

临床预测模型构建中的常用样本量估算方法介绍及实例分析*

卜志军¹ 丁申² 马文欣¹ 刘胜亿³ 鲁志林³ 朱静⁴ 全有翠⁵ 邹沂榭⁶
郝宇涵¹ 张应琳⁵ 刘欣欣⁷ 肖天琪⁴ 刘建平¹ 刘兆兰^{1#}

(1 北京中医药大学中医学院循证医学中心 北京 100029; 2 湖北中医药大学针灸骨伤学院;
3 湖北中医药大学第一临床学院; 4 湖北中医药大学中医临床学院; 5 北京中医药大学第二临床学院;
6 湖北中医药大学检验学院; 7 湖北中医药大学人文学院)

摘要: 临床预测模型是指利用多因素模型对患有某病或者未来某结局发生的概率进行预测。近年来, 中医药领域预测模型的相关研究逐渐兴起, 构建中医临床预测模型可以辅助中医师进行临床诊断和循证决策。恰当的样本量对于提高临床预测模型的精准度和可信度至关重要。样本量不足可能导致无法得到具有统计意义的结论, 而过大的样本量则可能造成资源浪费。本研究分析了包括每个变量至少 10 个事件经验法(10 events per variable, 10EPV)、每个参数至少 10 个事件(10 events per parameter, 10EPP)法则和四步法在内的临床预测模型样本量估算方法, 同时通过实例进行展示。本研究旨在为临床预测模型的样本量估算提供实用参考, 以促进中医药领域临床预测模型的相关研究。

关键词: 临床预测模型; 样本量估算; 中医学

中图分类号: R2-03; R311

doi: 10. 3969/j. issn. 2095-6606. 2024. 06. 008

Introduction and example analysis of common sample size estimation methods in clinical prediction model construction*

BU Zhijun¹, DING Shen², MA Wenxin¹, LIU Shengyi³, LU Zhilin³, ZHU Jing⁴, QUAN Youcui⁵,
ZOU Yiman⁶, HAO Yuhan¹, ZHANG Yinglin⁵, LIU Xinxin⁷, XIAO Tianqi⁴, LIU Jianping¹, LIU Zhaolan^{1#}

(1 Centre for Evidence-Based Medicine, School of Traditional Chinese Medicine, Beijing University of Chinese Medicine, Beijing 100029; 2 School of Acupuncture and Orthopaedic Traumatology, Hubei University of Chinese Medicine; 3 First Clinical School of Hubei University of Chinese Medicine; 4 Clinical School of Traditional Chinese Medicine, Hubei University of Chinese Medicine; 5 Second Clinical School of Beijing University of Chinese Medicine; 6 School of Laboratory Medicine, Hubei University of Chinese Medicine; 7 School of Humanities, Hubei University of Chinese Medicine)

Abstract: Clinical prediction models are utilized to forecast the likelihood of disease or future outcomes using multifactor models. In recent years, there has been a growing body of research on prediction models in the field of traditional Chinese medicine (TCM). The development of clinical prediction models in TCM can support practitioners in clinical diagnosis and evidence-based decision-making. Adequate sample size plays a pivotal role in enhancing the accuracy and reliability of clinical prediction models. Insufficient sample size may yield statistically insignificant findings, while an excessively large sample size may result in resource wastage. This study critically examines and deliberates on various method for

卜志军, 男, 在读硕士生

通信作者: 刘兆兰, 女, 博士, 教授, E-mail: lz1019@163.com

* 国家自然科学基金项目 (No. 82374298); 北京中医药大学学科后备带头人支持计划 (No. 90010960920033); 国家中医药管理局高水平中医药重点学科建设项目 (No. zyyzdxk-2023249)

estimating sample size in clinical prediction models, encompassing the 10 Events Per Variable (10EPV) principle, the 10 Events Per Parameter (10EPP) rule, and the Four Steps approach, with practical illustrations provided. The primary objective of this study is to furnish practical guidelines for sample size estimation in clinical prediction models and foster related research in the realm of TCM.

