

建立多变量多分类模型预测新生儿输血次数：一项回顾性研究

郭彬瀚¹ 陈剑¹ 江咏梅^{1*}

1 四川大学华西第二医院 检验科 四川成都 610041

摘要：**目的：**输血是新生儿诊疗过程中重要的一项手段。然而，由于患儿在不同出生条件下呈现出的个体差异以及妊娠期孕产妇可能潜在的并发症，对新生儿输血所产生的影响尚不明确。在该项研究中，我们旨在建立一个模型来评估新生儿在住院期间可能所需的输血次数，以达到早期预测和精准输血的目的。**方法：**从医院电子病历信息系统中收集相应指标（包括患儿入院期间与输血相关的指标以及患儿母亲孕期相关并发症等），根据住院期间患儿输血次数将数据划分为6组，分别为非输血组、输血1次、输血两次、输血3次、输血4次以及输血超过4次组。通过筛选在不同组别间具有统计学意义的指标后，我们基于多变量、随机森林和梯度增强树多分类任务算法构建了对应的模型来预测患儿的输血次数，并与其真实记录的输血次数进行吻合度比较。**结果：**构建的模型确定了以下变量是影响婴幼儿输血次数的关键因素：体重不足、早产、以及1、5和10分钟时的Apgar评分、妊娠期糖尿病。在不同模型预测性能评比中，与多变量多分类模型（宏平均加权AUC=0.7391）和梯度增强树多分类模型（宏平均加权AUC=0.8442）相比，随机森林模型（宏平均加权AUC=0.8540）表现出了最佳的预测性能。此外，仅由Apgar评分构建的随机森林模型（宏平均加权AUC=0.7967）也表现出良好的预测性能。**结论：**本研究中构建的多分类随机森林模型在预测新生儿输血次数方面取得了良好的效果，有助于在患儿诊疗中精准输血的管理。

四川大学华西第二医院 610041

通讯作者 江咏梅

第一作者 郭彬瀚（Tel 19938119203 E-mail: 2183983697@qq.com）