

中国健康管理协会团体标准

T/CHAA

人工智能全科临床辅助诊断系统技术要求

Technical requirement on the system for AI-aided clinical diagnosis
in general practice

2024-00-00 发布

2024-00-00 实施

中国健康管理协会发布

目 次

前 言	1
1 范围	2
2 规范性引用文件	2
3 术语和定义	2
3.1 辅助诊断系统	2
3.1.1 辅助诊断	2
3.2 疑似度	2
3.3 诊断模型	2
3.4 主诉模型	3
3.5 诊断准确率	3
3.6 诊断能力提升率	3
3.7 检测项	4
4 应用流程	4
5 性能要求	5
5.1 诊病范围	5
5.2 诊断准确率	6
5.3 诊断能力提升率	6
5.4 可及时更新	6
5.5 具有可解释性	6
5.6 具有泛化能力	6
5.7 对数据依赖程度低	6
5.8 提示必检项	6
5.9 具备推荐检测功能	6
5.10 能表达和处理不确定证据	7
5.11 提示每次诊断的检测完备度	7
5.12 自动生成自然语言和结构化电子病历	7

5.13 具有患者病情信息管理功能	7
6 安全性要求	7
6.1 保护隐私	7
6.2 可溯源纠错	7
6.3 服务器安全	7
6.4 服务器存放环境安全	7
6.5 网络安全	8
6.6 数据安全	8
附录：技术指标汇总	8

前 言

本标准按照 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。

本标准由中国健康管理协会提出。

本标准由国家卫生健康委员会归口。

本标准起草单位：清华大学、北京予同智能科技有限公司、清华大学长庚医院、北京协和医院、中国人工智能学会因果与不确定性人工智能专业委员会、中国人工智能学会智慧医疗专业委员会、清华大学互联网产业研究院、重庆市中医院、青岛胶州市卫生健康局、青岛崂山区卫生健康局、中国国际科技促进会医工结合分会、中国民营科技实业家协会元宇宙工作委员会、浙江省中质健康科技标准化研究院。

本标准主要起草人：张勤、张湛、张萍、焦洋、朱岩、班晓娟、赵建磊、徐晓东、田丰玮、钱峰、徐亭、余清。

人工智能全科临床辅助诊断系统技术要求

1 范围

本标准规定了人工智能全科临床辅助诊断系统相关技术要求。

本标准适用于提供全科临床辅助诊断系统产品和服务的单位参照执行。

2 规范性引用文件

下列文件对于本标准的应用是必不可少的。凡是注日期的引用文件，仅注日期的版本适用于本标准。凡是不注日期的引用文件，其最新版本(包括所有的修改单)适用于本标准。

GBT20004.1-2016 国家标准《团体标准化第1部分：良好行为指南》

GB/T 41867-2022 信息技术 人工智能 术语

GB/T 8566-2022 系统与软件工程 软件生存周期过程

GB/T 35273-2020 个人信息安全规范

GB/T 25000.51-2016 软件工程软件产品质量要求和评价

YY/T 0316-2016 风险管理对医疗器械的应用

3 术语和定义

下列术语和定义适用于本标准。

3.1 辅助诊断系统

辅助作为用户的医生诊断患者所患疾病的人工智能系统。

3.1.1 辅助诊断

为医生提供疾病诊断建议，包括在当前情况下优先收集患者的哪些临床信息或证据，解释这些证据以及怎样获取这些证据，根据已知证据给出可能疾病的疑似度及其排序，解释各种可能疾病与患者已知病情证据之间的关系，提示与所怀疑疾病无关的异常证据，提示检测完备度等，帮助医生做出疾病诊断，进而生成规范化电子病历。