Keywords: clinical prediction models; sample size estimation; traditional Chinese medicine

临床预测模型,又称为临床预测规则或风险评分,是指利用多因素分析预测疾病发生或特定结局概率的有力工具^[1]。临床预测模型不仅体现了现代医学的预防和个性化治疗理念,而且与中医学“治未病”的传统理论相契合^[2-3]。构建中医临床预测模型对于提高临床诊断的准确性、推动中医现代化具有重要意义^[4]。样本量估算是中医临床预测模型研究设计中的重要一环,可以有效避免资源浪费并防止模型过拟合^[5-6]。然而目前缺乏关于中医临床预测模型样本量估算方面的相关研究,根据2024年2月知网数据库检索显示,已发表的42篇中医药预测模型研究中,仅有4篇明确提及了样本量估算方法^[7-10]。基于此,本研究旨在深入讨论临床预测模型样本量估算方法,从而为临床预测模型的样本量估算提供实用参考,以促进中医药领域临床预测模型的相关研究。

1 每个变量至少10个事件经验法

每个变量至少10个事件(10 events per variable, 10EPV)经验法指的是,在构建预测模型时,对于每个预测变量,应至少有10个阳性结局事件(即对于回归方程中每个非截距项 β 对应10个阳性结果,其中阳性结果在总数中所占比例应小于阴性结果)。这种方法是基于经验得出的规则,用于确保模型的稳定性和预测能力^[11-12]。在本研究中,我们基于课题组前期关于中成药五味苦参肠溶胶囊治疗活动期溃疡性结肠炎的研究,进行了样本量的估算^[13]。该研究共纳入274例活动性溃疡性结肠炎患者,其中205例患者的临床数据用于构建模型,其余69例用于模型的验证。通过LASSO回归分析,本研究纳入了8个预测变量,包括红细胞沉降率、年龄、疾病类型、微小出血、脓桥、溃疡、疾病部位和疼痛,且以疾病缓解率作为主要的结局指标。在这些变量中,红细胞沉降率和年龄作为连续性指标,每个指标对应1个 β 系数。疾病类型和疼痛作为三分类指标,每个指标对应2个 β 系数,而疾病部位作为四分类指标,对应3个 β 系数。脓桥和脓液作为二分类指标,每个指标对应1个 β 系数。根据10EPV经验法模型中共涉及12个 β 系数,因此需

要至少120个阳性结局事件。在本研究的205例训练样本中,共有92个阳性结局事件,小于理想120个阳性结局事件,故提示样本量不充足。

2 每个参数至少10个事件法则

每个参数至少10个事件(10 events per parameter, 10EPP)法则是对10EPV经验法的一种补充和精细化,它要求在模型开发阶段,对于所有考虑的候选预测变量,每个参数应该至少对应一定数量的事件(通常事件数为10个),以确保模型的稳定性和可靠性^[14]。候选预测变量指的是在模型开发初期可能被考虑包括进模型的所有变量,而不仅仅是最终选定构建模型的那些变量^[15]。这是因为即使是在模型选择过程中被排除的变量,也可能影响模型的复杂度和最终的过拟合风险。在进行变量选择或优化模型时,研究人员可能会测试许多不同的变量组合,这些都应当在估算所需样本量时考虑在内。根据10EPP法则针对课题组前期研究进行样本量估算举例,该研究共收集17个候选预测变量^[13],除10EPV经验法提及最终纳入LASSO回归的8个预测变量外,还包括性别、是否前期治疗过、是否镜下糜烂、是否溃疡共4个二分类候选预测变量(每1个二分类候选预测变量对应1个 β),疾病的程度、腹泻次数、大便次数共3个三分类候选预测变量(每1个三分类候选预测变量对应2个 β),白蛋白以及病程2个连续性候选预测变量(每一个连续性候选预测变量对应1个 β),共12个 β 。结合10EPV经验法中的12个 β ,共计24个 β , $24 \times 10 = 240$ 。而训练集中的实际样本量为205名,小于理想样本量240,故提示样本量不充足。

3 四步法估算样本量

四步法估算样本量方法是由Riley等^[16]提出,共包括准确估计人群的平均期望风险/值,使得个体风险估计值的误差越小越好,最小化过度拟合所需的样本量,达到预期性能4个步骤。每一个步骤均可计算出样本量,最终选择四步中最大的样本量作为预测模型的理想样本量。本文以前期研究中的数据为例^[13],进行每一步的样本量估算举例。

