**临床预测模型Word写作模板**

**一、目录**

1. 题目
2. 摘要（中文+英文）
3. 关键词（中文+英文）
4. 前言  
    研究背景
5. 研究目的

* 研究方法  
   数据  
   数据来源

数据收集时间

* 研究对象
* 数据准备
* 结局指标与预测因素  
   结局指标定义与评估
* 预测因素筛选与定义
* 分析方法  
   数据划分方式
* 样本量估算依据
* 预测因素处理方法
* 模型类型选择与构建步骤
* 参数处理与异质性量化
* 模型性能评估指标与工具
* 类别不平衡处理（选用）
* 模型输出说明
* 训练集与测试集差异分析
* 开放科学相关说明  
   资金声明
* 利益冲突声明
* 研究方案与注册信息
* 数据与代码共享计划
* 结果  
   研究对象特征  
   纳入流程（流程图呈现）
* 基线特征描述（表格呈现）
* 开发集与测试集特征对比
* 模型相关结果  
   模型开发过程
* 模型详细描述
* 模型性能结果
* 讨论  
   主要结果解释
* 研究局限性分析
* 模型临床实用性探讨
* 未来研究方向
* 参考文献
* 附录（选用）  
   详细条目清单
* 补充数据表格
* 模型代码说明
* 补充说明  
   伦理审批信息
* 知情同意情况
* 图表说明

**二、各部分写作详细模板**

**（一）题目**

【模板】基于[具体算法，如：逻辑回归/随机森林/深度学习模型名称]的[目标人群，如：中老年高血压患者]【预测结局，如：脑卒中发生风险】临床预测模型的开发与验证研究

【写作要求】明确体现研究核心要素：模型算法类型、目标人群、预测结局、研究性质（开发/验证/两者兼具），确保文献检索时的高敏感度与精确度。

**（二）摘要（中文+英文）**

**1. 中文摘要**

【模板】

**背景**：[阐述预测结局的临床意义，如：脑卒中是中老年高血压患者的严重并发症，早期诊断对改善预后至关重要]；现有诊断方法存在[不足，如：敏感性低、依赖侵入性检查等]，基于[技术优势，如：机器学习的高维数据处理能力]构建预测模型具有重要价值。目的：开发并验证基于[算法名称]的[目标人群][诊断结局]临床预测模型，明确其诊断效能。

**方法**：以[数据来源，如：2019-2023年某三甲医院住院患者数据/多中心队列数据]为研究对象，按照[划分比例，如：7:3]随机分为开发集（n=XXX）和测试集（n=XXX）。以[金标准，如：临床诊断指南标准/病理检查结果]为结局指标，纳入[主要预测因素类别，如：人口学特征、实验室指标、影像学特征等]作为候选预测因素。通过[数据预处理方法，如：缺失值多重插补、异常值剔除]处理数据后，采用[算法名称]构建模型，运用[评估指标，如：ROC曲线下面积（AUC）、敏感性、特异性、阳性预测值、阴性预测值]评估模型性能。

**结果**：开发集共纳入XXX例研究对象，其中阳性结局XXX例（XX.X%）；测试集纳入XXX例，阳性结局XXX例（XX.X%）。最终模型纳入[具体预测因素，如：年龄、收缩压、糖化血红蛋白等]个预测因素。开发集AUC为XX.XX（95%CI：XX.XX-XX.XX），敏感性XX.X%，特异性XX.X%；测试集AUC为XX.XX（95%CI：XX.XX-XX.XX），敏感性XX.X%，特异性XX.X%。

**结论**：基于[算法名称]构建的[目标人群][预测结局]诊断预测模型具有良好的诊断效能和公平性，可为临床早期诊断提供可靠参考工具。

**2. 英文摘要**

【模板】略

**（三）关键词（中文+英文）**

【中文模板】临床预测模型；[目标疾病，如：脑卒中]；[算法名称，如：随机森林]；预测效能；[核心预测因素，如：血压]（3-8个，核心要素优先）

【英文模板】略

**（四）前言**

**1. 研究背景**

【模板】[目标疾病]是[阐述疾病负担，如：全球范围内高发的心血管疾病，我国患病率达XX.X%，每年新增病例XX万例]，其[早期诊断的重要性，如：早期确诊可使患者5年生存率提高XX%，但临床中约XX%的患者因早期症状不典型而延误诊断]。目前临床常用的[现有诊断方法，如：CT检查、生物标志物检测]存在[具体局限性，如：设备依赖度高、费用昂贵、敏感性仅为XX.X%等]，难以满足[临床需求，如：基层医疗机构筛查、大规模人群早筛]的需求。

