

ICS 点击此处添加 ICS 号

CCS 点击此处添加 CCS 号



团 体 标 准

T/ XXXX—XXXX

人工智能技术在牙颌面畸形诊治中的应用 专家共识

Expert Consensus on the Application of Artificial Intelligence Technology in the
Diagnosis and Treatment of Dentofacial Deformities

(征求意见稿)

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

2026 - XX - XX 发布

2026 - XX - XX 实施

中华口腔医学会 发布

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中华口腔医学会口腔正畸专业委员会，口腔医学计算机专业委员会及口腔美学专业委员会提出。

本文件由中华口腔医学会归口。

本文件起草单位（按首字拼音排序）：北京大学口腔医院，重庆医科大学附属口腔医院，福建医科大学附属口腔医院，华中科技大学同济医学院，吉林大学口腔医院，空军军医大学第三附属医院，南昌大学附属口腔医院，南京医科大学附属口腔医院，上海交通大学医学院附属第九人民医院，上海科技大学人工智能学院，沈阳市口腔医院，首都医科大学附属北京口腔医院，四川大学华西口腔医院，同济大学附属口腔医院，武汉大学口腔医学院，浙江大学计算机学院，中山大学附属口腔医院

本文件主要起草人：房兵、沈定刚、金作林、贺红、白玉兴、陈江、王林、李巍然、王军、金钊、陈莉莉、胡敏、曹阳、严斌、宋锦璘、李志华、夏伦果、谢贤聚、张宁、黄翠、韩向龙、张桂荣、刘伟才、刘峰、刘晓强、张思慧。

引 言

牙颌面畸形诊疗具有病程长、影响因素多、个体差异显著和多模态信息整合要求高等特点。临床决策不仅依赖病史、口内外检查、影像资料和模型分析，还需综合患者生长发育状态、软硬组织关系、功能表现、治疗目标、依从性及风险收益比进行判断。近年来，人工智能技术在口腔医学领域快速发展，已逐步应用于影像识别、自动测量、风险提示、诊疗流程管理和医患沟通等场景，为提高效率、改善一致性和支持精准诊疗提供了新的工具^[1]。

与此同时，人工智能在牙颌面畸形诊疗中的应用仍面临数据来源异质性强、样本代表性不足、标注标准不统一、模型泛化能力有限、透明性与可解释性不足、部署后性能漂移以及责任边界不清等问题^[2]。尤其是生成式人工智能和多模态大模型在信息整合、文本生成和候选方案输出方面具有潜在价值，但其结果仍可能存在错误、幻觉、逻辑不一致、忽略个体差异或与患者真实临床情况不符等风险，不得在脱离医师监督的情况下独立用于诊疗决策。

因此，有必要针对人工智能技术在牙颌面畸形诊疗中的应用边界、实施条件、验证要求、伦理合规和责任划分形成规范性意见，以促进其合理使用，减少误用和滥用风险，并推动相关研究和产品转化朝着临床真实需求导向发展^[3]。

本共识旨在为人工智能技术在牙颌面畸形诊疗中的研发、验证、部署和临床应用提供规范化建议，促进人工智能工具在提升诊疗效率、优化工作流程、辅助临床决策、支持患者管理及服务教学科研方面的合理应用，并降低不当依赖、超范围使用及安全风险。

人工智能技术在牙颌面畸形诊治中的应用专家共识

1 范围

本共识适用于人工智能技术在牙颌面畸形相关诊疗场景中的应用，包括但不限于

- a) 口腔及颅颌面影像的识别、分割、测量和标准化结果提取与整理；
- b) 牙颌面畸形及相关错殆问题的辅助筛查、辅助诊断和风险提示；
- c) 生长发育评估、骨龄评估、牙槽骨边界评估，以及颞下颌关节结构、气道状况等相关风险评估；
- d) 治疗方案辅助设计、疗效预测和治疗过程监测；
- e) 医患沟通、患者宣教、远程随访、教学培训和科研辅助；
- f) 与上述用途相关的数据治理、模型评价、部署管理和持续监测。

本共识不适用于将人工智能系统作为脱离执业医师监督的独立诊断者或独立治疗决策者；不适用于缺乏明确预期用途、缺乏基本验证证据、无法说明数据来源和适用边界、不能满足数据安全与隐私保护要求的人工智能工具或系统。

2 规范性引用文件

本文件没有规范性引用文件。

3 术语和定义

下列术语和定义适用于本文件。

3.1

人工智能 Artificial Intelligence

人工智能是指通过计算机系统模拟、扩展或辅助人类完成感知、识别、分析、推理、预测、生成和决策等任务的技术总称。

3.2

医疗人工智能 Medical Artificial Intelligence

医疗人工智能是指应用于疾病预防、筛查、诊断、治疗、康复、随访、教学、科研和管理等医学场景的人工智能技术与系统。

3.3

牙颌面畸形诊疗人工智能系统 Artificial Intelligence System for Diagnosis and Treatment of Dentofacial Deformities

牙颌面畸形诊疗人工智能系统是指用于牙颌面畸形相关信息获取、影像分析、辅助诊断、风险评估、治疗方案辅助设计、疗效监测、患者管理或沟通支持的人工智能工具或系统。

3.4

生成式人工智能 Generative Artificial Intelligence

生成式人工智能是指能够基于训练数据和输入提示自动生成文本、图像、报告、候选方案或其他内容的人工智能系统。其输出可用于辅助信息整合与表达，但不应直接替代临床判断。

3.5

多模态人工智能 Multimodal Artificial Intelligence

多模态人工智能是指能够联合处理文本、影像、三维模型、面像、时序信息及临床记录等多类型数据，并输出分析结果或建议的人工智能系统。

3.6

预期用途 Intended Use

预期用途是指人工智能系统被设计用于完成的具体任务、目标人群、使用者、使用环境和临床流程位置。

3.7

外部验证 External Validation

外部验证是指在独立于模型开发过程的数据集、机构、设备或人群中，对模型性能和泛化能力进行评价的过程。

3.8

人工智能可解释性 Artificial Intelligence Interpretability

人工智能可解释性是指人工智能系统的输出结果、主要依据、不确定性、适用边界及失效情形能够在合理范围内被临床使用者理解、说明和审查的特性。可解释性并不要求系统完全公开全部底层代码或参数，但应能够支持结果核验、异常识别、风险沟通和责任追溯。