3.1 准确估计人群的平均期望风险/值

样本量大小必须满足预测模型的截距能够被精确估计,从而确保所开发的模型可准确预测结果的总体比例。方法是在不包含预测变量的情况下,精确估计模型中的截距所需的样本量。对于二进制结果的估算,绝对误差通常设置为 0.05。

对于二分类结局, n 为样本量大小, ϕ 为结局事件比例,则 95%置信区间为^[17]:

$$95\% \text{ 置信区间} = \hat{\phi} \pm 1.96 \sqrt{\frac{\hat{\phi}(1 - \hat{\phi})}{n}}$$

绝对误差范围(δ)为:

$$\delta = 1.96 \sqrt{\frac{\hat{\phi}(1 - \hat{\phi})}{n}}$$

样本量估算公式为:

$$n = \left(\frac{1.96}{\delta}\right)^2 \hat{\phi}(1 - \hat{\phi})$$

由于前期研究^[13]中患者疗效好转事件发生阳性率为 0.448 78,在绝对误差为 0.05 的前提下,置信区间限制到 0.397 到 0.497,代入上述公式该研究所需要的样本量至少为 381 例。

$$n = \exp\left(\frac{-0.508 + 0.259\ln(0.44878) + 0.504\ln(8) - \ln(0.05)}{0.544}\right)$$
$$n = 454.007$$

3.3 最小化过度拟合所需的样本量

过度拟合是指模型在训练数据上表现较好,以至于捕捉到数据中的噪声和具体细节而非潜在的模式,导致模型在新数据集上表现不佳。过度拟合的模型常常缺乏泛化能力。收缩法(也称为惩罚或正则化)可通过减少开发模型的预测可变性来处理过拟合的问题,即减少极端预测情况(如预测概率接近 0 或 1)的发生^[18]。但收缩法并不能完全克服过拟合问题,因为收缩法是根据开发数据集进行估计,特别是当样本量很小时,往往是不精确的,故无法在特定应用中正确处理过拟合的大小^[19]。此外,所需的估计收缩率与模型的表现性能之间往往会出现负相关,所以无法通过选择合适的收缩率来消除过拟合的影响。因此,在模型开发中需要在较小收缩率($\leq 10\%$,即预期收缩因子 $S \geq 0.9$)下确定相对应的样本量和候选预测变量的数量。对于二元结果的样本量估算,要求研究人员分别在目标人群中指定预期结果的比例或比率^[17]。 R_{cs}^2 (Cox-Snell R^2)为评价模型性能的保守指标^[16],它反映模型信噪比,可影响多参数估计和模型潜在过拟合。二分类结局 R_{cs}^2

$$n = \left(\frac{1.96}{0.05}\right)^2 0.44878(1 - 0.44878) = 380.128$$

3.2 使得个体风险估计值的误差越小越好

样本量大小还需满足对预测值范围进行精确预测的要求。VanSmeden 等认为预测变量的数量、总样本的大小以及结局事件的比例是模型平均预测精度的 3 个主要因素^[17]。平均绝对预测误差(MAPE)的公式为:

$$\ln(\text{MAPE}) = -0.508 - 0.544\ln(n) + 0.259\ln(\phi) + 0.504\ln(P)$$

n 为样本量大小, ϕ 为预期结局事件比例(≤ 0.5), P 为预测变量的数量。由此得出样本量估算公式为:

$$n = \exp\left(\frac{-0.508 + 0.259\ln(\phi) + 0.504\ln(P) - \ln(\text{MAPE})}{0.544}\right)$$

基于 R 语言 Lasso 回归,前期研究中数据筛选出来的预测变量数为 8^[13],假设实现观测和真实结局概率之间的平均绝对误差为 0.05。代入上述公式,样本量至少为 455 人。二分类结局样本量公式如下:

范围是 $0 \sim \max(R_{cs}^2)$ 。 R_{cs}^2 需要根据外部证据和专家意见来告知其指定数值。

二分类结局样本量公式如下:

$$n = \frac{P}{(S - 1)\ln\left(1 - \frac{R_{cs}^2}{S}\right)}$$

在结局事件发生的比例是 0.448 78,使用 R 语言 Lasso 回归筛选出来的最终预测变量数为 8, $R_{cs}^2 = 0.2$,预期收缩率 = 0.1, $S = 1 - \text{预期收缩率}(10\%) = 0.9$ 的前提下,代入下列公式,可得中成药五味苦参肠溶胶囊治疗活动期溃疡性结肠炎样本量至少为 319 人。

$$n = \frac{8}{(0.9 - 1)\ln\left(1 - \frac{0.2}{0.9}\right)} = 318.326$$

3.4 预期达到的性能

预期达到的性能目的是保证开发模型和 R_{cs}^2 的乐观优化调整值之间差异小^[20],可用公式 $R^2\text{Nagelkerke} = R_{cs}^2 / \max(R_{cs}^2)$ 对 R_{cs}^2 进行限制,得出 $R^2\text{Nagelkerke}$ 值。乐观调整的 $R^2\text{Nagelkerke}$ 值是

模型在目标人群中拟合的更现实(近似无偏)估计。对于二分类结局,建议 $\delta \leq 0.05$, 二分类结局样本量公式如下:

$$R^2 \text{ Nagelkehe} = \frac{R_{cs}^2}{\max(R_{cs}^2)}$$

$$\max(R_{cs}^2) = 1 - \exp\left(\frac{2\ln L_{null}}{n}\right)$$

$$\ln L_{null} = E \ln\left(\frac{E}{n}\right) + (n - E) \ln\left(1 - \frac{E}{n}\right)$$

$$S = \frac{R_{cs}^2}{R_{cs}^2 + \delta \max(R_{cs}^2)}$$

$$n = \frac{P}{(S - 1) \ln\left(1 - \frac{R_{cs}^2}{S}\right)}$$

在结局事件发生的比例为 0.448 78, 使用 R 语言 Lasso 回归筛选出来的最终预测变量数为 8, 中成药五味苦参肠溶胶囊治疗活动期溃疡性结肠炎训练集中的阳性事件 $E = 92$, n_1 为训练集实际样本量 205 的情况下, 代入 $\ln L_{null}$ 的公式可得, $\ln L_{null} = 141.017 673 1$, $\max(R_{cs}^2) = 0.747$ 。 R_{cs}^2 的取值范围在 (0, 0.747), 假设 R_{cs}^2 取值为 0.2, 乐观 R2Nagelkerke 的预期乐观度仅为 0.05, 代入公式可得, $S = 0.842 6$, $n = 187.57$ 。即样本量需要 188 人。

$$\ln L_{null} = E \ln\left(\frac{E}{n}\right) + (n - E) \ln\left(1 - \frac{E}{n}\right) =$$

$$92 \ln\left(\frac{92}{205}\right) + (205 - 92) \ln\left(1 - \frac{92}{205}\right) = 141.0176731$$

$$\max(R_{cs}^2) = 1 - \exp\left(\frac{2\ln L_{null}}{n}\right) =$$

$$1 - \exp\left(\frac{2 \times (141.0176731)}{205}\right) = 0.747$$

$$S = \frac{R_{cs}^2}{R_{cs}^2 + \delta \max(R_{cs}^2)} = \frac{0.2}{0.2 + 0.05 \times 0.747} = 0.8426$$

$$n = \frac{P}{(S - 1) \ln\left(1 - \frac{R_{cs}^2}{S}\right)} =$$

$$\frac{8}{(0.8426 - 1) \ln\left(1 - \frac{0.2}{0.8426}\right)} = 187.57$$

根据以上 4 步, $n = \text{Max}(Nstep1, 2, 3, 4)$, 得出理想样本量为 454 人。而训练集中的实际样本量为 205 名, 小于理想样本量 454 人, 因此, 提示样本量不充足。

