
为什么很难建立准确的临床预测模型？

作者：陶立元 赵一鸣 来源：临床流行病学和循证医学

本文原地址：<https://www.iikx.com/news/statistics/512.html>

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

在今年中国医师协会循证医学分会的年会会前培训班上，来自北京儿童医院临床流行病和循证医学中心的彭晓霞教授讲课时提到了一篇题为《Why Is a Good Clinical Prediction Rule So Hard to Find?》的文章。听彭老师讲授获益很大，课后觉得有必要看一下这个述评文章，所以在这儿跟大家分享一下这篇编辑述评中的部分重要观点。

临床预测工具/模型(Clinical Prediction Rules, CPRs)近些年来在医学研究领域广受重视，很多团队都在尝试建立各种各样的疾病诊断或是治疗的预测模型，但是经常出现的是这些预测模型的预测准确性得不带很好的验证，同行也不太认可。所以才会让各种模型源源不断的产生，又如过眼云烟，昙花一现。那究竟是什么原因导致很多CPRs预测不准呢？在Grady等人写的《Why Is a Good Clinical Prediction Rule So Hard to Find?》评述中，作者就给出了自己的几点看法：

1)文章中认为这些CPRs的建立往往是基于单中心大样本的数据，通过多因素分析而产生的预测模型。模型在建立时应该选用未来模型真实应用场景中的研究对象数据来建模，且模型建立时应该包括可用来预测该结局的所有变量，同时要求这些变量能够被清晰稳定测量，即短时间内多次测量结果间不会有很大的变化。

2)预测模型往往在建模的数据中表现良好，而换一批数据后预测的准确性则没那么高。原因很简单，因为预测模型是通过建模数据拟合出来的，所以在建模数据中往往表现良好。但是如果换到实际的模型应用场景，往往实际的场景中病人的年龄、性别、患病率和疾病的严重程度等往往与建模样本不同。所以CPRs建立以后往往需要在不同的临床数据库中进行验证后，方可使用。Siontis等人的研究显示，在118个CPRs中仅有10个工具被在不同的人群中重复研究了4次及以上，而在这10个工具中仅有一个工具在不同数据库中预测结果一致。

3)尽管一个临床预测模型并不够精确，但是如果因为该模型的存在而改善了临床结局，那么这个模型也算是个不错的模型。我们可以使用RCT的方法去探索模型存在是价值，即一组研究对象用模型预测，另外一组不用模型，然后我们来观察研究的终点在两组是否是不一致的。最后作者建议我们应该审慎地使用新的预测模型，尤其是当这个模型还没有被很好的外部数据验证之前。

除了上述Grady等人的观点以外，科研论文中预测模型的建立和验证也有相应的报告规范(TRIPOD)，这一规范我们在本微信公众号之前的文章中有报告。除此之外，Steyerberg等人的文章《Towards better clinical prediction models: seven steps for development and an ABCD for validation》指出预测的建立需要经过7个步骤，预测模型的验证需要经历ABCD共4个过程，暂且起名 $7+4=$ “11步预测模型建立法”。

Steyerberg等人的文章指出预测模型建立的7个步骤分别是：1)提出研究问题，检查数据；2)合理编码预测变量；3)规范化模型参数；4)估计模型；5)模型的预测准确性；6)模型的验证；7)模型的展示方法。预测模型验证ABCD四个方面如下图：

Table 2 Overview of four measures (ABCD) for model performance

Aspect	Measure	Visualization	Characteristics
Calibration	A: alpha Calibration-in-the-large	Calibration plot	Intercept in plot; compares mean observed with mean predicted
	B: beta Calibration slope		Regression slope in plot; related to shrinkage of regression coefficients
Discrimination	C-statistic	ROC curve	Interpretation as the probability of correct classification for a pair of subjects with and without the endpoint
Clinical usefulness	Decision-curve analysis NB	Decision curve	Net true-positive classification rate by using a model over a range of thresholds

感兴趣的读者可以参考下面几篇文章：

- 1、 Why Is a Good Clinical Prediction Rule So Hard to Find? (PMID：22025427)
- 2、 Good Prediction Rules Are Hard to Find Did We Succeed? (PMID：23070111)
- 3、 Towards better clinical prediction models: Seven steps for development and an ABCD for validation(PMID：24898551)

更多 统计方法 请访问 <https://www.iikx.com/news/statistics/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](https://www.iikx.com)转发