

## RESEARCH ARTICLE

# Artificial intelligence for detecting temporomandibular joint osteoarthritis using radiographic image data: A systematic review and meta-analysis of diagnostic test accuracy

Liang Xu<sup>1,2</sup>, Jiang Chen<sup>1,3\*</sup>, Kaixi Qiu<sup>4</sup>, Feng Yang<sup>3</sup>, Weiliang Wu<sup>1</sup>

**1** The School of Stomatology, Fujian Medical University, Fuzhou, Fujian, China, **2** Department of Stomatology, The First Affiliated Hospital of Fujian Medical University, Fuzhou, Fujian, China, **3** School and Hospital of Stomatology, Fujian Medical University, Fuzhou, Fujian, China, **4** Fuzhou No. 1 Hospital Affiliated with Fujian Medical University, Fuzhou, Fujian, China

 These authors contributed equally to this work.

\* [jiangchen@fjmu.edu.cn](mailto:jiangchen@fjmu.edu.cn)



## OPEN ACCESS

**Citation:** Xu L, Chen J, Qiu K, Yang F, Wu W (2023) Artificial intelligence for detecting temporomandibular joint osteoarthritis using radiographic image data: A systematic review and meta-analysis of diagnostic test accuracy. PLoS ONE 18(7): e0288631. <https://doi.org/10.1371/journal.pone.0288631>

**Editor:** Jonas Bianchi, University of the Pacific Arthur A Dugoni School of Dentistry, UNITED STATES

**Received:** March 20, 2023

**Accepted:** July 2, 2023

**Published:** July 14, 2023

**Copyright:** © 2023 Xu et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its [Supporting Information](#) files.

**Funding:** This work was supported by Fujian Provincial Department of Education (JAT200161). The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

## Abstract

In this review, we assessed the diagnostic efficiency of artificial intelligence (AI) models in detecting temporomandibular joint osteoarthritis (TMJOA) using radiographic imaging data. Based upon the PRISMA guidelines, a systematic review of studies published between January 2010 and January 2023 was conducted using PubMed, Web of Science, Scopus, and Embase. Articles on the accuracy of AI to detect TMJOA or degenerative changes by radiographic imaging were selected. The characteristics and diagnostic information of each article were extracted. The quality of studies was assessed by the QUADAS-2 tool. Pooled data for sensitivity, specificity, and summary receiver operating characteristic curve (SROC) were calculated. Of 513 records identified through a database search, six met the inclusion criteria and were collected. The pooled sensitivity, specificity, and area under the curve (AUC) were 80%, 90%, and 92%, respectively. Substantial heterogeneity between AI models mainly arose from imaging modality, ethnicity, sex, techniques of AI, and sample size. This article confirmed AI models have enormous potential for diagnosing TMJOA automatically through radiographic imaging. Therefore, AI models appear to have enormous potential to diagnose TMJOA automatically using radiographic images. However, further studies are needed to evaluate AI more thoroughly.

## Introduction

Temporomandibular joint osteoarthritis (TMJOA), a severe subtype of temporomandibular joint disorder, is characterized by progressive absorption of articular cartilage, remodeling of the subchondral bone, and chronic pain [1]. The global incidence of TMJOA is reported to be 8% to 16% [2–4]. The disease severely affects patients' quality of life, causing excruciating pain and imposing a heavy social and economic burden on individuals and families [5]. Therefore,

**Competing interests:** The authors have declared that no competing interests exist.

the importance of early TMJOA diagnosis lies in its potential to enhance treatment efficacy, alleviate symptoms, implement preventive measures, preserve joint function, and optimize healthcare resource utilization. Timely identification and intervention can significantly improve patient outcomes and quality of life while potentially reducing the burden of more invasive and costly treatments.

Presently, the diagnosis of osteoarthritis mainly depends on medical history, disease characteristics, and digital imaging. However, the first two methods sometimes provide limited information about the joint status of patients with TMJOA. Therefore, medical imaging, such as magnetic resonance imaging (MRI), cone-beam computed tomography (CBCT), and orthopantomogram (OPG), is often necessary to assess osteoarthritis [6–8]. The X-ray passed through the temporomandibular joint area and was subsequently detected by a detector. The attenuated X-ray signal is converted into an electrical signal, which is then converted into computed tomography (CT) and OPG images. The principle of nuclear magnetic resonance is what MRI uses. Radio frequency excites nuclei within an external magnetic field to produce a signal and convert it into a medical image. The clinician can detect the morphological changes of bone components in the TMJ with OPG, MRI, and CBCT images.

The above-mentioned noninvasive imaging modalities have been widely used in diagnosing TMJOA [9–11]. Studies have indicated that OPG can only detect apparent erosion, sclerosis, and osteophytes [12]. Although OPG is not sensitive enough to TMJOA [13], it is still an effective tool for the preliminary screening of TMJ. MRI and CBCT were selected to carefully evaluate temporomandibular joint structure because they can identify osseous changes such as erosion, osteophytes, and sclerosis [14]. CBCT is the preferred method of bone evaluation compared to MRI and performs better in diagnosing TMJ osteoarthritis [15]. MRI is also frequently used in TMJOA studies because of advantages in bone evaluation [16]. It not only assesses the condyle morphology but also identifies abnormalities in the bone marrow. Nevertheless, even with such rapid advances in radiological imaging, accurate diagnosis of TMJOA remains challenging. The diagnosis of TMJOA is particularly complex because of the diverse morphological changes of the condyle, such as erosion, flattening, sclerosis, osteophyte, and subcortical cyst [17]. As the diagnostic accuracy of degenerative changes in the condyle depends largely on the radiologists' expertise and quality of the scanner itself, the interpretation process is susceptible to both misdiagnosis and missed diagnosis [18]. Studies have shown that CBCT has higher accuracy in diagnosing degenerative bone changes than MRI and OPG, with a sensitivity of 0.7–0.9 [19].

Additionally, improvements in technology to extract more accurate and meaningful diagnostic information have made physicians' work more complex. Therefore, rapid and accurate diagnosis of TMJOA has gradually become a research hotspot.

AI is becoming increasingly popular, especially deep learning (DL), owing to remarkable data mining and processing progress. It is considered a reliable method for combining clinical data and physician reports from electronic medical records to improve the accuracy of various medical tasks [20]. AI has been applied to segment or diagnose lesions automatically in medical images and can be combined with the results of other medical tests to determine disease prognosis [21–23]. In this review, we aimed to assess the diagnostic efficiency of AI models in detecting TMJOA using radiographic image data.

## Materials and methods

Our study protocol was registered on PROSPERO (CRD42023396713). This systematic review and meta-analysis were completed following the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) guidelines, including search strategy, eligibility criteria,

data extraction, risk of bias assessment, and data analysis ([S1 Checklist](#)). In addition, the QUADAS-2 tool [24] was used to assess the quality of the included studies.

### Search strategy

Two authors (QKX and WWL) separately performed PubMed, Web of Science, Scopus, and Embase database search using standard search formulas. A comprehensive search was conducted across four databases using a combination of predetermined keywords. The keywords used for the search were ("artificial intelligence" [Mesh] OR "machine learning" [Title/Abstract] OR "neural networks, computer" [Title/Abstract] OR "deep learning" [Title/Abstract]) AND ("temporomandibular joint"[Title/Abstract] OR "osteoarthritis"[Title/Abstract]) AND "sensitivity and specificity"[Title/Abstract]. Additional information regarding the search strategy can be found in the [S1 File](#). Additional articles were identified by manual search. Then, a full-text review was conducted to determine whether the identified literature met the inclusion criteria. If there was any disagreement during the search process, it was settled via consultation with the third author (YF).