3.9

联邦学习 Federated Learning

联邦学习是指在多机构或多节点协作场景下，数据不直接离开原始存储位置，由各参与方在本地完成模型训练，仅共享模型参数、梯度或其他中间更新信息，并通过中心服务器或协同机制完成聚合的机器学习方法。联邦学习有助于减少原始数据集中汇聚，但并不当然免除隐私保护、数据安全、算法偏倚和责任划分要求。

3.10

可信计算 Trusted Computing

可信计算是指在不直接暴露原始数据内容的前提下，通过安全硬件、密码学协议或受控执行环境等技术手段，对数据进行分析、协同训练或结果验证的技术路径。其主要目的是在多中心协作中提升数据处理过程的安全性、可控性和可审计性。

3.11

人工智能模型 Artificial Intelligence Model

指通过训练形成、能够对输入数据进行识别、分割、测量、预测或生成的核心推理单元。

3.12

人工智能软件/系统 Artificial Intelligence Software/System

指以一种或多种人工智能模型、算法或规则引擎为核心功能，并结合输入输出接口、部署环境、日志管理、权限控制和人工复核流程，用于实现特定医疗用途的软件产品或临床应用系统。

4 总体原则

4.1 临床主导原则

人工智能在牙颌面畸形诊疗中的定位应为临床辅助工具，不应作为独立诊疗决策主体，其作用是支持医师识别信息、整理信息、提高效率和提示风险，而不是替代医师作出最终诊断和治疗决策。

推荐意见：人工智能输出的所有关键结果，均应纳入临床医师主导的综合诊断框架之中进行解释和使用；最终诊疗责任由具有相应资质的临床医师承担。

4.2 安全有效原则

人工智能系统进入临床应用前，应具备与其预期用途相匹配的技术验证和临床证据，证明其在目标人群和目标场景中具有可接受的准确性、稳定性和安全性。

推荐意见：拟进入临床使用的人工智能系统，应在代表性数据上完成必要验证；缺乏基本证据支持的系统不宜直接纳入临床流程。

4.3 场景限定原则

人工智能系统只能在其设计用途和验证范围内使用。不同场景对准确性、可解释性、容错率和人工复核要求不同，不应将一个场景中表现良好的模型直接迁移到其他场景。

推荐意见：使用前应明确人工智能系统的任务边界、适用对象、输入要求、输出形式和失效场景；超出训练分布或超出验证范围的病例应谨慎使用或避免使用。

4.4 人机协同原则

人工智能的临床价值不在于完全自动化，而在于提升“人机协同团队”的整体表现。临床使用中应重视医师对输出结果的理解、核查、修正和否决能力。

推荐意见：对可能影响治疗方向、时机和风险控制的人工智能输出结果，需要设置人工复核与人工纠偏环节，不能机械接受人工智能结论。

4.5 透明可追溯原则

人工智能系统应尽可能向使用者明确说明其预期用途、训练与验证概况、适用条件、性能范围、已知局限性、版本信息及更新机制。临床使用过程应保留必要记录，确保结果可追溯^[4]。

推荐意见：医疗机构使用人工智能系统时，需要建立输入、输出、人工修正和最终临床决定的留痕机制，以便质量控制、复盘分析和责任追踪。对高风险应用场景，透明可追溯要求应与可解释性要求协同落实，不应仅停留于日志留存或版本记录层面。

4.6 隐私保护与合规原则

患者数据采集、传输、存储、标注、训练、部署和共享全过程，均应符合国家现行法律法规、伦理审查和医疗信息安全要求。

推荐意见：应落实数据最小必要使用、去标识化处理、分级授权、访问控制和安全审计；涉及第三方平台或云端服务时，应事先评估数据合规与安全风险。

4.7 持续监测原则

人工智能系统的临床应用不是一次性准入行为，而是覆盖设计、开发、部署、更新和退役的全生命周期管理过程。

推荐意见：部署后应建立常态化性能监测、异常上报、版本管理、偏差识别和更新评估机制，防止性能漂移、过拟合再训练和模型失效带来新的临床风险。

4.8 可解释性原则

人工智能系统应在与其预期用途和风险等级相匹配的范围内具备可解释性，使临床医师能够理解输出结果的含义、识别异常输出、判断是否需要人工复核，并在必要时向患者作出适当说明。

推荐意见：对直接影响诊断分型、拔牙决策、手术适应证判断和生长改良时机选择等高风险场景，应优先采用能够提供关键输入因素、置信度或不确定性提示、错误案例分析和失效场景说明的系统。开发者宜提供必要的解释信息和性能边界说明；医疗机构在准入评估与部署监测中，应将可解释性和可审计性纳入评价内容。

4.9 伦理优先原则

人工智能在牙颌面畸形诊疗中的研发、验证、部署和使用，应始终以患者安全、有效获益、隐私保护、公平可及和责任明确为前提，不应以提升效率为名削弱知情同意、人工监督或风险控制。对儿童青少年、复杂畸形、涉及显著面容改变和潜在长期不可逆后果的场景，应实施更严格的伦理审查和风险沟通。

推荐意见：涉及研究使用、数据二次利用、跨机构共享、自动化生成或远程管理等情形时，应依法履行伦理审查、知情告知/同意或其他合法处理基础；对可能对患者权益产生重大影响的自动化决策，不应仅依赖系统自动输出。

5 主要场景及推荐意见

本章所述应用场景可根据人工智能输出对临床决策的影响程度，按工作性分层大致分为低风险、中风险和高风险三类。低风险场景主要指信息整理、定点、测量、结构化报告、教学支持等，人工智能输出通常不直接改变治疗方案；中风险场景主要指筛查、风险提示、远程随访、进一步检查建议和部分疗效预测等，人工智能输出可能影响后续检查和管理路径，但需经医师复核；高风险场景主要指拔牙决策、正颌手术适应证判断、生长改良时机选择、关键支抗设计及其他可能直接改变治疗路径或显著影响患者安全的场景，人工智能不应独立作出最终结论。本文参考国际医疗器械监管机构论坛（International Medical Device Regulators Forum, IMDRF）提出的作为医疗器械的软件（Software as a Medical Device,

SaMD) 框架中的风险分类逻辑, 但不等同于国家药品监督管理局或IMDRF的正式监管类别^[5,6]。上述分层为便于临床使用设置的工作性分层, 不对应正式监管分类。