临床预测模型通过整合多元预测因素量化疾病发生概率，为疾病早期诊断提供客观依据，已在[相关领域，如：肿瘤、糖尿病]诊断中展现出良好应用前景。然而，现有针对[目标疾病]的临床预测模型多基于[传统方法，如：逻辑回归算法]构建，未充分利用[先进技术优势，如：高维影像学数据、多组学数据]；同时，部分模型存在[报告缺陷，如：预测因素筛选过程不透明、未进行外部验证、未考虑不同人群公平性]等问题，导致模型临床转化困难。

此外，[目标疾病]的发病风险在[不同人群，如：不同性别、地域、年龄组]中存在显著差异，已知[具体健康不平等问题，如：农村地区患者因医疗资源可及性低，诊断延迟率高于城市地区]，但现有模型多未纳入人群差异因素，可能导致特定群体诊断效能下降。基于此，本研究拟整合[数据类型，如：临床常规指标、影像学特征]，采用[算法名称]构建针对[目标人群][目标疾病]的临床预测模型，以期解决现有诊断方法的不足，提高早期诊断准确性。

**2. 研究目的**

【模板】（1）基于[数据来源]开发基于[算法名称]的[目标人群][目标疾病]诊断预测模型，筛选核心诊断预测因素；（2）在[验证数据集来源]中验证模型的诊断效能，包括区分度、校准度及临床实用性；（3）评估模型在[不同亚组，如：不同性别、年龄分层、地域]中的诊断公平性，为模型的临床推广提供依据。

**（五）研究方法**

**1. 数据**

**（1）数据来源**

【模板】开发集数据来源于[具体来源，如：20××年×月-20××年×月××医院住院及门诊就诊患者、某多中心队列研究“XX研究”数据库]；外部验证集数据来源于[具体来源]。明确数据获取的伦理审批情况：[伦理审批编号，如：××伦理委员会审批号XX号]，所有研究对象均签署知情同意书（若为回顾性数据，说明豁免知情同意的审批意见）。

**（2）数据收集时间**

【模板】开发集研究对象招募起始时间：XXXX年XX月，招募终止时间：XXXX年XX月，随访终止时间：XXXX年XX月（若为横断面研究，说明诊断指标与结局指标的评估时间间隔，均在XX小时/天内完成）；外部验证集招募起始时间：XXXX年XX月，招募终止时间：XXXX年XX月，随访终止时间：XXXX年XX月。详细记录数据收集的具体时间节点，以确保诊断预测因素与结局指标的时间关联性（诊断预测模型本质为横断面关系，需保证两者评估时间相同或相近）。

**（3）研究对象**

【模板】纳入标准：（1）[人口学特征，如：年龄≥18岁，性别不限]；（2）[疾病相关，如：因疑似[目标疾病]就诊，接受过[金标准检查]；（3）[数据相关，如：具有完整的[核心预测因素，如：临床症状、实验室指标、影像学数据]及结局指标信息；（4）[其他，如：自愿参与本研究并配合数据收集]。

排除标准：（1）[疾病相关，如：合并[严重基础疾病，如：终末期肾病、恶性肿瘤晚期]可能影响结局判断；（2）[数据相关，如：关键预测因素缺失率超过XX%、存在明显数据录入错误且无法溯源修正；（3）[其他，如：研究期间退出、拒绝提供后续随访信息]。

**（4）数据准备**

【模板】采用[软件名称，如：R 4.3.1软件、SPSS 26.0软件]进行数据整理与预处理。缺失值处理：首先统计各预测因素的缺失率，对于缺失率<XX%的变量，采用[方法，如：多重插补法（m=5次）、均值/中位数填补（针对连续变量）、众数填补（针对分类变量）]；缺失率≥XX%的变量予以剔除。异常值处理：通过[方法，如：箱线图（四分位距法）、Z值法（|Z|>3.29）]识别异常值，经临床专家审核确认后，[处理方式，如：修正录入错误、视为缺失值处理]。

数据质量检查：制定统一的数据核查标准，由2名经过培训的研究人员独立进行数据核查，包括[核查内容，如：逻辑一致性检查（如年龄与出生日期匹配性）、范围合理性检查（如血压值在合理区间内）]，不一致之处通过查阅原始病历进行核实修正，一致率达到XX%以上。同时，比较不同社会人口群体（如不同性别、年龄组、地域）的数据预处理过程，确保处理方法的一致性，减少偏倚。

**2. 结局指标与预测因素**

**（1）结局指标定义与评估**

【模板】以[金标准，如：《XX疾病诊断指南（XXXX年版）》中的临床诊断标准、病理检查结果、影像学确诊结果]作为结局指标，定义[阳性结局，如：确诊为[目标疾病]，阴性结局：排除[目标疾病]诊断]。结局指标的评估时间范围：研究对象首次就诊后XX小时内完成金标准评估。