### Eligibility criteria

The literature published in PubMed, Web of Science, Scopus, and Embase between January 2010 and January 2023 on the detection of TMJOA using AI in radiographic images was included. There were no restrictions on the countries where the studies were conducted, but only articles published in English were included. The inclusion criteria according to PICOS was as follows: P, patient with TMJOA or degenerative change of condyle; I, AI models including deep learning, machine learning, and radiomic; C, not applicable; O, sensitivity, specificity, AUC value; S, prospective or retrospective study. Excluded criteria were narrative reviews, letters, reviews, editorials, protocol studies, guides, systematic reviews, and meta-analyses.

### Data extraction

The following data were extracted from the included literature: types of AI models, research characteristics, and outcome measurements. To obtain diagnostic accuracy data, a 2×2 confusion matrix, sensitivity, specificity, accuracy, true positive, false positive, true negative, false negative, and the area under the receiver operating characteristic curve (AUROC) were extracted or reconstructed. For further analysis, the following data were extracted: authors, publication year, country, sex, study style, imaging modality, total images, the sample size of test data, and techniques.

### Quality assessment and publication bias

Two independent reviewers (QKX and YF) performed quality assessments of selected studies using the QUADAS-2 criteria. When there was a disagreement, the third author made the final decision based on the criteria. Publication bias was assessed by funnel plot of diagnostic AUC. Asymmetric shape of the funnel plot of included studies indicated study heterogeneity.

### Data analysis

Sensitivity and specificity were calculated by true positives, false positives, true negatives, and false negatives. Forest plots of sensitivity and specificity and a summary receiver operating characteristic (SROC) curve were generated using Stata 15.1. Meta-regression analysis was conducted to estimate the source of heterogeneity when  $I^2$  was  $\geq 50\%$ .

## Results

### Search results

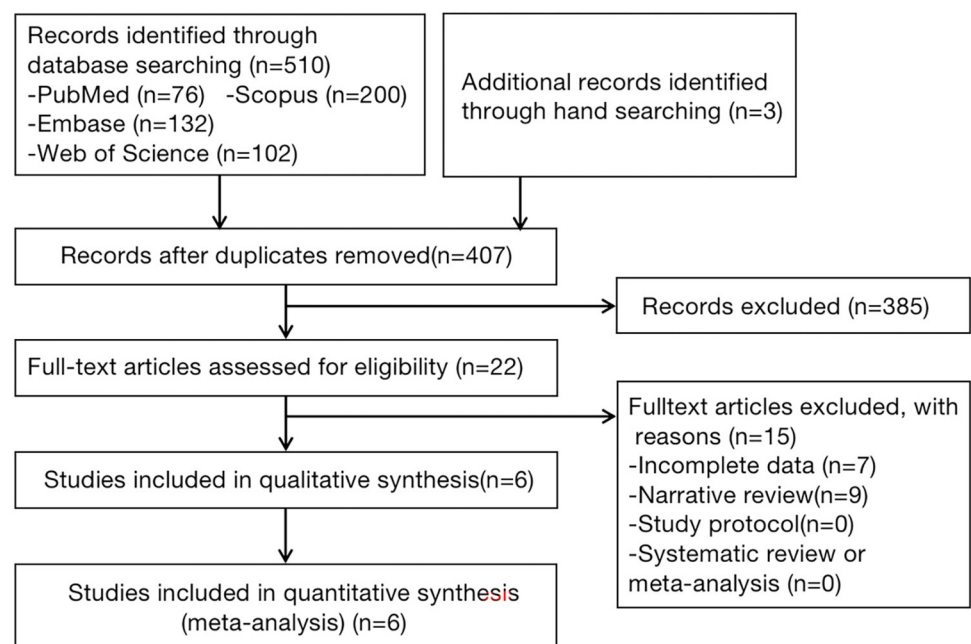
We retrieved 510 articles from the four databases. Three additional articles were identified through manual screening. After removing 103 duplicates, we analyzed the titles and abstracts of the remaining 407 articles. Twenty-two articles of interest were identified according to the inclusion and exclusion criteria. Sixteen of these articles were subsequently excluded due to incomplete data or a narrative review style. Finally, six studies [25–30] were included in the systematic review. The literature screening process is shown in Fig 1.

### Study characteristics

All six included studies were retrospective (Table 1). MRI was used in one study, CBCT in two studies, and OPG in three studies. Five studies used the Diagnostic Criteria for Temporomandibular Disorders published by Schiffman as the reference standard [31], whereas one study did not offer a reference standard. The research was completed in Belgium, Iran, and Korea, with 66.6% (4/6) of the studies conducted in South Korea. The selected studies utilized one of two AIs: KNN (2/6, 33.3%) or convolutional neural networks (CNN) (4/6, 66.7%).

### Quality assessment and publication bias

To evaluate the quality of the studies, we applied the QUADAS-2 risk checklist to test the bias risk in each study (Fig 2). The risk of bias in patient selection was low in one half (3/6, 50%) of the studies and unclear in the other half. The same was true for the risk of bias in the index test, flow, and timing. The reference standard test included all six studies with a low risk of bias (100%). Applicability concerns regarding patient selection were low in five (83.3%) studies and unclear in one (16.7%). Applicability concerns in the index test were low in one study (16.7%), high in one study (16.7%), and unclear in four studies (66.6%). Applicability concerns in the reference standard were low in four (66.7%) and high in two (33.3%) studies. The funnel plot assessment (Fig 3) showed no significant publication bias ( $P = 0.68$ ).



