肺腺癌生存模型，换一批患者还能用吗？

讲完一张生存曲线，导师可能会问：“给另一批患者用同一个模型，还能分出谁的风险更高吗？”

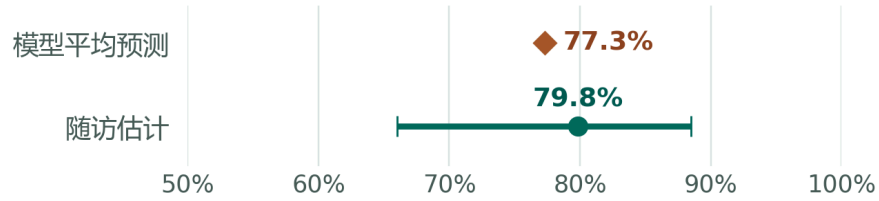
TCGA-LUAD · 建模	CPTAC-LUAD · 验证
455 人	72 人
全随访期 156 例死亡	全随访期 19 例死亡

我们先定好入组规则，让 BioAgent 用年龄和病理分期为 I–III 期患者建立 Cox 生存模型。固定系数后，Agent 给第二批 75 人评分，保存预测，再核对随访。3 人生存状态未知，最终 72 人可评估。

能排出风险，也要核对生存率

外部队列的 **C 指数为 0.767**（95% 区间 0.652–0.860）。在可比较的患者对中，模型约有 77% 的机会给较早死亡者更高的风险评分。

同一批 72 人的两年总生存率



横线为随访估计的 95% 区间（66.0%–88.5%）。

带一张能复核的图去组会

两年生存率的预测值与随访估计只差 2.5 个百分点。这足以说明模型预测得准吗？

整段随访记录了 19 例死亡，两年内有 11 例。两年时仍有 **33 人在险**，即活到了这个时点且仍有随访。有人更早结束随访，因此用 Kaplan–Meier 法（KM）估计生存率。图中的区间较宽，两组数值接近也只说明这批患者整体的情况，个体预测仍需更多证据。

用公开病例表，亲手复核一次

解压结果包，把 **外部可评估病例.csv** 上传到 BioAgent，再发送下面这段任务。

> 请用 `os_days` 和 `os_event` 估计 730.5 天时的 KM 生存率和 95% 区间，与 `predicted_24m` 的均值画在同一张图上。列出患者数、两年内死亡数和两年时在险人数，附计算代码。

下载公开结果包（1.1 MB） **登录 BioAgent**

> **转给导师时可以这样说**
>
> 老师，我想用这份肺腺癌数据复核一张图。模型预测两年生存率约 77%，两年时还在随访的只有 33 人。下次组会我把图和人数一起讲，再看咱们的随访够不够做类似验证，您看可以吗？

原始数据 TCGA 临床终点资料 · CPTAC 肺腺癌研究

文中数值来自云生信对两个公开队列的教学分析。