (2019).
108. Sharma, S. et al. From occlusion to insight: object search in semantic shelves using large language models. Preprint at arXiv:2302.12915 (2023).



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023