**Fig 1. PRISMA flowchart of the included articles.**

<https://doi.org/10.1371/journal.pone.0288631.g001>

**Table 1. Characteristics of individual studies.**

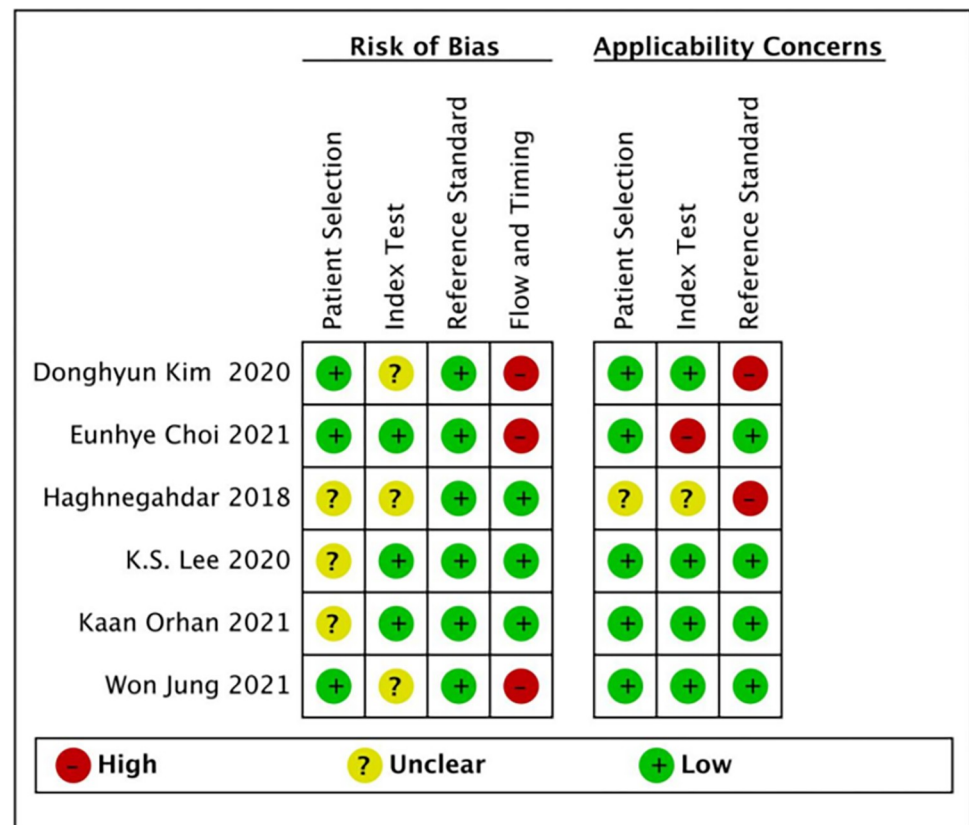
| author         | year | country | study style         | total images | Sample size of test data | age and gender                                       | AI Techniques                   | Female in sample (%) |
|----------------|------|---------|---------------------|--------------|--------------------------|--|---------------------------------|----------------------|
| Eunhye Choi    | 2021 | Korea   | retrospective study | 1599         | 272                      | 210males, 979females; mean ± SD age, 37.1 ± 16.0y;   | Keras ResNet (CNN)              | 82.3                 |
| K.S. Lee       | 2020 | Korea   | retrospective study | 3514         | 300                      | 84 males, 230 females; mean ± SD age, 39.5 ± 18.2y;  | single-shot detector (SSD)(CNN) | 73.2                 |
| Won Jung       | 2021 | Korea   | retrospective study | 858          | 172                      | 142 males, 376 females; mean ± SD age, 47.3 ± 20.1y; | EfficientNet-B7 (CNN)           | 72.6                 |
| Donghyun Kim   | 2020 | Korea   | retrospective study | 1932         | 231                      | 700 male and 592 female; mean age: 43.3 years        | Fine-tuned VGG16 network(CNN)   | 45.8                 |
| Kaan Orhan     | 2021 | Belgium | retrospective study | 856          | 18                       | 34 male and 73 female; mean age: 38 years ± 17.97;   | k-nearest neighbors (KNN)       | 68.2                 |
| Haghnegahdar A | 2018 | Iran    | retrospective study | 264          | 264                      | unclear  | k-nearest neighbors (KNN)       | unclear              |

MRI, magnetic resonance imaging; CBCT, cone-beam computed tomography; OPG, orthopantomogram; TMJOA, temporomandibular joint osteoarthritis; CNN, convoluted neural networks; KNN, k-nearest neighbor

<https://doi.org/10.1371/journal.pone.0288631.t001>

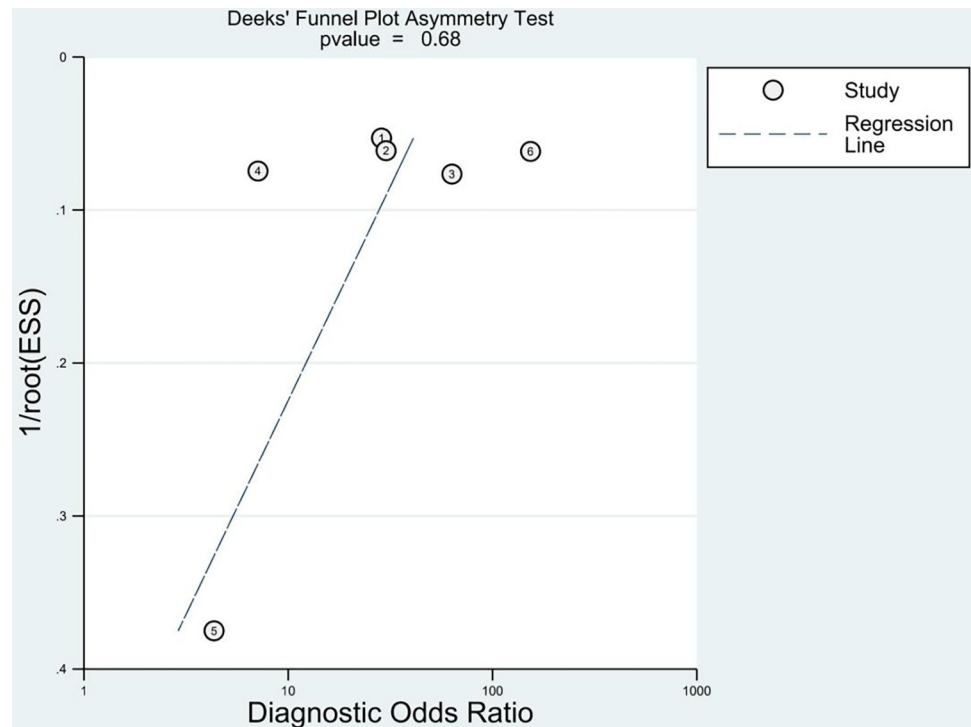
### Diagnostic accuracy

The included studys’ sensitivity and specificity ranged from 0.54 to 0.94 and from 0.50 to 0.91, respectively. The pooled sensitivity and specificity for AI models were 0.80 (95% confidence interval [CI]: 0.67–0.89) with severe heterogeneity (89%) and 0.90 (95% CI: 0.87–0.92) with



**Fig 2. Quality assessment by QUADAS-2 tool.**

<https://doi.org/10.1371/journal.pone.0288631.g002>



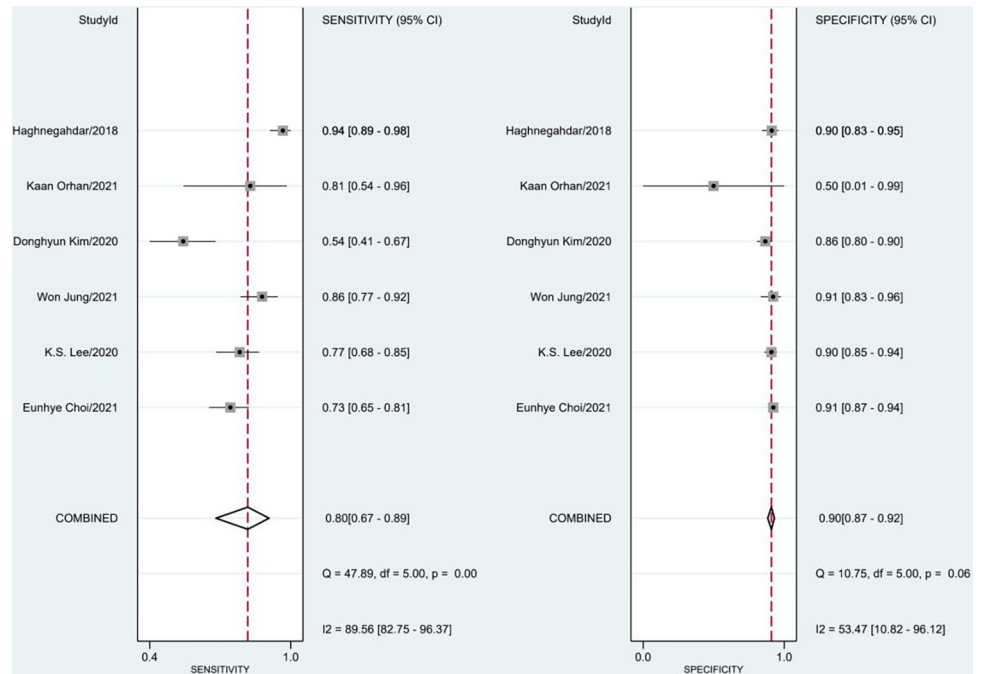
**Fig 3. Funnel plot for diagnostic accuracy of AI in detection of TMJOA.** TMJOA, temporomandibular joint osteoarthritis; AI, artificial intelligence.