4 R 语言 pmsampsize 包实现样本量估算

基于上述四步法公式, R 语言 pmsampsize 包可以实现样本量估算。操作步骤如下, 首先安装 pmsampsize 包, 然后加载 pmsampsize 包。由于前期课题组研究中的结局指标为二分类结局(疗效是否好转)^[13], 因此, 在代码中输入种类为 b (type = b), 调整 R_{cs}^2 设定为 0.2 (csrsquare = 0.2), 预测变量数目为 8 (parameters = 8), 事件发生率为 0.448 78 (prevalence = 0.448 78), R 语言 pmsampsize 包估算样本量结果如图 1A 所示, 其中, Criteria 1 对应的是四步法中第三步(样本量 $n = 319$), Criteria 2 对应的是四步法中第四步(样本量 $n = 188$), Criteria 3 对应的是四步法中第一步(样本量 $n = 381$)。而四步法中的第二步则需要使用公式或者 shiny 包实现在线计算, shiny 包的在线网站是: <https://mvansmeden.shinyapps.io/BeyondEPV/>。在该网站上需要输入预测变量的数量、结局事件发生率, 以及相对平均预测平方误差 (rMPSE) 的数值, 当 rMPSE 输入 0.07, MAPE 数值与步骤 2 中公式的 0.05 保持一致。当预测变量数量输入为 8、结局事件发生率输入为 0.448 78, rMPSE 输入 0.07 时, 得出结果为 440 人。对比公式结果, 发现公式计算与网页计算器计算存在一定的差异, 可能是因为数值计算的精度、舍入规则等差异导致公式计算所得出的样本量更大。

此外, R 语言 pmsampsize 包还提供了关于连续性结局指标以及生存数据类型结局事件样本量计算方法。针对于连续性结局, 输入种类为 c (type = c), 假设 R_{cs}^2 设定为 0.2 (rsquared = 0.2), 预测变量数目为 8 (parameters = 8), 总体结局的平均值为 1.9 (intercept = 1.9), 总体结局的标准差为 0.6 (sd = 0.6)。可得出如图 1B 所示的结果, 即共需 242 例样本。针对于生存数据类型结局指标, 输入种类为 s (type = s), 假设调整后的 R_{cs}^2 设定为 0.2 (csrsquare = 0.2), 预测变量数目为 8 (parameters = 8), 结局事件发生率为 0.065 (rate = 0.065), 进行预测的关键时间点为 2 个月 (timepoint = 2), 预期的平均随访时长为 2.07 个月 (meanfup = 2.07), 预测的关键时间点与预期的平均随访时长的单位要保持一致, 结果表明至少需要 319 例样本(如图 1C)。

5 临床预测模型构建中的常用样本量估算方法对比

表 1 对于临床预测模型构建常用的样本量估算方法进行了对比, 给出了样本量估算公式与特点对比。

```
(A)
> pmsampsize(type = "b", csrsquared = 0.2, parameters = 8, prevalence = 0.44878)
NB: Assuming 0.05 acceptable difference in apparent & adjusted R-squared
NB: Assuming 0.05 margin of error in estimation of intercept
NB: Events per Predictor Parameter (EPP) assumes prevalence = 0.44878

Samp_size Shrinkage Parameter CS_Rsq Max_Rsq Nag_Rsq EPP
Criteria 1 319 0.900 8 0.2 0.747 0.268 17.90
Criteria 2 188 0.843 8 0.2 0.747 0.268 10.55
Criteria 3 381 0.900 8 0.2 0.747 0.268 21.37
Final 381 0.900 8 0.2 0.747 0.268 21.37

(B)
> pmsampsize(type = "c", rsquared = 0.2, parameters = 8, intercept = 1.9, sd = 0.6)
NB: Assuming 0.05 acceptable difference in apparent & adjusted R-squared
NB: Assuming MMOE <= 1.1 in estimation of intercept & residual standard deviation
SPP - Subjects per Predictor Parameter

Samp_size Shrinkage Parameter Rsq SPP
Criteria 1 233 0.900 8 0.2 29.12
Criteria 2 129 0.838 8 0.2 16.12
Criteria 3 242 0.903 8 0.2 30.25
Criteria 4* 242 0.903 8 0.2 30.25
Final 242 0.903 8 0.2 30.25

Minimum sample size required for new model development based on user inputs = 242
* 95% CI for intercept = (1.83, 1.97), for sample size n = 242

(C)
> pmsampsize(type = "s", csrsquared = 0.2, parameters = 8, rate = 0.065,
+ timepoint = 2, meanfup = 2.07)
NB: Assuming 0.05 acceptable difference in apparent & adjusted R-squared
NB: Assuming 0.05 margin of error in estimation of overall risk at time point = 2
NB: Events per Predictor Parameter (EPP) assumes overall event rate = 0.065

Samp_size Shrinkage Parameter CS_Rsq Max_Rsq Nag_Rsq EPP
Criteria 1 319 0.900 8 0.2 0.555 0.361 5.37
Criteria 2 255 0.878 8 0.2 0.555 0.361 4.29
Criteria 3 * 319 0.900 8 0.2 0.555 0.361 5.37
Final SS 319 0.900 8 0.2 0.555 0.361 5.37

Minimum sample size required for new model development based on user inputs = 319,
corresponding to 660.3 person-time** of follow-up, with 43 outcome events
assuming an overall event rate = 0.065 and therefore an EPP = 5.37

* 95% CI for overall risk = (0.087, 0.155), for true value of 0.122 and sample size n
= 319
**where time is in the units mean follow-up time was specified in
```