5.1 影像学分析

人工智能可用于牙颌面畸形诊疗相关二维与三维影像资料的识别、分割、定点、测量和结构化分析。侧位头影测量片、曲面断层片、CBCT/CT、MRI等影像学资料。二维影像主要用于头影测量、萌出与阻生筛查、生长发育和气道的初步评估; CBCT/CT主要用于牙、牙根、牙槽骨、颌骨、气道、埋伏牙、骨性不对称及颞下颌关节等三维硬组织和空间关系分析; MRI主要用于颞下颌关节盘-髁关系、关节盘形态、早期髁突改变、特发性髁突吸收风险及其他软组织/关节病变的辅助评估^[7-9]。MRI宜用于存在明确临床指征的情形, 例如颞下颌关节疼痛、开口受限、锁结、持续或伴症状的弹响/杂音、下颌运动异常、咬合进行性变化, 或临床怀疑关节盘移位、早期髁突改变、活动性髁突吸收, 以及正颌术前需评估颞下颌关节状态时。

推荐意见: 对标准化程度较高、重复性强的定点、分割和测量任务, 人工智能可作为提高效率 and 一致性的辅助工具。对可能直接影响治疗决策的输出结果, 如骨性关系判断、牙根位置评估、牙槽骨边界分析、埋伏牙定位、气道评估、颞下颌关节结构判断及正颌相关三维硬组织分析, 临床医生应进行人工复核并在必要时进行人工校正。CBCT/CT的使用应基于临床指征并遵循最小必要视野和剂量优化原则; MRI宜用于颞下颌关节症状、疑似关节盘移位、特发性髁突吸收、关节病变或正颌术前TMJ评估等有明确指征的情形。当影像质量不佳、伪影明显、存在解剖变异、术后改变或复杂畸形时, 不建议仅依据自动输出直接形成最终测量结论、影像诊断意见或治疗决策^[10]。

5.2 模型分析

人工智能可用于口内扫描及数字牙颌模型的自动分割、牙位标记、牙弓宽度和长度测量、拥挤度与间隙分析、Bolton比例分析、覆胎/覆盖分析、咬合接触关系评估、牙列对称性分析, 以及治疗前后牙位变化的自动追踪与比较^[2]。

推荐意见: 对基于口内扫描或数字牙颌模型的标准化测量和牙位分析任务, 人工智能可作为提高效率、减少操作者差异和支持病例归类的辅助工具。对边界型拔牙判断、复杂咬合重建、正颌联合正畸病例以及需要结合牙根、牙槽骨、关节和软组织信息综合判断的病例, 不应仅依据模型分析结果作出最终诊断或治疗决策。模型分析结果应与临床检查、影像学分析和患者个体情况综合解释。

5.3 辅助筛查与辅助诊断

人工智能可用于牙颌面畸形及相关错胎问题的筛查、病例标记、分层提示和风险预警, 如牙列拥挤、间隙、覆胎、覆盖异常、骨性关系异常、阻生牙风险和部分牙周支持组织不足的提示^[11]。

推荐意见: 人工智能可作为筛查、分诊和风险提示工具, 不应直接替代最终的诊断。牙颌面畸形的最终诊断仍应基于病史、临床检查、影像学分析、模型分析及必要的功能评估综合作出^[12]。

5.4 生长发育、骨龄及相关分析评估

人工智能可用于手腕片及颈椎成熟度、部分骨龄阶段、生长发育趋势以及牙槽骨界限、牙根移动风险、气道狭窄、颞下颌关节结构异常和牙槽骨吸收风险的辅助提示^[13]。

推荐意见: 骨龄和生长发育判断具有连续性和动态性, 人工智能输出可作为辅助参考, 但不能替代对临床生长潜力、家族史、既往发育轨迹和随访资料的综合判断^[14]。对生长加速期、边界期或人工智能输出与临床表现、人工判读结果或既往随访轨迹不一致的病例, 需要动态复查和人工复核。对气道、关节和牙周相关风险的评估, 应与临床检查和必要的影像检查结合判断^[10]。

5.5 治疗方案辅助设计与决策支持

人工智能可整合影像学数据、数字牙颌模型/口内扫描数据、面像数据、模型分析结果及病史资料, 辅助生成初步分析结果、病例归类、候选治疗路径和风险提示, 用于支持而非替代临床决策^[15]。

推荐意见: 人工智能可用于治疗方案初筛、病例归类、隐形矫治器设计辅助、牙移动目标位辅助生成及方案比较, 但对拔牙与否、拔牙牙位、支抗设计、扩弓边界、牙根移动安全性、正颌手术适应证及生长改良时机等高风险问题, 不建议由人工智能独立决定^[16]。最终方案应由临床医生在循证证据、临床经验和患者价值偏好共同基础上形成^[11]。

5.6 颜面美学评估、疗效预测与医患沟通

人工智能可用于软组织标志点识别、面部比例分析、对称性评估、部分治疗后外观变化预测和沟通图示生成，辅助医患沟通和治疗预期管理。

推荐意见：颜面美学评价具有明显个体性和主观性，人工智能可作为量化分析和沟通支持工具，但不应将其输出作为唯一的美学判断依据。对面像模拟、微笑模拟、外观预测或治疗后效果展示，应明确标注为“预测”或“仿真示意”，不得将预测图像等同于真实疗效，不得用于夸大疗效、弱化风险或替代知情同意过程。当生成式人工智能用于沟通图示、风险解释、术前预期展示或宣教文本生成时，相关输出应明确标注其“辅助生成”属性，并经临床医师审核后使用，不得将自动生成内容直接等同于正式医学意见或真实疗效承诺。

5.7 治疗过程监测、远程随访与健康管理

人工智能可用于治疗过程中的牙位变化追踪、附件识别、阶段性疗效评估、佩戴依从性提示、异常情况预警、远程随访和复诊提醒等。

推荐意见：对标准化程度较高的随访任务，人工智能有助于提高过程管理效率；对牙根吸收、牙周风险、支抗失控、复发倾向和关节症状等问题，人工智能提示仅可作为辅助参考，不应替代必要的临床复诊和专业检查。用于远程随访的系统应明确采集要求、适用人群、预警阈值和人工转诊路径^[17]。

5.8 教学培训与科研辅助

人工智能可用于病例整理、图像标注、知识检索、教学示教、虚拟现实和数字人辅助模拟训练、科研设计支持以及多中心协作研究管理。

推荐意见：教学和科研场景中可合理使用人工智能提升效率，但其生成的文字、图像、分析结果或文献总结，应经过教师或研究者审核，不应直接作为学术结论、考核依据或论文结果。使用生成式人工智能构建教学案例、病例摘要、问答脚本或模拟患者材料时，应明确标注其生成属性，不得与真实患者资料混淆。涉及学术发表、基金申请和学生培养时，应遵循学术诚信与伦理规范，明确人工智能的使用边界^[18]。