评估方法一致性保障：对于需主观解释的结局指标（如影像学结果），由2名具有[资质，如：5年以上相关临床工作经验、副主任医师及以上职称]的专家独立评估，采用[一致性检验方法，如：Kappa检验]评估一致性，Kappa值≥XX表明一致性良好；若存在分歧，通过第三方专家会诊达成共识。

**（2）预测因素筛选与定义**

【模板】参考[依据，如：相关临床指南、既往研究文献、专家共识]，初步纳入候选预测因素，分为以下几类：（1）人口学特征：如年龄（连续变量，岁）、性别（分类变量，男/女）、地域（分类变量，城市/农村）、文化程度（分类变量，小学及以下/初中/高中及以上）；（2）临床症状与体征：如[具体症状，如：头痛（有/无）、血压（连续变量，mmHg）、心率（连续变量，次/分）]；（3）实验室指标：如[具体指标，如：血常规中的白细胞计数（连续变量，×10⁹/L）、生化指标中的肌酐（连续变量，μmol/L）]，检测时间为就诊后XX小时内，检测仪器为[仪器型号，如：贝克曼AU5800全自动生化分析仪]；（4）影像学特征：如[具体特征，如：CT影像中的病灶大小（连续变量，mm）、病灶密度（分类变量，高密度/等密度/低密度）]，由专业影像科医师提取；（5）其他：如[具体因素，如：吸烟史（有/无）、家族病史（有/无）]。

对所有预测因素进行明确操作定义，例如“吸烟史”定义为：每天吸烟≥1支，连续吸烟≥6个月，或既往吸烟达到上述标准但已戒烟<1年；“高血压”定义为：收缩压≥140mmHg和（或）舒张压≥90mmHg，或既往确诊高血压并规律服药。对于需主观判断的预测因素（如临床症状严重程度），制定标准化评估量表，由培训后的评估人员进行评分，确保评估一致性。

**3. 分析方法**

**（1）数据划分方式**

【模板】采用[随机划分方法，如：简单随机抽样、分层随机抽样（按结局事件发生情况分层）]将开发集数据按[比例，如：7:3]划分为训练集（用于模型构建）和内部验证集（用于模型初步优化）；外部验证集数据不参与模型构建，仅用于模型外部有效性验证。数据划分过程采用[软件/方法，如：R软件caret包、Python sklearn库]实现，设置随机种子（如seed=1234）以保证结果可重复。明确说明数据划分时是否考虑聚类效应（如多中心数据按中心分层），若考虑，采用[方法，如：按中心进行分层随机抽样]。

**（2）样本量估算依据**

【模板】参考诊断预测模型样本量估算指南，基于以下参数进行估算：（1）预计结局事件发生率P=XX.X%（根据预实验/既往文献获得）；（2）候选预测因素数量k=XX个；（3）事件数与预测因素数比例至少为10:1（或20:1）；（4）α=0.05，β=0.2（把握度80%）。通过[公式/软件，如：nQuery Advisor 7.0软件、PASS 15.0软件]计算得出，开发集至少需要XX例样本，其中阳性结局XX例；外部验证集样本量至少为开发集的XX%（或XX例），以保证验证效能。本研究实际纳入开发集XX例（阳性XX例）、外部验证集XX例（阳性XX例），满足样本量要求。

**（3）预测因素处理方法**

【模板】连续变量：首先进行正态性检验（如Shapiro-Wilk检验），符合正态分布的变量以均数±标准差描述，组间比较采用t检验；不符合正态分布的变量以中位数（四分位数）描述，组间比较采用Wilcoxon秩和检验。分类变量以频数（百分比）描述，组间比较采用χ²检验或Fisher确切概率法（当期望频数<5时）。

预测因素筛选：采用[筛选方法，如：单因素分析结合多因素逐步回归法（向前/向后/逐步法）、LASSO回归（用于高维数据降维）、随机森林变量重要性评分、专家咨询法（针对临床意义明确的因素）]进行预测因素筛选。单因素分析中P<XX（如0.10）的变量纳入多因素筛选；LASSO回归通过交叉验证确定最优λ值，剔除系数为0的变量；随机森林通过变量重要性评分（如Gini系数、袋外误差减少量）筛选评分前XX%的变量；最终结合临床意义，由[XX名专家，如：3名副主任医师及以上职称的临床专家]审核确定纳入模型的最终预测因素。

**（4）模型类型选择与构建步骤**

【模板】根据数据特征与研究需求，选择[算法名称，如：逻辑回归（LR）、随机森林（RF）、支持向量机（SVM）、人工神经网络（ANN）、XGBoost]构建诊断预测模型。以XGBoost模型为例，构建步骤如下：（1）将筛选后的预测因素导入模型，设置初始参数（如学习率η=0.1，最大树深度max\_depth=6，最小样本权重和min\_child\_weight=1，子样本比例subsample=0.8，列采样比例colsample\_bytree=0.8）；（2）采用[交叉验证方法，如：10折交叉验证]对模型参数进行优化，以[目标指标，如：验证集AUC最大]为目标确定最优参数组合；（3）基于最优参数构建最终XGBoost模型；（4）同时构建传统逻辑回归模型作为对照，比较不同模型的诊断效能。