<https://doi.org/10.1371/journal.pone.0288631.g003>

moderate heterogeneity (53%) (Fig 4). According to the SROC curve, the AUC was 0.92 (95% CI: 0.89–0.94) (Fig 5). We performed a meta-regression analysis to explore the sources of heterogeneity and found that imaging modality, ethnicity, sex, age, AI techniques, and sample size were all possible causes of heterogeneity (Table 2).

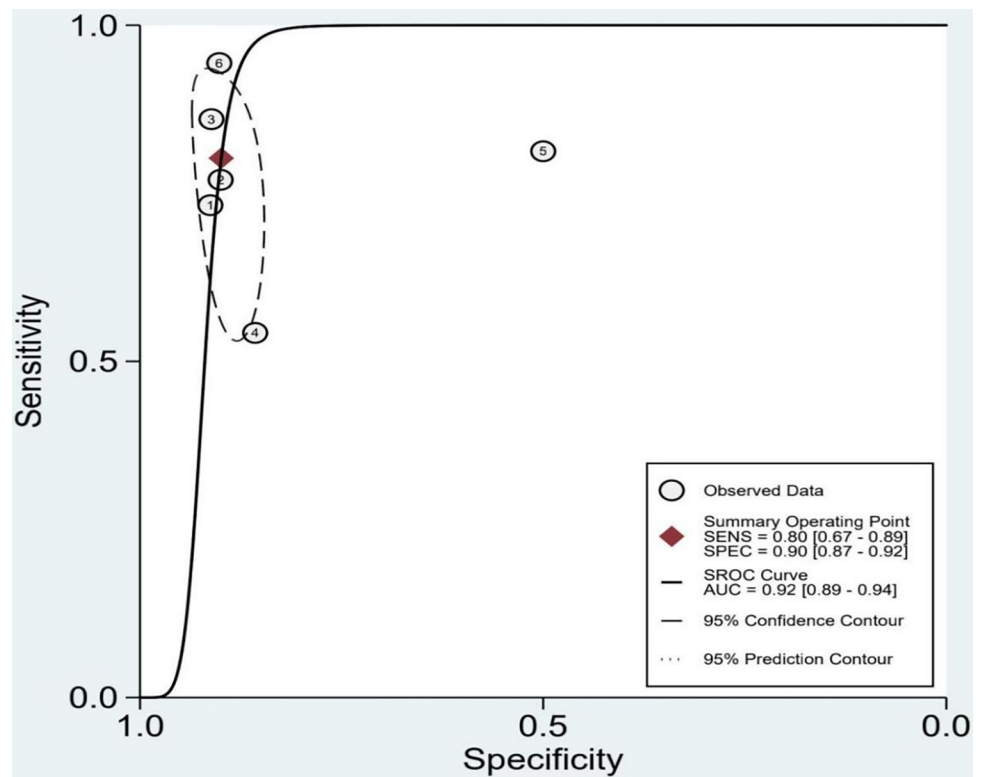
## Discussion

AI has been used to study knee and hip arthritis, mainly for autodetection, classification, and segmentation. Artificial intelligence (AI) has shown promising applications in the diagnosis of temporomandibular arthritis (TMJ arthritis). AI algorithms have been developed to analyze medical imaging data, such as X-rays, CT, and MRI, to aid in the detection and diagnosis of TMJ arthritis. These algorithms can assist in identifying characteristic features, measuring joint space, assessing bone changes, and detecting early signs of arthritis. By training models on large datasets, AI algorithms can learn to distinguish healthy TMJ joints from those affected by arthritis, enabling more accurate and efficient diagnosis. It is important to note that while AI shows promise in TMJ arthritis diagnosis, its clinical application is still evolving. Further research, validation, and refinement of AI models are needed to ensure their accuracy, reliability, and integration into routine clinical practice. To our knowledge, only a few studies and meta-analyses for the auto-detection of TMJOA have been conducted on temporomandibular osteoarthritis. de Dumast et al. constructed a neural network algorithm to classify TMJOA based on an imaging dataset [32]. Ribera et al. also designed a similar deep learning model [33]. However, the diagnostic accuracy differed appreciably between these studies, ranging from 78.0% to 92.4%. We have retrieved two meta-analyses that bear similarities to our current article. The first study evaluated TMD [34], while the second investigated TMJOA [35]. It is



**Fig 4. Meta-analysis of sensitivity and specificity for AI.**

<https://doi.org/10.1371/journal.pone.0288631.g004>



**Fig 5. SROC for diagnostic accuracy of AI in detection of TMJOA.**

<https://doi.org/10.1371/journal.pone.0288631.g005>

Table 2. Meta-regression analysis for diagnostic accuracy of AI in detection of TMJOA.

| Covariate/Subgroup     | Studies(n) | Sensitivity (95%CI) | P-value | Specificity (95%CI) | P-value |
|------------------------|------------|---------------------|---------|---------------------|---------|
| Image modality         |            |                     | 0.86    |                     | <0.001  |
| OPG                    | 3          | 0.73 (0.58–0.88)    |         | 0.90 (0.87–0.93)    |         |
| The rest               | 3          | 0.87 (0.77–0.97)    |         | 0.90 (0.86–0.94)    |         |
| Number of total images |            |                     | <0.001  |                     | <0.001  |
| >1000                  | 3          | 0.69 (0.59–0.80)    |         | 0.89 (0.86–0.92)    |         |
| ≤1000                  | 3          | 0.90 (0.84–0.96)    |         | 0.90 (0.85–0.94)    |         |
| Number of test data    |            |                     |         |                     |         |
| >200                   | 4          | 0.78 (0.65–0.91)    | <0.001  | 0.90 (0.87–0.92)    | <0.001  |
| ≤200                   | 2          | 0.85 (0.69–1.00)    |         | 0.90 (0.83–0.97)    |         |
| Type of AI techniques  |            |                     |         |                     |         |
| CNN                    | 4          | 0.74 (0.63–0.85)    | <0.001  | 0.90 (0.87–0.93)    | <0.001  |
| KNN                    | 2          | 0.91 (0.84–0.99)    |         | 0.88 (0.81–0.96)    |         |
| Mean age               |            |                     | 0.2     |                     | <0.001  |
| >40                    | 2          | 0.73 (0.58–0.88)    |         | 0.88 (0.84–0.93)    |         |
| ≤40                    | 3          | 0.76 (0.64–0.88)    |         | 0.91 (0.88–0.94)    |         |
| Female in sample (%)   |            |                     | <0.001  |                     | <0.001  |
| >65                    | 4          | 0.79(0.73–0.84)     |         | 0.91(0.88–0.93)     |         |
| ≤65                    | 1          | 0.54(0.73–0.84)     |         | 0.86(0.81–0.91)     |         |

<https://doi.org/10.1371/journal.pone.0288631.t002>

noteworthy that both studies included additional diagnostic biomarkers beyond medical images to establish the diagnosis of TMD and TMJOA. This limitation reduced the scientific value of these two meta-analyses. The present systematic review and meta-analysis aimed to explore the diagnostic rate of AI models developed for the detection of TMJOA on medical images. Six articles were included in our research, comprising 523 images with osteoarthritis and 734 images from controls. After the synthesis, the pooled specificity, sensitivity, and accuracy were 0.80, 0.90, and 0.92, respectively. This high-pooled DTA shows that AI models can successfully differentiate between patients with and without degenerative changes. Researchers have confirmed a similar conclusion in other fields that artificial intelligence is more accurate and reliable than radiologists [36].