6 数据治理与标注规范

6.1 数据采集

牙颌面畸形人工智能研究通常涉及多模态数据，包括二维影像、三维口腔及面部数据、临床文本和随访信息。数据质量直接决定模型性能上限和泛化性能。

推荐意见：数据采集应尽量标准化，明确纳入与排除标准、设备信息、采集参数、时间点和临床背景；应尽可能覆盖不同年龄、性别、地区、错殆类型、疾病复杂程度和设备来源，以提升样本代表性和模型泛化能力。涉及科研用途、模型训练、跨机构共享或超出原始诊疗目的的数据二次使用时，应依法履行伦理审查，并根据具体处理情形选择知情同意或者法律法规规定的其他合法处理基础；同时应做好去标识化处理与最小必要处理。

6.2 数据存储与安全管理

数据治理应覆盖原始数据、预处理数据、标注数据、训练数据、测试数据和输出结果的全生命周期^[19]。

推荐意见：应建立分级分类管理、访问授权、加密存储、操作审计、备份恢复和最小必要使用机制；原始数据、标注结果和模型输出宜分别管理并保留版本信息。涉及跨机构协作、第三方平台或云端部署时，应事先明确数据权属、责任边界和安全要求^[20]。对医疗健康、生物识别等敏感个人信息，应开展个人信息保护影响评估，建立脱敏规范和再识别风险控制流程。

6.3 跨境数据流动、第三方云平台与外部模型依赖的合规要求

在牙颌面畸形人工智能研发、验证和部署过程中，如使用境外基础模型、境外云平台、境外API接口、外部托管算力或第三方预训练权重，应首先识别是否涉及向境外提供个人信息或其他依法受管控的

数据，并根据数据类型、处理目的、处理方式和接收方情况，依法选择相应合规路径。涉及患者医疗健康信息等敏感个人信息的，应采取更严格的审慎标准。

推荐意见：医疗机构或研发团队使用第三方云平台、外部模型或相关服务时，应明确数据流向、存储位置、处理角色、权限边界、日志留存、模型更新方式和退出机制；涉及患者敏感个人信息的，宜优先采用本地部署、专有云、专线隔离或经评估可控的境内托管方式^[21]。使用境外基础模型或外部接口进行推理、微调、蒸馏或提示增强时，不应将其视为单纯技术工具而忽视数据合规、隐私保护和安全管理义务；涉及相关场景时，应参考国家现行法律法规及配套规范执行^[22]。

6.4 数据标注与参考标准

高质量标注是医疗人工智能开发的基础。牙颌面畸形相关任务中的许多标注具有专业性强、主观性高和跨学科依赖的特点。标注人员需经过专业系统培训后方可开展相关工作。

推荐意见：标注前应制定统一、文档化的标注规范和操作手册；关键任务宜采用多名专家独立标注与一致性评价，并设置仲裁机制。参考标准应与模型预期用途匹配，并解释其局限性。对于定点、分割、分型和方案判断等任务，应报告标注一致性情况^[23]。标注一致性指标及报告建议见附录表3。

6.5 数据共享与多中心协作

多中心高质量数据协作有助于提高模型稳健性和外部适用性，但合作必须在合法、合规和可追责前提下开展。

推荐意见：鼓励在权威机构牵头下建立标准化的数据协作机制，统一数据字典、采集标准、标注规范和评价指标；在满足伦理和安全要求的前提下，可探索联邦学习、可信计算等多中心协作路径，但不应以共享为名降低数据质量控制和隐私保护要求。数据共享、委托处理或共同处理应根据合作模式签订相应协议，明确处理目的、方式、范围、期限、安全措施、双方权利义务、违约责任和退出机制；涉及向其他处理者提供个人信息、委托处理、敏感个人信息处理或跨境传输的，还应根据具体情形履行单独同意、个人信息保护影响评估以及数据出境相关程序等要求。不应在未依法履行相关程序、或不具备法律法规规定的合法处理基础的情况下共享患者敏感个人信息^[19]。人工智能应用工作性风险分层与管理要求见附录表4。

7 人工智能模型与软件的研发、验证与评价

7.1 模型研发与训练

模型研发应围绕明确的临床问题开展，避免“为技术而技术”。不同任务可采用不同算法架构与训练策略^[24]。

推荐意见：模型开发前应明确问题定义、预期用途、输入输出形式、终端用户、临床流程位置和失败后果；应重视过拟合、类别不平衡、数据泄漏和算法偏倚等问题。模型选择不应仅以单次性能指标高低为依据，而应综合考虑临床可用性、稳定性、可维护性和风险等级^[25]。使用开源基础模型、第三方预训练权重、外部API或云端推理服务进行微调、蒸馏或部署时，应核查其来源、许可条件、更新机制、训练数据声明、依赖组件安全性、版本锁定策略和可审计性；模型更替、权重升级、提示模板变化和第三方服务迁移均应纳入变更控制与再验证范围。

7.2 技术验证与临床验证

技术验证不能替代临床验证。模型在实验室环境中表现良好，并不代表其在真实临床流程中同样有效。

推荐意见：用于临床应用的人工智能模型，在转化为软件产品或临床系统前，应完成与预期用途相匹配的技术验证和临床验证；高风险用途应进一步开展外部验证，并尽量在多中心、前瞻性或真实世界场景中评价其临床效能、失败模式及对工作流程的影响^[26]。

7.3 性能评价指标

不同类型任务应采用相匹配的评价指标。仅报告准确率不能充分反映模型临床价值。

推荐意见：对分类任务，宜报告敏感度（sensitivity）、特异度（specificity）、阳性预测值（positive predictive value, PPV）、阴性预测值（negative predictive value, NPV）、受试者工作特征曲线下面积（area under the receiver operating characteristic curve, AUC）及95%置信区间；对分割任务，宜报告Dice相似系数（Dice similarity coefficient, DSC）或交并比（intersection over union, IoU）、边界误差及失败率；对回归和测量任务，宜报告平均绝对误差（mean absolute error, MAE）、均方根误差（root mean square error, RMSE）、临床可接受误差范围和一致性指标；对临床流程任务，还应评价时间效益、人工修正率、决策一致性和对患者结局的潜在影响。

7.4 外部验证、亚组分析与真实世界评估

牙颌面畸形诊疗存在明显的人群异质性和设备异质性，单中心结果往往难以直接推广。

推荐意见：宜对不同年龄段、性别、疾病复杂程度、设备来源和图像质量分层报告模型表现，识别模型可能表现欠佳的亚组和场景。对拟广泛推广的系统，应重视外部验证和真实世界连续病例评价，避免仅基于理想化数据集得出应用结论^[27]。