图 1 R 语言 pmsampsize 包针对不同结局类型实现样本量估算结果

(A) 二分类样本量估算结果, (B) 连续性样本量估算结果, (C) 生存时间样本量估算结果。

表 1 临床预测模型样本量估算方法的对比

方法名称	公式	特点
“10EPV”经验法	连续性最终纳入模型中的变量对应一个 β , 分类型最终纳入模型中的变量对应的是分类数目减去 1 个 β , β 数目相加后 $\times 10$ 得出总的样本量	“10EPV”经验法简便, 但因忽视复杂性而过于保守
“10EPP”法则	与 10EPV 计算方法相似, 但变量选择的是候选变量	是对 10EPV 经验法的一种补充和精细化, 要求在模型开发阶段考虑所有的候选预测变量
Step1. 准确估计人群的平均期望风险/值	$n = \left(\frac{1.96}{\delta}\right)^2 \phi(1 - \phi)$	实现准确估计结局整体风险或平均结果值需要的样本量
Step2. 使得个体风险估计值的误差越小越好	$n = \exp\left(\frac{-0.508 + 0.259\ln(\phi) + 0.504\ln(P) - \ln(MAPE)}{0.544}\right)$	使所有个体的预测值具有较小平均误差需要的样本量
Step3. 最小化过度拟合所需的样本量	$n = \frac{P}{(S - 1)\ln\left(1 - \frac{R_{cs}^2}{S}\right)}$	在开发模型时注意在较小收缩(又称惩罚或正则化)率(≤ 0.1 , 即预期收缩因子 $S \geq 0.9$)下确定样本量和预测变量的数量
Step4. 预期达到的性能(如回归任务的 R 方等)	$R^2_{Nagelke} = \frac{R_{cs}^2}{\max(R_{cs}^2)}, \max(R_{cs}^2) = 1 - \exp\left(\frac{2\ln L_{null}}{n}\right)$ $\ln L_{null} = E\ln\left(\frac{E}{n}\right) + (n - E)\ln\left(1 - \frac{E}{n}\right), S = \frac{R_{cs}^2}{R_{cs}^2 + \delta\max(R_{cs}^2)}$ $n = \frac{P}{(S - 1)\ln\left(1 - \frac{R_{cs}^2}{S}\right)}$	保证开发模型和 R_{cs}^2 的乐观优化调整值之间差异较小所需要的样本量

6 讨论

近年来,随着建模技术的更新以及多学科交叉融合的发展,中医药领域内开展了较多中医临床预测模型的相关研究。构建中医临床预测模型对于实现个性化治疗、提高诊疗准确性和效率以及促进中医药循证医学基础具有重要意义。作为临床试验设计一个极为重要的环节,样本量估算直接关系到研究结论的可靠性、可重复性以及研究效率的高低。

然而中医药临床预测模型研究中涉及样本量估算的研究较少。目前预测模型的样本量计算方法主要包括10EPV经验法、10EPP法则以及四步法。本文基于课题组前期研究中的数据对上述方法进行实例验证,结果显示,基于同样的数据,使用四步法所需要样本量最大。四步法综合考虑了准确估计总体平均风险、个体风险预测的精确度、最小化过拟合的样本量需求,以及达到预定性能指标的各个方面,通过这4个维度逐一估算样本量,并采用这些步骤中最大的数值作为最终估算结果。相比于10EPV经验法和10EPP法则,四步法通过更多的维度以确保研究的可靠性和减少过拟合的风险,因此,在样本量计算上更加精细和保守。本研究为构建中医临床预测模型样本量估算提供了方法学借鉴和案例实践演示,期待未来出现更多的关于中医药领域的高质量临床预测模型研究。