详细描述模型的基本原理：如XGBoost是一种基于梯度提升决策树（GBDT）的集成学习算法，通过迭代构建多棵决策树，利用梯度下降法最小化损失函数（如对数损失函数），从而提高模型的预测精度。明确模型构建所使用的软件与包，如R软件xgboost包（版本1.7.5.1）、Python sklearn库（版本1.2.2）。

**（5）****模型性能评估指标与工具**

【模板】模型性能从区分度、校准度、临床实用性三个维度进行评估：（1）区分度：即模型区分阳性与阴性结局的能力，主要指标包括AUC（ROC曲线下面积）、敏感性、特异性、阳性预测值（PPV）、阴性预测值（NPV），AUC≥0.8表明区分度良好；绘制ROC曲线，通过Youden指数（敏感性+特异性-1）确定最佳截断值。（2）校准度：即模型预测概率与实际发生概率的一致性，采用[方法，如：校准曲线（Calibration curve）、Hosmer-Lemeshow检验（针对逻辑回归）、Brier评分（越小越好）]评估；校准曲线通过将预测概率分为XX组（如10组），计算每组实际发生率与预测概率的均值，绘制拟合曲线并计算校准优度。（3）临床实用性：采用[方法，如：决策曲线分析（DCA）、临床影响曲线（CIC）]评估；DCA曲线横坐标为阈值概率，纵坐标为净获益，通过比较模型与“全部治疗”“全部不治疗”策略的净获益，确定模型的临床应用价值；CIC曲线展示在不同阈值概率下，被正确诊断的阳性患者数与被误诊的阴性患者数。

所有评估指标均同时在训练集、内部验证集、外部验证集中计算，以全面评估模型的稳定性与泛化能力。评估工具采用[软件/包，如：R软件pROC包（AUC计算）、calibrate包（校准曲线）、rmda包（DCA分析）]。

**（6）类别不平衡处理（选用）**

【模板】本研究开发集中阳性结局发生率为XX.X%（<XX%，如30%），存在类别不平衡问题。采用[处理方法，如：过采样法（SMOTE算法，合成少数类样本）、欠采样法（随机剔除多数类样本，使两类样本比例达到1:1）、加权法（为少数类样本分配更高的权重，如权重=多数类样本量/少数类样本量）、集成方法（EasyEnsemble算法）]处理类别不平衡。详细描述处理过程：如采用SMOTE算法时，设置近邻数k=5，合成新样本后阳性与阴性样本比例为1:1；处理后重新构建模型，并比较处理前后模型的AUC、敏感性、特异性等指标变化，评估处理效果。同时说明处理方法对样本量的影响，如欠采样后训练集样本量从XXX例减少至XXX例。

**4. 模型输出说明**

【模板】模型最终输出为[输出类型，如：疾病诊断的概率值（范围0-1）、分类结果（阳性/阴性）]。若为概率值，需明确说明概率对应的临床意义，即“预测概率≥[最佳截断值，如：0.35]时，判定为疾病诊断阳性，提示存在[目标疾病]患病风险；预测概率<[最佳截断值]时，判定为诊断阴性”，其中最佳截断值来源于开发集ROC曲线的Youden指数最大值。若为分类结果，需同步报告该分类对应的预测概率范围及判断依据。

模型输出的呈现形式包括：（1）可视化工具：如列线图（Nomogram），通过[软件名称，如：R软件rms包、Stata软件nomogram命令]绘制，将各预测因素的得分与总得分、对应预测概率相关联，实现“输入预测因素→查找得分→计算总得分→对应预测概率”的可视化操作；（2）数字化工具：如基于[开发工具，如：Python Flask框架、R Shiny包]构建的网页版计算器或手机端小程序，用户可直接输入预测因素数值（如年龄、血压值），系统自动计算并输出预测结果及置信区间；（3）统计表格：在研究报告中呈现典型案例的预测结果，包括各预测因素具体数值、模型输出概率、实际诊断结局及一致性判断。

此外，需明确模型输出的使用步骤：（1）收集待评估对象的[具体预测因素，如：年龄、性别、收缩压、白细胞计数]等信息，确保信息完整且符合预测因素的操作定义；（2）通过[呈现形式，如：列线图查分、计算器输入]获取预测结果；（3）结合临床实际情况（如患者基础疾病、症状体征）对预测结果进行综合判断，避免单纯依赖模型输出做出诊断决策。