7.5 研究设计与报告规范

高质量证据形成不仅依赖模型性能，也依赖透明、完整和可评价的研究设计与报告^[26]。

推荐意见：开展人工智能相关研究时，可根据研究类型参照TRIPOD+AI、STARD-AI、SPIRIT-AI/CONSORT-AI和CLAIM 2024等国际通行报告规范，以提高研究的透明性、可重复性和可比较性^[25]。研究报告应明确模型训练数据来源、标注规范、验证流程及局限性，便于同行复核与推广。

8 软件/系统部署与使用管理

8.1 部署前评估

临床部署前，应从技术、流程、人员和管理四个层面进行准备。部署前评估宜结合本共识表4的工作性风险分层，分层确定本地化测试深度、人员培训要求和人工复核策略。

推荐意见：医疗机构在引入人工智能软件或系统前，应明确其预期用途、准入状态、技术要求、接口方式、数据流向、维护责任和应急处置方案^[28]；同时应完成必要的本地化测试和人员培训，确认其适配本机构病例构成和 workflows 后再投入常规使用^[29]。

8.2 使用过程中的人工复核与流程嵌入

人工智能的输出价值取决于其是否能被合理嵌入临床流程，而不是游离于流程之外形成额外负担。

推荐意见：临床使用中应明确哪些结果可直接参考、哪些结果必须人工复核、哪些异常情况需要升级处理；对关键节点输出结果，应保留人工否决和修正权限。人工智能不应削弱临床思考，也不应取代必要的医患沟通和知情过程。

8.3 版本管理与持续监测

模型更新、输入分布变化和临床环境变化均可能引起性能变化^[4]。

推荐意见：应对模型版本、更新时间、更新内容、验证结果和适用范围变化进行记录；重要更新后宜重新评估其临床适用性。对部署后的性能下降、系统性偏差、异常误报漏报或数据漂移，应及时分析原因并采取整改措施，必要时暂停相关模块使用。

8.4 不良事件与异常处理

人工智能输出错误在高风险场景下可能影响诊断方向、治疗计划和患者安全。

推荐意见：医疗机构应建立人工智能相关异常事件记录和上报机制，对明显错误输出、系统失效、患者投诉和潜在安全事件进行闭环管理；对可疑由系统引发的重大风险，应及时启动人工接管和风险评估程序^[19]。应预先明确人工接管的触发条件、责任岗位（人）、处置流程、响应时限和上报路径；异常事件处置后宜根据事件严重程度开展复盘和根因分析，形成纠正和预防措施，必要时暂停相关模块使用并重新评估。

9 生成式人工智能与多模态大模型的通用原则与风险防控

9.1 适用定位

生成式人工智能和多模态大模型可用于信息整合、非结构化资料整理、候选建议生成和表达支持，但其定位应为临床辅助工具，而非独立诊疗主体^[30]。

推荐意见：对诊断结论、拔牙决策、手术适应证判断、生长改良时机选择和处方生成等高风险任务，不应仅依据生成式人工智能输出直接作出最终临床决定^[31]。

9.2 幻觉与输出核查

生成式人工智能可能输出看似合理但实际错误、虚构依据、遗漏禁忌证或与患者资料不一致的内容。

推荐意见：对生成式人工智能生成的结论、摘要、建议、参考文献和图示，均应进行事实核查、逻辑核查和个体适配性核查；无法追溯依据或明显超出临床常识的内容，不应作为临床决策依据^[32]。

9.3 数据输入与平台边界

大模型应用常涉及提示词、临床文本、影像或其他敏感数据输入，对平台安全性和数据流向提出更高要求。

推荐意见：对使用生成式人工智能服务的医疗机构、教育科研机构和研发团队，应区分“向境内公众提供服务”与“机构内部研发/内部应用”两类场景；前者应遵守生成式人工智能服务相关规定，后者虽不直接适用该办法，但仍应遵守数据安全、个人信息保护、伦理审查和医疗器械/软件治理等要求^[20]。

9.4 留痕、版本管理与责任

生成式人工智能输出具有动态性和不确定性，同一问题在不同时间可能产生不同答案。

推荐意见：将生成式人工智能用于临床辅助时，宜保留关键提示词、版本信息、主要输出内容和人工修订记录，以保证过程可追溯。涉及模型升级、外部接口变更或提示词模板调整时，宜评估其对输出稳定性的影响。最终使用内容必须由临床医师或责任人员审核确认^[4]。

10 伦理、隐私、合规与责任边界

10.1 患者利益优先与知情透明

人工智能应用应始终以患者利益为首要原则，避免以提高效率为名牺牲安全性、解释充分性和患者权益^[33]。

推荐意见：当人工智能在诊疗流程中发挥重要辅助作用时，宜在适当范围内向患者说明其辅助性质、潜在价值和局限性；涉及自动生成内容、远程评估或重要风险提示时，应确保患者获得充分、准确和可理解的信息^[20]。

10.2 公平性与偏倚控制

训练数据不平衡、参考标准不一致和临床流程偏差，均可能导致模型对特定人群表现不佳。

推荐意见：模型开发和应用中应关注年龄、性别、地区、设备、疾病复杂程度等因素导致的性能差异，对可能受不公平影响的人群进行额外评估，不应以单一审美标准、单一人群样本或单一中心数据替代广泛适用性。

10.3 责任划分

人工智能应用涉及研发机构、产品提供方、医疗机构、临床使用者等多方主体。

推荐意见：研发机构或产品提供方应对系统的设计质量、更新维护、说明书完整性和预期性能负责；医疗机构应对准入评估、部署管理、培训和持续监测负责；临床医师应对具体病例中的解释、采纳、修正和最终诊疗决定负责。人工智能系统本身不具备责任主体资格^[23]。

10.4 能力建设

人工智能的规范应用以使用者具备基本理解能力为前提。

推荐意见：应加强口腔正畸及相关领域医师、技师、研究人员和管理者的人工智能素养培训，重点包括系统适用边界、常见失效模式、过度依赖风险、结果解读和伦理合规要求^[25]。