3.2 疑似度

辅助诊断系统计算所得疾病的疑似程度，用大于0、小于等于1的百分数表示，疑似度越大，该疾病的可能性越大。

3.3 诊断模型

辅助诊断系统所包含的医学知识、数据、以及诊断过程所采用的算法的集合。

3.4 主诉模型

针对一个或几个相关的主诉症状或主诉临床表现构建的诊断模型，该模型所包含的疾病与医院科室无关，每个疾病均可能导致主诉模型中的至少一个主诉症状或主诉临床表现。

3.5 诊断准确率

针对每个主诉模型，由测试方（第三方三甲医院，即未参与诊断模型构建且未提供诊断模型构建所需任何数据的三甲医院）检索具有主诉模型所包含的主诉症状或主诉临床表现的住院且治好出院的病历（以下简称病历），对病历按照主诉模型中所包含的疾病进行分类，在每个疾病分类中进行病历随机抽样，经测试方判断为合格（病历记录的病情信息支持诊断排序第一的疾病）的病历作为测试病历，随机抽样直到测试病历数达到 10 个、或已经没有病历为止，根据测试病历中记载的病情信息，用辅助诊断系统进行疾病诊断，诊断结果按疑似度排序为第一的疾病与病历中记载的排序第一的（含并列第一）疾病相同的，记为诊断准确；不相同的，由测试方的临床专家仔细斟酌，认为辅助诊断系统诊断结果正确的，记为准确，否则记为错误。每个疾病的测试病历中诊断为准确的病历数除以测试病历总数，即为主诉模型中该疾病的诊断准确率：

$$\text{疾病诊断准确率} = \frac{\text{测试病历中诊断为准确的病历数}}{\text{测试病历总数}} \times 100\%$$

每个疾病的诊断准确率之和除以测试疾病总数为主诉模型诊断准确率：

$$\text{主诉模型诊断准确率} = \frac{\text{主诉模型中每个疾病的诊断准确率之和}}{\text{主诉模型中测试疾病总数}} \times 100\%$$

其中未找到测试病历的疾病不参与诊断准确率计算。上述诊断准确率须由测试方盖章确认。

所有主诉模型中诊断准确率最低的为辅助诊断系统的诊断准确率。

同一疾病在不同主诉模型中诊断准确率最低的为该疾病的诊断准确率。

3.6 诊断能力提升率

使用辅助诊断系统（年使用 100 次及以上）的医生群体在使用辅助诊断系统后诊断能力的提升幅度。计算方法为：在相同疾病颗粒度下，医生群体在使用辅助诊断系统后的一年时长内诊断出的疾病数除以该群体在未使用辅助诊断系统的最近一年诊断出的疾病数再减 1：