The sensitivity of DL in Kim's literature was only 0.54. We carefully reviewed the literature and found that although model 1 and model 2 could identify condyle region and morphology, respectively, their detection efficiency was relatively low. Adjusting the hyperparameters of the model may result in the desired performance. In Orhans's literature, the low specificity may be due to insufficient samples in the test set, which is much lower than in other articles. The MRI images in this article were used to analyze osseous changes and disc displacement. Therefore, the samples were selected from patients with temporomandibular joint disorder. The test set's samples were insufficient because the proportion of TMJOA patients in the total sample was low. However, we agree with the authors on the design of the deep learning model. Therefore, these two articles were eventually included in our meta-analysis.

In our analysis, heterogeneity was observed among the included studies. Therefore, we performed a meta-regression analysis to identify the source of heterogeneity and revealed that imaging modality, ethnicity, sex, age, AI techniques, and sample size might account for heterogeneity. OPG (50.0%), CBCT (33.3%), and MRI (16.7%) were selected as the study images in the included literature. Imaging modalities may be the primary source of heterogeneity. It is generally believed that CBCT is superior to other imaging modalities for diagnosing TMJOA [16]. In our meta-analysis, the included articles based on OPG were less accurate (78%–88%)



than those based on CBCT (86%–92%). This is similar to the conclusion from Kaimal S [37], who compared the diagnostic effects of OPG, MRI, and CBCT for TMJOA. Their results showed that the inter-observer reliabilities based on OPG, MRI, and CBCT had a Kappa of 0.16, 0.47, and 0.71, respectively. OPG performance can be poor because the condyles may overlap with the articular eminence or cervical vertebrae in OPG images. In the process of deep learning, the pixels of the overlapping region increase, thus affecting the calculation results. However, CBCT does not have a similar situation because of its three-dimensional nature. Since the included literature using MRI was insufficient, this paper does not discuss it here. The diagnostic efficiency of MRI was inferior to that of CBCT, which was similar to the results of previous studies [16]. The short relaxation time of hard tissue resulted in poor MRI performance [38]. However, many articles still use MRI to study TMJOA [39, 40]. The advantage of MRI is that it can evaluate hard and soft tissues simultaneously to analyze TMJ more comprehensively.

Sample sizes ( $P < 0.001$ ), including the size of the test data and techniques used to train deep learning ( $P < 0.001$ ), were also identified as sources of heterogeneity. In general, the performance of deep learning models improves as the amount of data increases [41, 42]. Several experiments have used training sets of different sizes to train deep learning models to determine the ideal sample size [43, 44]. Their results revealed that insufficient sample sizes lead to undertraining, which can affect the final accuracy. A similar situation was observed in this study. The accuracy fluctuated greatly (78%–92%) in the articles with small total sample sizes, while it fluctuated slightly (78%–84%) in the articles with large total sample sizes. The influence of the size of the test data on accuracy was also similar.

Additionally, the performance of deep learning is directly related to the inherent characteristics of deep learning [45, 46]. The included studies constructed two deep learning models: CNN and KNN. The accuracy of CNN was higher than that of KNN, which may be due to their different structures and characteristics. CNN is essentially a mathematical model whose structure and operation logic refer to a biological nervous system, so information can be distributed and processed in parallel [47]. The KNN algorithm, also known as KNN or K-NN, is a supervised learning model for classification [48]. KNN classifies or predicts groupings of individual data points by proximity. Owing to KNN's relatively simple architecture and algorithm of the KNN, small sample sizes and sample imbalance would lead to classification bias.

In contrast, for larger sample sizes, CNN has better classification ability than KNN. Early studies indicate that the average recognition rate of neural networks is higher than that of the KNN classification method [49, 50]. Two studies in our meta-analysis used small samples to train KNN. Combined with the influence of unbalanced samples, the accuracy of KNN may be lower than that of CNN. This suggests that the architecture of deep learning affects performance. Therefore, different deep-learning algorithms should be compared for different application scenarios to determine the most appropriate one.

Although the prevalence of degenerative joint disorder varies widely [51], the incidence of TMJOA is closely correlated with sex and age. Incidence and severity are higher in women [52]. TMJOA is at least twice as common in women than men [53]. Cellular sexual dimorphism, hormones, genetic factors, and immune modulation mechanisms may contribute to the sex disparity observed in TMJOA [54–56]. Numerous studies have confirmed the strong correlation between age and TMJOA [57, 58]. TMJOA has significantly different peak age characteristics, which are approximately 30 and 55 years [59]. This relationship reflects the inherent accumulation of tissue damage owing to a gradual decline in cellular adaptation.

Although the title of our search include artificial intelligence, almost all the literature included in the research used deep learning technology. Deep learning does not require artificial feature sets compared to traditional machine learning. Feature extraction is data-driven,

which enables deep learning to extract deeper features. Meanwhile, higher precision and stronger robustness make deep learning more widely used. Nevertheless, deep learning has its limits. First, deep learning does not negate any data nor detect hidden biases in the data, which can lead to unobjective results. Second, deep learning is susceptible to counterattack, which leads to radically different judgments. Third, it can only find correlations between events but cannot explain causation. Finally, the performance of deep learning depends on the size of the data set, which requires high computational power. Despite these problems, deep learning is still a promising tool.

This meta-analysis only included six papers, which may lead to the need for more objective conclusions. However, this is because the research on deep learning in temporomandibular osteoarthritis is still exploratory, and more literature meeting the inclusion criteria are needed. For this reason, we expanded the search scope to include LILACS, Scopus, Web of Science, and other databases, but no other literature meeting the inclusion criteria was retrieved. We will regularly update this paper in the later stages to ensure its timeliness.

## Conclusion

AI models appear to have enormous potential to diagnose TMJOA automatically using radiographic images. Although AI still has shortcomings in automatic diagnosis and various models differ in accuracy, its high average accuracy still makes it an auxiliary means to avoid misdiagnosis or missed diagnosis. However, further studies are needed to evaluate AI more thoroughly.