参考文献:

- [1] 王燕,李文静,吕红芝,等. 骨质疏松骨折临床预测模型的建立与评价[J]. 中华老年骨科与康复电子杂志, 2022, 8(6): 350-360.
- [2] 王中瑞,朱蓉,甄倩,等. 中老年2型糖尿病合并稳定型心绞痛患者中西医预后模型探讨[J]. 中国实验方剂学杂志, 2024, 30(14): 138-144.
- [3] 孟庆岩,刘圆圆,杨柳,等. 《黄帝内经》五运六气历法探析[J]. 北京中医药大学学报, 2021, 44(10): 874-878.
- [4] 林树元,刘畅,李煜,等. 中医在人工智能时代的挑战与经方智能化研究思路[J]. 中华中医药杂志, 2019, 34(2): 448-451.
- [5] 陈平雁. 临床试验中样本量确定的统计学考虑[J]. 中国卫生统计, 2015, 32(4): 727-731, 733.
- [6] 胡晶,李博,张会娜,等. 针灸临床试验的样本量估计[J]. 中国针灸, 2021, 41(10): 1147-1152.
- [7] 侯鑫. COPD稳定期合并肌少症独立危险因素探究及风险

预测模型的建立[D]. 沈阳: 辽宁中医药大学, 2023.

- [8] 马馨蕊. COPD稳定期气虚体质患者 GOLD 分级水平的预测模型研究[D]. 沈阳: 辽宁中医药大学, 2022.
- [9] 陈莹. 冠心病稳定型心绞痛脾虚痰浊证严重程度的预测模型研究[D]. 沈阳: 辽宁中医药大学, 2019.
- [10] 杨法智. 基于中医体质分布的急性脑梗死复发风险预测模型研究[D]. 济南: 山东中医药大学, 2023.
- [11] PEDUZZI P, CONCATO J, KEMPER E, et al. A simulation study of the number of events per variable in logistic regression analysis[J]. J Clin Epidemiol, 1996, 49(12): 1373-1379.
- [12] VITTINGHOFF E, MCCULLOCH C E. Relaxing the rule of ten events per variable in logistic and Cox regression[J]. Am J Epidemiol, 2007, 165(6): 710-718.
- [13] BU Z J, HUANG Z R, CHEN Y R, et al. Development and multiple visualization methods for the therapeutic effect prediction model of five-flavor Sorghum Flavesces enteric-coated capsules in the treatment of active ulcerative colitis: a study on model development and result visualization[J]. Eur J Integr Med, 2023, 63: 102297.
- [14] RILEY R D, SNELL K I, ENSOR J, et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes[J]. Stat Med, 2019, 38(7): 1276-1296.
- [15] OGUNDIMU E O, ALTMAN D G, COLLINS G S. Adequate sample size for developing prediction models is not simply related to events per variable[J]. J Clin Epidemiol, 2016, 76: 175-182.
- [16] RILEY RD, ENSOR J, SNELL KIE, et al. Calculating the sample size required for developing a clinical prediction model[J]. BMJ, 2020, 368: m441.
- [17] VAN SMEDEN M, MOONS K G, DE GROOT J A, et al. Sample size for binary logistic prediction models: Beyond events per variable criteria[J]. Stat Methods Med Res, 2019, 28(8): 2455-2474.
- [18] PAVLOU M, AMBLER G, SEAMAN S R, et al. How to develop a more accurate risk prediction model when there are few events[J]. BMJ, 2015, 351: h3868.
- [19] VAN H J C. Shrinkage and penalized likelihood as methods to improve predictive accuracy[J]. Statistica Neerlandica, 2001, 55(1): 17-34.
- [20] NAGELKERKE N J D. A note on a general definition of the coefficient of determination[J]. Biometrika, 1991, 78(3): 691-692.

(收稿日期:2024-03-08)

责任编辑:王利敏