10.5 监管准入、分类界定与临床使用边界

人工智能产品在牙颌面畸形诊疗中的临床应用，不仅取决于技术性能和临床价值，还取决于其是否属于医疗器械、是否完成相应分类界定、是否满足注册审查要求以及是否在核准的预期用途范围内使用。根据国家药品监督管理局已发布的《人工智能医用软件产品分类界定指导原则》，人工智能医用软件是指基于医疗器械数据、采用人工智能技术实现其医疗用途的独立软件；含人工智能软件组件的医疗器械，其分类界定可参照执行。现行指导原则体系中，已将《医疗器械软件注册审查指导原则(2022年修订版)》和《人工智能医疗器械注册审查指导原则》纳入审评要求^[6]。对高风险工作性场景，医疗机构在完成产品属性和准入要求判断后，仍宜在院内另行建立更严格的使用权限、人工复核和异常事件管理机制。

推荐意见：拟用于牙颌面畸形诊疗的人工智能产品，在进入临床常规使用前，应明确其产品属性、分类界定、预期用途、适用人群、输入输出形式和临床使用边界；对属于医疗器械管理范畴的产品，应依法完成相应注册、备案或分类界定程序后方可按核准范围使用。医疗机构不应将研究型系统、教学型系统或仅用于辅助整理信息的模型，等同于已完成医疗器械准入的临床产品使用。对超说明书、超适用范围或模型频繁更新但未重新评估的系统，应谨慎引入临床流程。

11 不推荐或应谨慎应用的情形

11.1 不推荐直接应用的情形

以下情形不推荐将人工智能输出直接作为诊断或治疗依据：

- a) 系统缺乏明确预期用途或基本验证证据；
- b) 输入数据质量明显不达标，存在严重伪影、缺失或采集不规范；
- c) 患者属于训练和验证覆盖明显不足的人群；
- d) 综合征相关畸形、重大颅颌面发育异常、术后重建等高度复杂病例；
- e) 生成式人工智能输出无法追溯依据、存在明显幻觉或与患者真实资料明显不符。

11.2 应谨慎应用的情形

以下情形应在资深医师充分核验前提下谨慎使用人工智能：

- a) 边界型拔牙病例；
- b) 生长改良时机依赖长期动态判断的病例；
- c) 合并明显牙周问题、关节问题或口腔伴全身疾病的病例；
- d) 多模态资料缺失或信息不完整的病例。

11.3 人工智能与临床判断不一致时的处理

人工智能输出与临床判断不一致时，不应简单以系统结果替代人工判断，也不应在缺乏依据时完全忽略潜在风险提示。

推荐意见：当人工智能结果与临床结论冲突时，应优先核查输入质量、适用边界、病例复杂程度和模型失效可能性；必要时应补充检查、组织会诊或进行动态随访，不应直接接受人工智能输出结果。

12 结语

人工智能正在推动牙颌面畸形诊疗向数字化、智能化和精准化发展，但其真正临床价值取决于是否能够在明确边界内解决实际问题，并与临床 workflow 深度融合。现阶段，人工智能更适合作为提高效率、优化流程、辅助判断和支持沟通的工具，而非独立诊疗决策主体。未来应在高质量数据治理、多中心验证、真实世界评估、规范监管、持续监测和复合型人才培养的基础上，推动人工智能在牙颌面畸形诊疗中成为安全、有效的医师工作伙伴，并可持续应用。

附 录 A
(资料性)
多模态数据采集最低技术记录项

表A.1 多模态数据采集最低技术记录项

数据模态	最低技术记录项	标准化建议	原始文件与元数据留存要求
通用项	受试者唯一编码；采集日期与时间；采集地点；设备厂商/型号；软件版本；操作者；校准状态；临床阶段（初诊/治疗中/治疗后/随访）；采集目的；预处理流程；去标识化方式	同一研究队列或同一中心宜尽量采用同一设备、同一协议、同一体位与同一预处理流程；参数变更应留痕	保留原始文件、去标识化后文件、质控记录、日志文件和版本信息
CBCT	设备型号；软件版本；体素大小（mm）；FOV尺寸与类别；kVp；mA；曝光/扫描时间；重建参数；患者体位；是否采用低剂量/脉冲模式；采集适应证	FOV宜与ROI匹配；同队列尽量同协议；高分辨率任务可考虑更小体素，常规三维正畸评估宜优先使用能满足诊断任务的低剂量协议；不宜仅为生成头影侧位片或全景片而额外进行高剂量CBCT	优先保留原始DICOM；同时保存检查说明、FOV分类、重建方式和源序列UID
头影侧位片	设备/探测器类型；软件版本；kVp；mA；曝光时间；几何放大信息或装置固定参数；患者体位；是否使用cephalostat；图像后处理方式	尽量采用自然头位；Frankfort平面近似平行地面；正中矢状面平行探测器；牙列最大牙尖交错位；唇自然闭合；同一队列尽量同设备、同协议、同体位	保留原始 DICOM或原始数字影像；保留后处理前图像及放大率/几何信息
曲面断层片/其他2D影像	设备型号；软件版本；kVp；mA；曝光时间；患者定位方式；后处理参数	同中心同研究尽量统一采集协议；若与AI训练相关，应记录是否存在金属伪影、运动伪影、切层偏差	保留原始DICOM、后处理版本及质控记录
口内扫描（IOS）	扫描仪型号；软件版本；校准日期/状态；扫描策略；扫描范围（单颌/双颌/咬合）；操作者；是否重扫/拼接；明显缺失区和伪影；输出格式；单位	采集前宜清洁牙面并控制唾液；同队列尽量采用同一扫描仪、同一策略和同一操作者培训标准；研究启动前宜完成本地预实验并预设内部质控门槛	保留原始扫描工程文件；导出STL/Ply/Obj时保留单位、坐标系、对齐方式和版本信息
三维牙列模型/网格	文件格式；单位；坐标系；朝向；配准矩阵；来源模态；生成软件；版本；时间戳	不宜只保留“裸STL”；跨平台交换时应附带 sidecar manifest	保留源DICOM或源IOS文件、导出网格文件及 sidecar manifest（JSON/XML）
二维面像/摄影	相机/手机型号；镜头/等效焦距；图像分辨率；文件格式；背景；光源类型；拍摄距离；相机方向；拍摄位（正面静息、正面微笑、侧面等）；是否关闭美颜/HDR/自动塑形	宜保持背景、光照、拍摄距离和焦距一致；患者取自然头位；Frankfort平面近似平行地面；镜头平行地面；若用DSLR/无反，可参考f/8、1/200s、ISO 100作为外面像的工作设置	保留原始图像；不宜仅留压缩后社交媒体图片；应记录是否做过裁剪、缩放、白平衡或色彩校正
三维面像	设备型号；软件版本；采集方式（结构光/立体摄影等）；分辨率或点云密度；采集时长；表情状态；头位；遮挡情况（头发、眼镜、口罩）	宜统一中性表情、自然头位、稳定照明；避免明显表情变化和头位偏斜；前后随访尽量同设备同协议	保留原始点云/网格、纹理贴图及配准信息
临床文本/结构化病历	数据来源；抽取时间；字段字典；缺失值处理；结构化规则；术语映射方法	多中心研究宜统一变量字典和编码方式；对自由文本抽取规则应固定并版本化	保留原始病历索引、抽取脚本/规则版本和字段映射表