## Supporting information

**S1 Checklist. PRISMA checklist.**  
(PDF)

**S1 File. Search strategy.**  
(DOCX)

## Author Contributions

**Conceptualization:** Liang Xu, Jiang Chen.

**Data curation:** Liang Xu, Jiang Chen, Kaixi Qiu, Feng Yang, Weiliang Wu.

**Formal analysis:** Liang Xu, Jiang Chen.

**Investigation:** Jiang Chen, Kaixi Qiu, Feng Yang.

**Methodology:** Kaixi Qiu.

**Project administration:** Liang Xu, Jiang Chen, Weiliang Wu.

**Resources:** Liang Xu, Kaixi Qiu.

**Writing – original draft:** Jiang Chen.

**Writing – review & editing:** Liang Xu, Jiang Chen, Kaixi Qiu.

## References

1. Wang XD, Zhang JN, Gan YH, Zhou YH. Current understanding of pathogenesis and treatment of TMJ osteoarthritis. *J Dent Res*. 2015; 94(5):666–73. Epub 20150305. <https://doi.org/10.1177/0022034515574770> PMID: 25744069.

2. Overman VP. National Institute of Dental and Craniofacial Research. *Int J Dent Hyg*. 2008; 6(3):249. <https://doi.org/10.1111/j.1601-5037.2008.00329.x> PMID: 18768032.
3. Valesan LF, Da-Cas CD, Réus JC, Denardin ACS, Garanhani RR, Bonotto D, et al. Prevalence of temporomandibular joint disorders: a systematic review and meta-analysis. *Clin Oral Investig*. 2021; 25(2):441–53. Epub 20210106. <https://doi.org/10.1007/s00784-020-03710-w> PMID: 33409693.
4. Kalladka M, Quek S, Heir G, Eliav E, Mupparapu M, Viswanath A. Temporomandibular joint osteoarthritis: diagnosis and long-term conservative management: a topic review. *J Indian Prosthodont Soc*. 2014; 14(1):6–15. Epub 20130922. <https://doi.org/10.1007/s13191-013-0321-3> PMID: 24604992; PubMed Central PMCID: PMC3935038.
5. Kim TY, Shin JS, Lee J, Lee YJ, Kim MR, Ahn YJ, et al. Gender Difference in Associations between Chronic Temporomandibular Disorders and General Quality of Life in Koreans: A Cross-Sectional Study. *PLoS One*. 2015; 10(12):e0145002. Epub 20151216. <https://doi.org/10.1371/journal.pone.0145002> PMID: 26673219; PubMed Central PMCID: PMC4686021.
6. Bäck K, Ahlqwist M, Hakeberg M, Dahlström L. Occurrence of signs of osteoarthritis/arthritis in the temporomandibular joint on panoramic radiographs in Swedish women. *Community Dent Oral Epidemiol*. 2017; 45(5):478–84. Epub 20170712. <https://doi.org/10.1111/cdoe.12312> PMID: 28699681.
7. Toshima H, Ogura I. Characteristics of patients with temporomandibular joint osteoarthritis on magnetic resonance imaging. *J Med Imaging Radiat Oncol*. 2020; 64(5):615–9. Epub 20200601. <https://doi.org/10.1111/1754-9485.13054> PMID: 32478461.
8. Campos MI, Campos PS, Cangussu MC, Guimarães RC, Line SR. Analysis of magnetic resonance imaging characteristics and pain in temporomandibular joints with and without degenerative changes of the condyle. *Int J Oral Maxillofac Surg*. 2008; 37(6):529–34. Epub 20080428. <https://doi.org/10.1016/j.ijom.2008.02.011> PMID: 18440778.
9. Bianchi J, Gonçalves JR, de Oliveira Ruellas AC, Ashman LM, Vimort JB, Yatabe M, et al. Quantitative bone imaging biomarkers to diagnose temporomandibular joint osteoarthritis. *Int J Oral Maxillofac Surg*. 2021; 50(2):227–35. Epub 20200627. <https://doi.org/10.1016/j.ijom.2020.04.018> PMID: 32605824; PubMed Central PMCID: PMC7765741.
10. Ottersen MK, Abrahamsson AK, Larheim TA, Arvidsson LZ. CBCT characteristics and interpretation challenges of temporomandibular joint osteoarthritis in a hand osteoarthritis cohort. *Dentomaxillofac Radiol*. 2019; 48(4):20180245. Epub 20190128. <https://doi.org/10.1259/dmfr.20180245> PMID: 30633546; PubMed Central PMCID: PMC6592579.
11. Crow HC, Parks E, Campbell JH, Stucki DS, Daggy J. The utility of panoramic radiography in temporomandibular joint assessment. *Dentomaxillofac Radiol*. 2005; 34(2):91–5. <https://doi.org/10.1259/dmfr/24863557> PMID: 15829691.
12. Poveda-Roda R, Bagan J, Carbonell E, Margaix M. Diagnostic validity (sensitivity and specificity) of panoramic X-rays in osteoarthritis of the temporomandibular joint. *Cranio*. 2015; 33(3):189–94. Epub 20140731. <https://doi.org/10.1179/2151090314Y.0000000018> PMID: 25079854.
13. Oliveira SR, Oliveira RDS, Rodrigues ED, Junqueira JLC, Panzarella FK. Accuracy of Panoramic Radiography for Degenerative Changes of the Temporomandibular Joint. *J Int Soc Prev Community Dent*. 2020; 10(1):96–100. Epub 20200211. [https://doi.org/10.4103/jispcd.JISPCD\\_411\\_19](https://doi.org/10.4103/jispcd.JISPCD_411_19) PMID: 32181226; PubMed Central PMCID: PMC7055348.
14. Schnabl D, Rottler AK, Schupp W, Boisserée W, Grunert I. CBCT and MRT imaging in patients clinically diagnosed with temporomandibular joint arthralgia. *Heliyon*. 2018; 4(6):e00641. Epub 20180605. <https://doi.org/10.1016/j.heliyon.2018.e00641> PMID: 30003149; PubMed Central PMCID: PMC6040602.
15. Hilgenberg-Sydney PB, Bonotto DV, Stechman-Neto J, Zwir LF, Pachêco-Pereira C, Canto GL, et al. Diagnostic validity of CT to assess degenerative temporomandibular joint disease: a systematic review. *Dentomaxillofac Radiol*. 2018; 47(5):20170389. Epub 20180302. <https://doi.org/10.1259/dmfr.20170389> PMID: 29480019; PubMed Central PMCID: PMC6196046.
16. Grossmann E, Remedi MP, Ferreira LA, Carvalho AC. Magnetic Resonance Image Evaluation of Temporomandibular Joint Osteophytes: Influence of Clinical Factors and Artrogenic Changes. *J Craniofac Surg*. 2016; 27(2):334–8. <https://doi.org/10.1097/SCS.0000000000002377> PMID: 26825745.
17. Morales H, Cornelius R. Imaging Approach to Temporomandibular Joint Disorders. *Clin Neuroradiol*. 2016; 26(1):5–22. Epub 20150915. <https://doi.org/10.1007/s00062-015-0465-0> PMID: 26374243.
18. Massilla Mani F, Sivasubramanian SS. A study of temporomandibular joint osteoarthritis using computed tomographic imaging. *Biomed J*. 2016; 39(3):201–6. Epub 20160831. <https://doi.org/10.1016/j.bj.2016.06.003> PMID: 27621122; PubMed Central PMCID: PMC6138784.
19. Larheim TA, Hol C, Ottersen MK, Mork-Knutsen BB, Arvidsson LZ. The Role of Imaging in the Diagnosis of Temporomandibular Joint Pathology. *Oral Maxillofac Surg Clin North Am*. 2018; 30(3):239–49. Epub 20180601. <https://doi.org/10.1016/j.coms.2018.04.001> PMID: 29866450.