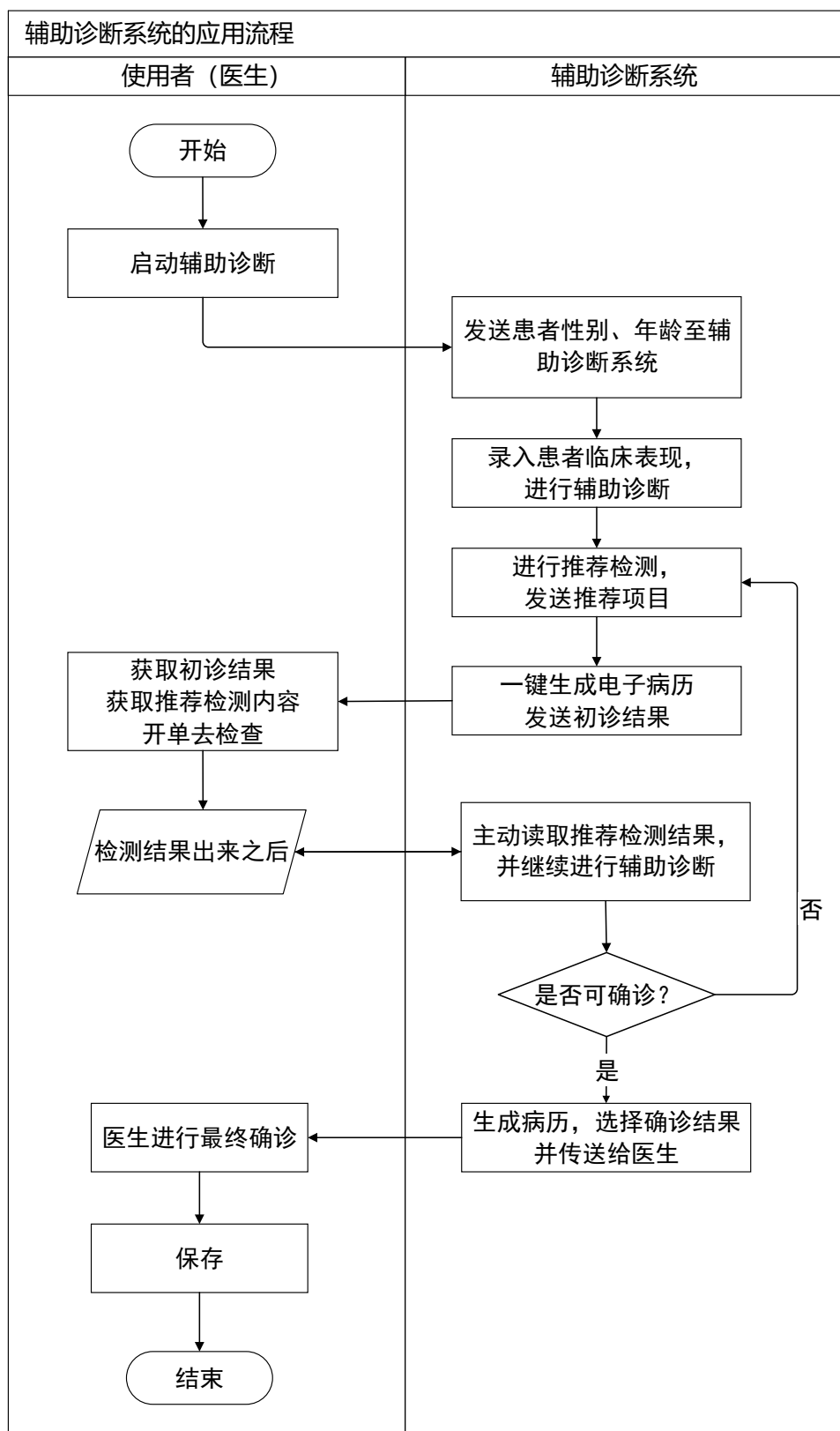
$$\text{诊断能力提升率} = \left(\frac{\text{使用辅助诊断系统（年使用100 次以上）的医生群体一年时长内诊断出的疾病数}}{\text{使用辅助诊断系统（年使用100 次以上）的医生群体在未使用辅助诊断系统的最近一年诊断出的疾病数}} - 1 \right) \times 100\%$$

所谓疾病颗粒度，指划分疾病的粗细程度，通常由 ICD 疾病代码的层级确定。

3.7 检测项

某项检查（例如患者性别、体温）可有不同的结果（例如男或女、体温正常、低热或高热），所有可能的结果构成一个检测项（例如性别为一个检测项、体温为一个检测项）。

4 应用流程



5 性能要求

5.1 诊病范围

覆盖 50 个以上主诉症状或主诉临床表现、1000 个以上疾病、10000 个以上 ICD 疾病代码。

5.2 诊断准确率

诊断准确率不低于 95%，其中对每个疾病的诊断准确率不低于 80%。

5.3 诊断能力提升率

经两年且两个以上县域应用，统计结果显示使用辅助诊断系统的医生群体的诊断能力提升率大于 50%。

5.4 可及时更新

系统每次更新时长不超过 24 小时，更新周期根据实际需要而定，实际需要包括但不限于与权威机构最新发布的临床指南相符（不晚于其发布后的 5 个工作日）。

5.5 具有可解释性

包括诊断结果对医学专业人士可解释、诊断模型对医学专业人士可解释、诊断模型更新对医学专业人士可解释。所谓诊断结果可解释，就是辅助诊断系统做出诊断的理由可被医学专业人士理解、证实或证伪，可解释为什么辅助诊断系统得出了该诊断结论；所谓诊断模型可解释，就是诊断模型中的医学知识是怎么表示的、其中的参数的物理意义、以及诊断过程所采用的算法的物理意义三者均可被医学专业人士理解、证实或证伪；所谓诊断模型更新可解释，就是诊断模型中的知识、数据和算法的任何变更发生在何处、变更的理由、以及变更对诊断模型产生的影响三者均可被医学专业人士理解、证实或证伪。