数据模态	最低技术记录项	标准化建议	原始文件与元数据留存要求
时序随访数据	随访时间点；记录频率；事件定义；设备与数据源；缺失原因；异常值处理方法	宜明确时间窗和事件判定标准；同一指标前后应尽量同一采集方式	保留时间戳、源系统标识、更新日志和数据清洗规则

附 录 B
(资料性)
建议采用的 CBCT FOV 记录方式

表B.1 建议采用的 CBCT FOV 记录方式

FOV类别	建议记录方式
小FOV	球形直径或圆柱高度 $\leq 10\text{cm}$
中FOV	球形直径或圆柱高度 $>10\text{cm}$ 且 $\leq 15\text{cm}$
大FOV	球形直径或圆柱高度 $>15\text{cm}$

附录 C
(资料性)
标注一致性指标建议表

表C.1 标注一致性指标建议表

标注任务类型	典型场景	建议主要指标	建议补充指标	报告建议
二分类/ 名义分类 (2名标注者)	是否拔牙、是否埋伏牙、是否存在骨开裂/骨开窗	Cohen' s κ	observed agreement (原始一致率); 95% CI	应说明类别定义、是否失衡、分歧处理规则
二分类/ 名义分类 (≥ 3 名标注者)	骨型分类、错殆分类、风险分层	Fleiss' κ 或 Conger' s κ	observed agreement; 必要时 Gwet' s AC1	应说明标注者数量、是否独立、仲裁流程
序位分类	生长发育分期、严重程度分级、美学等级	加权 κ ; 广义加权 κ ; 必要时 Gwet' s AC2	observed agreement	应说明权重方案(线性 / 二次)
连续测量	线距、角度、骨量、软组织比例	ICC或CCC	均值差、标准差、95%一致性界限(Bland - Altman)	应明确ICC模型类型与置信区间
头影标志点/ 三维标志点	2D/3D landmark annotation	平均径向误差(MRE)或平均欧氏距离	1mm、2mm、3mm 容许误差内成功检测率(SDR); ICC(用于派生测量)	应按标志点逐点报告,而非只给总体均值
分割任务	气道、颞下颌关节、牙根、颌骨、面部区域	Dice 相似系数(DSC)或IoU	Hausdorff distance 95(HD95); 平均表面距离(ASD); 失败率	宜报告多专家一致性和仲裁后共识掩膜生成方法
框选/ 检测任务	阻生牙检测、病灶框选、关键结构ROI	标注者间IoU	中心点距离; 框面积差; 失败率	应说明IoU阈值和异常样本处理
治疗方案排序/ 优先级	候选方案排序、风险优先级	加权 κ 或 Kendall' s W	一致率	应明确排序规则与并列项处理方式
文本抽取/ 结构化标签	从病历中抽取变量、风险因子或结局事件	Cohen' s κ / Fleiss' κ	精确率、召回率(如作为抽取流程质控)	应同时给出字段字典与抽取规则版本

附 录 D
(资料性)
人工智能应用工作性风险分层与管理要求

表D.1 人工智能应用工作性风险分层与管理要求

风险层级	工作性判定标准	典型场景	准入与验证要求	使用权限与人工复核强度
低风险	AI输出主要用于信息整理、定点、测量、结构化报告或教学支持，错误通常不会单独改变治疗路径，且较易被临床人员发现和纠正	自动头影定点、模型分割、牙弓测量、结构化报告草拟、教学示教辅助	宜完成本地化测试、质控和人员培训；对研究型系统宜明确“非临床最终结论”属性	可在常规临床流程中辅助使用；宜抽样复核或对异常值重点复核
中风险	AI输出会影响进一步检查、分诊、随访管理或部分临床管理路径，但通常仍保留充分人工复核和纠偏空间	辅助筛查、骨龄初筛、远程随访预警、依从性监测、复发风险提示	宜完成场景特异性验证、SOP建立和异常上报路径；必要时结合本地连续病例评估	对异常/阳性/边界结果应人工复核；关键节点宜由专科医师确认
高风险	AI输出直接参与或实质影响不可逆、高代价、长期影响明显的治疗决策，错误可能显著影响患者安全或预后	拔牙决策、正颌手术适应证判断、骨段移动规划、关键支抗设计、生长改良时机选择、TMJ/ICR高风险判断	宜优先使用完成相应准入或完成充分验证的产品；应建立专门SOP、版本控制、异常事件闭环和必要时MDT评估机制	每例均应由资深医师人工复核；不应自动采纳为最终结论；必要时多学科会诊
<p>说明：上述低、中、高风险分层为本共识为便于临床使用而设置的工作性分层，参考医疗软件风险分类的通用思路，重点考察人工智能输出对临床决策的影响程度、相应医疗情境的严重性、错误后果的可逆性以及人工复核空间；该分层不等同于NMPA或IMDRF的正式监管分类，也不替代医疗器械产品属性判断、分类界定、注册审查和机构准入程序。</p>				