20. Kononenko I. Machine learning for medical diagnosis: history, state of the art and perspective. *Artif Intell Med*. 2001; 23(1):89–109. [https://doi.org/10.1016/s0933-3657\(01\)00077-x](https://doi.org/10.1016/s0933-3657(01)00077-x) PMID: 11470218.
21. Brosset S, Dumont M, Bianchi J, Ruellas A, Cevidanes L, Yatabe M, et al. 3D Auto-Segmentation of Mandibular Condyles. *Annu Int Conf IEEE Eng Med Biol Soc*. 2020; 2020:1270–3. <https://doi.org/10.1109/EMBC44109.2020.9175692> PMID: 33018219; PubMed Central PMCID: PMC7771389.
22. Karhade AV, Schwab JH, Bedair HS. Development of Machine Learning Algorithms for Prediction of Sustained Postoperative Opioid Prescriptions After Total Hip Arthroplasty. *J Arthroplasty*. 2019; 34(10):2272–7.e1. Epub 20190613. <https://doi.org/10.1016/j.arth.2019.06.013> PMID: 31327647.
23. Lee JH, Kim DH, Jeong SN, Choi SH. Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. *J Periodontol Implant Sci*. 2018; 48(2):114–23. Epub 20180430. <https://doi.org/10.5051/jpis.2018.48.2.114> PMID: 29770240; PubMed Central PMCID: PMC5944222.
24. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011; 155(8):529–36. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009> PMID: 22007046.
25. Choi E, Kim D, Lee JY, Park HK. Artificial intelligence in detecting temporomandibular joint osteoarthritis on orthopantomogram. *Sci Rep*. 2021; 11(1):10246. Epub 20210513. <https://doi.org/10.1038/s41598-021-89742-y> PMID: 33986459; PubMed Central PMCID: PMC8119725
26. Lee KS, Kwak HJ, Oh JM, Jha N, Kim YJ, Kim W, et al. Automated Detection of TMJ Osteoarthritis Based on Artificial Intelligence. *J Dent Res*. 2020; 99(12):1363–7. Epub 20200701. <https://doi.org/10.1177/0022034520936950> PMID: 32609562.
27. Jung W, Lee KE, Suh BJ, Seok H, Lee DW. Deep learning for osteoarthritis classification in temporomandibular joint. *Oral Dis*. 2023; 29(3):1050–9. Epub 20211101. <https://doi.org/10.1111/odi.14056> PMID: 34689379.
28. Kim D, Choi E, Jeong HG, Chang J, Youm S. Expert System for Mandibular Condyle Detection and Osteoarthritis Classification in Panoramic Imaging Using R-CNN and CNN. *Applied Sciences*. 2020; 10:7464. <https://doi.org/10.3390/app10217464>
29. Orhan K, Driesen L, Shujaat S, Jacobs R, Chai X. Development and Validation of a Magnetic Resonance Imaging-Based Machine Learning Model for TMJ Pathologies. *Biomed Res Int*. 2021; 2021:6656773. Epub 20210705. <https://doi.org/10.1155/2021/6656773> PMID: 34327235; PubMed Central PMCID: PMC8277497.
30. Haghnegahdar AA, Kolahi S, Khojastepour L, Tajeripour F. Diagnosis of Temporomandibular Disorders Using Local Binary Patterns. *J Biomed Phys Eng*. 2018; 8(1):87–96. Epub 20180301. <https://doi.org/10.1259/dmfr/25151578> PMID: 29732343; PubMed Central PMCID: PMC5928314.
31. Schiffman E, Ohrbach R, Truelove E, Look J, Anderson G, Goulet JP, et al. Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) for Clinical and Research Applications: recommendations of the International RDC/TMD Consortium Network\* and Orofacial Pain Special Interest Group†. *J Oral Facial Pain Headache*. 2014; 28(1):6–27. <https://doi.org/10.11607/jop.1151> PMID: 24482784; PubMed Central PMCID: PMC4478082.
32. de Dumast P, Mirabel C, Cevidanes L, Ruellas A, Yatabe M, Ioshida M, et al. A web-based system for neural network based classification in temporomandibular joint osteoarthritis. *Comput Med Imaging Graph*. 2018; 67:45–54. Epub 20180501. <https://doi.org/10.1016/j.compmedimag.2018.04.009> PMID: 29753964; PubMed Central PMCID: PMC5987251.
33. Ribera NT, de Dumast P, Yatabe M, Ruellas A, Ioshida M, Paniagua B, et al. Shape variation analyzer: a classifier for temporomandibular joint damaged by osteoarthritis. *Proc SPIE Int Soc Opt Eng*. 2019; 10950. Epub 20190313. <https://doi.org/10.1117/12.2506018> PMID: 31359900; PubMed Central PMCID: PMC6663087.
34. Almășan O, Leucuța DC, Hedeșiu M, Mureșanu S, Popa Ș L. Temporomandibular Joint Osteoarthritis Diagnosis Employing Artificial Intelligence: Systematic Review and Meta-Analysis. *J Clin Med*. 2023; 12(3). Epub 20230125. <https://doi.org/10.3390/jcm12030942> PMID: 36769590; PubMed Central PMCID: PMC9918072.
35. Jha N, Lee KS, Kim YJ. Diagnosis of temporomandibular disorders using artificial intelligence technologies: A systematic review and meta-analysis. *PLoS One*. 2022; 17(8):e0272715. Epub 20220818. <https://doi.org/10.1371/journal.pone.0272715> PMID: 35980894; PubMed Central PMCID: PMC9387829.
36. Bedrikovetski S, Dudi-Venkata NN, Maicas G, Kroon HM, Seow W, Carneiro G, et al. Artificial intelligence for the diagnosis of lymph node metastases in patients with abdominopelvic malignancy: A systematic review and meta-analysis. *Artif Intell Med*. 2021; 113:102022. Epub 20210202. <https://doi.org/10.1016/j.artmed.2021.102022> PMID: 33685585.