5.6 具有泛化能力

诊断准确率不依赖数据独立同分布假设，不因应用场景与测试场景不同而降低。

5.7 对数据依赖程度低

诊断准确率对构建诊断模型的数据的数量和数值准确度不敏感。

5.8 提示必检项

所谓必检项，就是为了尽可能避免漏诊，针对主诉症状或主诉临床表现提示必须检测其状态的检测项。

5.9 具备推荐检测功能

推荐下一步应当优先做哪些检测并根据每项检测的推荐度对推荐检测项排序，供医生参考，以便尽可能高效精准地收集患者的临床证据，医生可根据推荐度排序并结合当地实际情况选择做哪些检测。推荐度是一个大于 0 小于 1 的百分数，推荐度越大，检测的综合价值越高；临床证据是检测项的检测结果，包括各种症状、体征、实验室检查、影像学检查、病理检查、影响发病率的风险因素等。

5.10 能表达和处理不确定证据

观测到的证据隶属于检测项的多个结果的，属于每个结果的不确定程度由隶属度（观测值隶属于检测项不同结果的概率值）表示，要求辅助诊断系统能够处理这些不确定证据，诊断计算过程不再产生新的误差。

5.11 提示每次诊断的检测完备度

与当前诊断有关的所有检测项中，结果已知检测项的关注度之和除以结果已知和结果未知检测项的关注度之和为检测完备度：

$$\text{检测完备度} = \frac{\text{结果已知检测项的关注度之和}}{(\text{结果已知检测项的关注度之和}) + (\text{结果未知检测项的关注度之和})} \times 100\%$$

在主诉模型中，关注度是指检测项的检测结果对于诊断引起主诉症状或主诉临床表现的疾病的重要程度，是一个正实数，由医学专家根据经验在主诉模型中给定，其绝对值无意义，仅相对值有意义。

5.12 自动生成自然语言和结构化电子病历

根据录入的临床信息和诊断结果，自动生成符合卫生健康部门规范和患者个体情况的自然语言电子病历和结构化电子病历。其中结构化电子病历用于各种统计分析，亦可直接导入辅助诊断系统再次诊断。

5.13 具有患者病情信息管理功能

可保存患者当前就诊记录、可查看患者的历次就诊记录、可读取患者就诊记录继续录入新的临床证据或修改临床证据进行诊断。

6 安全性要求

6.1 保护隐私

诊断模型的构建和辅助诊断系统的使用过程严格保护个人隐私，不涉及除患者性别、年龄等与诊断有关的信息之外的任何其他个人信息。

6.2 可溯源纠错

辅助诊断系统如果出现诊断错误，可以追溯，知道错在哪里和为什么错，并能确保改正错误后不会再现同样的错误。

6.3 服务器安全

承载辅助诊断系统的服务器符合当地有关政府部门规定的安全要求。

6.4 服务器存放环境安全

辅助诊断系统服务器的存放环境符合当地有关政府部门规定的安全要求。

6.5 网络安全

辅助诊断系统运行的网络环境符合当地有关政府部门规定的安全要求。

6.6 数据安全

辅助诊断系统所产生的数据存储符合当地有关政府部门规定的安全要求。

附录：技术指标汇总

名称	参数值
辅助诊断系统覆盖的主诉症状或主诉临床表现数量	> 50
辅助诊断系统去重后覆盖的疾病种类	> 1000
辅助诊断系统去重后覆盖的 ICD 疾病代码数量	> 10000
辅助诊断系统的诊断准确率	$\geq 95\%$
辅助诊断系统每个疾病的诊断准确率	$\geq 80\%$
两年以上县域试点应用统计结果显示使用辅助诊断系统的医生群体的诊断能力提升率	> 50%
辅助诊断系统每次更新时长	< 24 小时
辅助诊断系统保持与权威机构最新发布的临床指南相符的期限	< 5 个工作日