参 考 文 献

- [1] World Health Organization. Ethics and governance of artificial intelligence for health: WHO Guidance[EB/OL]. (2021-06-28)[2026-04-10]. <https://www.who.int/publications/i/item/9789240029200>.
- [2] Nordblom N F, Büttner M, Schwendicke F. Artificial intelligence in orthodontics: critical review[J]. *Journal of Dental Research*, 2024, 103(6): 577-584.
- [3] Wei Y, Wu Y J, Deng J J, et al. Expert consensus on the application of artificial intelligence in stomatology[J]. *Chinese Journal of Clinical Thoracic and Cardiovascular Surgery*, 61(2): 151-161.
- [4] Government of Canada. Transparency for machine learning-enabled medical devices: guiding principles[EB/OL]. (2024-06-13)[2026-03-15]. <https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/transparency-machine-learning-guiding-principles.html>.
- [5] International Medical Device Regulator Forum. Software as a Medical Device: Possible Framework for Risk Categorization and Corresponding Considerations[EB/OL]. (2014-09-18)[2026-04-10]. <https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd>
- [6] 国家药监局. 人工智能医用软件产品分类界定指导原则[EB/OL]. (2021-07-01)[2026-04-10]. https://www.cdr-adr.org.cn/ylqx_1/Medical_zcfg/zcfg_zdyz_m/202107/t20210713_48808.html.
- [7] Liu J, Hao J, Lin H, et al. Deep learning-enabled 3D multimodal fusion of cone-beam CT and intraoral mesh scans for clinically applicable tooth-bone reconstruction[J]. *Patterns*, 2023, 4(9): 100825.
- [8] Hao J, Liao W, Zhang Y L, et al. Toward clinically applicable 3-dimensional tooth segmentation via deep learning[J]. *Journal of dental research*, 2022, 101(3): 304-311.
- [9] Tian Y, Hao J, Wang M, et al. Automatic jawbone structure segmentation on dental CBCT images via deep learning[J]. *Clinical Oral Investigations*, 2024, 28(12): 663.
- [10] Tejani A S, Klontzas M E, Gatti A A, et al. Checklist for artificial intelligence in medical imaging (CLAIM): 2024 update[J]. *Radiology: Artificial Intelligence*, 2024, 6(4): e240300.
- [11] Gracea R S, Winderickx N, Vanheers M, et al. Artificial intelligence for orthodontic diagnosis and treatment planning: A scoping review[J]. *Journal of dentistry*, 2025, 152: 105442.
- [12] Sounderajah V, Guni A, Liu X, et al. The STARD-AI reporting guideline for diagnostic accuracy studies using artificial intelligence[J]. *Nature medicine*, 2025, 31(10): 3283-3289.
- [13] Li H, Li H, Yuan L, et al. The psc-CVM assessment system: A three-stage type system for CVM assessment based on deep learning[J]. *BMC Oral Health*, 2023, 23(1): 557.
- [14] Kazimierczak W, Jedliński M, Issa J, et al. Accuracy of artificial intelligence for cervical vertebral maturation assessment—a systematic review[J]. *Journal of Clinical Medicine*, 2024, 13(14): 4047.
- [15] Shi Z, Meng Z, Chen R, et al. LETA: Tooth Alignment Prediction Based on Dual-branch Latent Encoding[J]. *IEEE transactions on visualization and computer graphics*, 2025, 31(9): 4805-4820.
- [16] Takeshita W M, Silva T P, De Souza L L T, et al. State of the art and prospects for artificial intelligence in orthognathic surgery: A systematic review with meta-analysis[J]. *Journal of Stomatology, Oral and Maxillofacial Surgery*, 2024, 125(6): 101787.
- [17] Snider V, Homsy K, Kusnoto B, et al. Clinical evaluation of artificial intelligence driven remote monitoring technology for assessment of patient oral hygiene during orthodontic treatment[J]. *American Journal of Orthodontics and Dentofacial Orthopedics*, 2024, 165(5): 586-592.
- [18] Thorat V, Rao P, Joshi N, et al. Role of Artificial Intelligence (AI) in Patient Education and Communication in Dentistry[J]. *Cureus*, 2024, 16(5): e59799.
- [19] 全国人民代表大会. 中华人民共和国数据安全法（2021年6月10日第十三届全国人民代表大会常务委 员会第二十九次会议通过） [EB/OL]. (2021-06-10)[2026-04-10]. http://www.npc.gov.cn/c2/c30834/202106/t20210610_311888.html.
- [20] 全国人民代表大会. 中华人民共和国个人信息保护法（2021年8月20日第十三届全国人民代表 大会常务委 员会第三十次会议通过） [EB/OL]. (2021-08-20)[2026-04-10]. http://www.npc.gov.cn/npc/c2/c30834/202108/t20210820_313088.html.
- [21] 中华人民共和国国家互联网信息办公室. 个人信息出境标准合同办法[EB/OL]. (2023-02-22)[2026-04-10]. https://www.cac.gov.cn/2023-02/24/c_1678884830036813.htm.
- [22] 国家互联网信息办公室. 促进和规范数据跨境流动规定[M]. 2023. https://www.cac.gov.cn/2024-03/22/c_1712776611775634.htm

- [23] 国家卫生健康委, 教育部, 科技部, 国家中医药局. 涉及人的生命科学和医学研究伦理审查办法 [EB/OL]. (2023-02-18)[2026-04-10]. https://www.gov.cn/zhengce/zhengceku/2023-02/28/content_5743658.htm.
- [24] International Medical Device Regulators Forum. Good machine learning practice for medical device development: Guiding principles[EB/OL]. (2025-01-29)[2026-03-15]. <https://www.imdrf.org/documents/good-machine-learning-practice-medical-device-development-guiding-principles>.
- [25] 国家药品监督管理局. 人工智能医疗器械注册审查指导原则[EB/OL]. (2022-03-09)[2026-04-10]. <https://www.cmde.org.cn/flfg/zdyz/zdyzwbk/20220309091014461.html>.
- [26] Collins G S, Moons K G M, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods[J]. *BMJ*, 2024, 385: e078378.
- [27] Moons K G M, Damen J A A, Kaul T, et al. PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods[J]. *BMJ*, 2025, 388: e082505.
- [28] Liu X, Cruz Rivera S, Moher D, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension[J]. *The Lancet. Digital Health*, 2020, 2(10): e537-e548.
- [29] Decide-Ai Steering Group. DECIDE-AI: new reporting guidelines to bridge the development-to-implementation gap in clinical artificial intelligence[J]. *Nature Medicine*, 2021, 27(2): 186-187.
- [30] World Health Organization. Ethics and governance of artificial intelligence for health: large multi-modal models[EB/OL]. (2025-03-25)[2026-04-10]. <https://www.who.int/publications/i/item/9789240084759>
- [31] Meng Z, Hao J, Dai X, et al. DentVLM: A multimodal vision-language model for comprehensive dental diagnosis and enhanced clinical practice[EB/OL]. (2025-09-27)[2026-04-10]. <https://doi.org/10.48550/arXiv.2509.23344>.
- [32] Baxmann M, Kárpáti K, Baráth Z. The potentials and challenges of integrating generative artificial intelligence (AI) in dental and orthodontic education: a systematic review[J]. *BMC oral health*, 2025, 25(1): 905.
- [33] Shaw J, Ali J, Atuire C A, et al. Research ethics and artificial intelligence for global health: perspectives from the global forum on bioethics in research[J]. *BMC medical ethics*, 2024, 25(1): 46.
-