37. Kaimal S, Ahmad M, Kang W, Nixdorf D, Schiffman EL. Diagnostic accuracy of panoramic radiography and MRI for detecting signs of TMJ degenerative joint disease. *Gen Dent*. 2018; 66(4):34–40. PMID: [29964246](https://pubmed.ncbi.nlm.nih.gov/29964246/); PubMed Central PMCID: PMC9488601.
38. Detterbeck A, Hofmeister M, Hofmann E, Haddad D, Weber D, Hölzing A, et al. MRI vs. CT for orthodontic applications: comparison of two MRI protocols and three CT (multislice, cone-beam, industrial) technologies. *J Orofac Orthop*. 2016; 77(4):251–61. Epub 20160420. <https://doi.org/10.1007/s00056-016-0028-2> PMID: [27098643](https://pubmed.ncbi.nlm.nih.gov/27098643/)
39. Bertram S, Rudisch A, Innerhofer K, Pümpel E, Grubwieser G, Emshoff R. Diagnosing TMJ internal derangement and osteoarthritis with magnetic resonance imaging. *J Am Dent Assoc*. 2001; 132(6):753–61. <https://doi.org/10.14219/jada.archive.2001.0272> PMID: [11433854](https://pubmed.ncbi.nlm.nih.gov/11433854/).
40. Takaoka R, Yatani H, Senzaki Y, Koishi Y, Moriguchi D, Ishigaki S. Relative risk of positional and dynamic temporomandibular disc abnormality for osteoarthritis-magnetic resonance imaging study. *J Oral Rehabil*. 2021; 48(4):375–83. Epub 20210111. <https://doi.org/10.1111/joor.13138> PMID: [33372318](https://pubmed.ncbi.nlm.nih.gov/33372318/).
41. Bach SH, He B, Ratner A, Ré C. Learning the Structure of Generative Models without Labeled Data. *Proc Mach Learn Res*. 2017; 70:273–82. PMID: [30882087](https://pubmed.ncbi.nlm.nih.gov/30882087/); PubMed Central PMCID: PMC6417840.
42. Samala RK, Chan HP, Hadjiiski L, Helvie MA. Risks of feature leakage and sample size dependencies in deep feature extraction for breast mass classification. *Med Phys*. 2021; 48(6):2827–37. Epub 20210412. <https://doi.org/10.1002/mp.14678> PMID: [33368376](https://pubmed.ncbi.nlm.nih.gov/33368376/); PubMed Central PMCID: PMC8601676.
43. Brui E, Efimtcev AY, Fokin VA, Fernandez R, Levchuk AG, Ogier AC, et al. Deep learning-based fully automatic segmentation of wrist cartilage in MR images. *NMR Biomed*. 2020; 33(8):e4320. Epub 20200511. <https://doi.org/10.1002/nbm.4320> PMID: [32394453](https://pubmed.ncbi.nlm.nih.gov/32394453/); PubMed Central PMCID: PMC7784718.
44. Fang Y, Wang J, Ou X, Ying H, Hu C, Zhang Z, et al. The impact of training sample size on deep learning-based organ auto-segmentation for head-and-neck patients. *Phys Med Biol*. 2021; 66(18). Epub 20210914. <https://doi.org/10.1088/1361-6560/ac2206> PMID: [34450599](https://pubmed.ncbi.nlm.nih.gov/34450599/).
45. Bahrami A, Karimian A, Arabi H. Comparison of different deep learning architectures for synthetic CT generation from MR images. *Phys Med*. 2021; 90:99–107. Epub 20210929. <https://doi.org/10.1016/j.ejmp.2021.09.006> PMID: [34597891](https://pubmed.ncbi.nlm.nih.gov/34597891/).
46. Gautam R, Sharma M. Prevalence and Diagnosis of Neurological Disorders Using Different Deep Learning Techniques: A Meta-Analysis. *J Med Syst*. 2020; 44(2):49. Epub 20200104. <https://doi.org/10.1007/s10916-019-1519-7> PMID: [31902041](https://pubmed.ncbi.nlm.nih.gov/31902041/).
47. Putra RH, Doi C, Yoda N, Astuti ER, Sasaki K. Current applications and development of artificial intelligence for digital dental radiography. *Dentomaxillofac Radiol*. 2022; 51(1):20210197. Epub 20210708. <https://doi.org/10.1259/dmfr.20210197> PMID: [34233515](https://pubmed.ncbi.nlm.nih.gov/34233515/); PubMed Central PMCID: PMC8693331.
48. Yu J, Wang X, Chen X, Guo J. Automatic Premature Ventricular Contraction Detection Using Deep Metric Learning and KNN. *Biosensors (Basel)*. 2021; 11(3). Epub 20210304. <https://doi.org/10.3390/bios11030069> PMID: [33806367](https://pubmed.ncbi.nlm.nih.gov/33806367/); PubMed Central PMCID: PMC8000997.
49. Kim EK, Wu JT, Tamura S, Close R, Taketan H, Kawai H, et al. Comparison of neural network and k-NN classification methods in medical image and voice recognitions. *Med J Osaka Univ*. 1993; 41-42(1–4):11–6. PMID: [7476650](https://pubmed.ncbi.nlm.nih.gov/7476650/).
50. Ay B, Turker C, Emre E, Ay K, Aydin G. Automated classification of nasal polyps in endoscopy video-frames using handcrafted and CNN features. *Comput Biol Med*. 2022; 147:105725. Epub 20220613. <https://doi.org/10.1016/j.compbiomed.2022.105725> PMID: [35716434](https://pubmed.ncbi.nlm.nih.gov/35716434/).
51. Pantoja LLQ, de Toledo IP, Pupo YM, Porporatti AL, De Luca Canto G, Zwir LF, et al. Prevalence of degenerative joint disease of the temporomandibular joint: a systematic review. *Clin Oral Investig*. 2019; 23(5):2475–88. Epub 20181011. <https://doi.org/10.1007/s00784-018-2664-y> PMID: [30311063](https://pubmed.ncbi.nlm.nih.gov/30311063/).
52. Alzahrani A, Yadav S, Gandhi V, Lurie AG, Tadinada A. Incidental findings of temporomandibular joint osteoarthritis and its variability based on age and sex. *Imaging Sci Dent*. 2020; 50(3):245–53. Epub 20200916. <https://doi.org/10.5624/isd.2020.50.3.245> PMID: [33005582](https://pubmed.ncbi.nlm.nih.gov/33005582/); PubMed Central PMCID: PMC7506092.
53. Zhao YP, Ma XC. [Temporomandibular disorders related pain interaction with age, sex and imaging changes of osteoarthritis]. *Zhonghua Kou Qiang Yi Xue Za Zhi*. 2006; 41(12):757–8. PMID: [17349201](https://pubmed.ncbi.nlm.nih.gov/17349201/).
54. Xue XT, Zhang T, Cui SJ, He DQ, Wang XD, Yang RL, et al. Sexual dimorphism of estrogen-sensitized synoviocytes contributes to gender difference in temporomandibular joint osteoarthritis. *Oral Dis*. 2018; 24(8):1503–13. Epub 20180820. <https://doi.org/10.1111/odi.12905> PMID: [29806726](https://pubmed.ncbi.nlm.nih.gov/29806726/).
55. Robinson JL, Cass K, Aronson R, Choi T, Xu M, Buttenbaum R, et al. Sex differences in the estrogen-dependent regulation of temporomandibular joint remodeling in altered loading. *Osteoarthritis Cartilage*.

2017; 25(4):533–43. Epub 20161127. <https://doi.org/10.1016/j.joca.2016.11.008> PMID: 27903449; PubMed Central PMCID: PMC5359071.

56. Kang JH. Transcriptomes in peripheral blood of young females with temporomandibular joint osteoarthritis. *Sci Rep.* 2021; 11(1):8872. Epub 20210423. <https://doi.org/10.1038/s41598-021-88275-8> PMID: 33893371; PubMed Central PMCID: PMC8065155.
57. Koç N. Evaluation of osteoarthritic changes in the temporomandibular joint and their correlations with age: A retrospective CBCT study. *Dent Med Probl.* 2020; 57(1):67–72. <https://doi.org/10.17219/dmp/112392> PMID: 31997586.
58. Lamot U, Strojjan P, Šurlan Popovič K. Magnetic resonance imaging of temporomandibular joint dysfunction-correlation with clinical symptoms, age, and gender. *Oral Surg Oral Med Oral Pathol Oral Radiol.* 2013; 116(2):258–63. <https://doi.org/10.1016/j.oooo.2013.04.019> PMID: 23849379.
59. Manfredini D, Piccotti F, Ferronato G, Guarda-Nardini L. Age peaks of different RDC/TMD diagnoses in a patient population. *J Dent.* 2010; 38(5):392–9. Epub 20100125. <https://doi.org/10.1016/j.jdent.2010.01.006> PMID: 20100537.