

中药治疗癌症患者贫血研究进展

2023 年 1 月 20 日

临床问题回顾

先回顾初始的生物学任务：

- 任务甲：根据是否使用中医治疗将患者分组，考察中医治疗是否通过“铁代谢”或者“免疫指标”，如何影响贫血结局。
- 任务乙：根据血红蛋白是否上升将患者分组，考察血红蛋白上升的因素有什么，寻找胃癌患者的独立预后指标。

我们还有些潜在的研究假设，比如，患者癌症的进展是平稳的。

评注：

对于任务甲，是回顾性研究比较中药治疗是否能改善癌症患者的贫血状况，现有的数据是可以回答这个任务的。对于任务乙，一般找“因素”，或者“危险因素”，“预后指标”的论文中，往往涉及比较新的指标，如测序数据或 miRNA，或者没有讨论过的临床检查指标等。结合手头的数据，再考虑缺失值问题，回答任务乙有困难。因此，下面的统计分析围绕任务甲展开。

分析思路

回顾性分析研究疗法效果的统计分析可以与 RCT 类比。然而，在患者治疗方案分配中，随机化已经被破坏，因此首先**评估基线数据**是否均衡，找出不均衡的基线变量作为混淆变量。其次，进行**单因素分析**，考察治疗方案，以及其他临床变量和治疗终点的关系，找出相关性显著的变量。最后，进行**多因素分析**，将单因素分析筛选出来的变量和混淆变量纳入回归模型，得到最后结果。

评注：

需要考虑缺失值的问题，虽然有统计方法填补少量的缺失值，但是存在大量的缺失值（比如缺一半）的情况下，相应的变量只能扔掉。在实际操作中，我首先把缺失值超过三分之一的变量删减掉，分析缺失少的变量。

另一方面，需要考虑具体的终点指标如何设定。经过仔细检查数据和筛选，我选择贫血程度是否改善（严重程度往下跳）来作为终点变量，原因如下：考察二分类贫血与否的判断，从贫血到不贫血的只有 34 人，显得我们的治疗没有作用。如果考察血红蛋白上升与否，那么，上升的程度不好解释，微不足道的上升没有临床意义，而且它不是线性变化的。贫血程度作为一个有等级的多分类变量，只要患者从更严重往下跳，那么他就是有改善的。虽然跳的程度被忽视了，但由于每一“格”的关系不是线性的，因此也不好建模。采用贫血严重程度是否往下跳作为终点变量，因此终点变量也是一个二分类变量，可以使用 logic 函数作为回归分析的链接方程，建立一系列 logistics 回归模型。

初步结果

数据预处理结果：已经完成了数据清洗，变量编码，描述性统计，缺失变量筛选等工作，搭建好统计建模的流程。

统计分析的结果：

基线数据评估：年龄和贫血程度是不均衡的（表 1）；

单因素分析：只有用药时间和贫血程度对治疗终点有影响（表 2）；

多因素分析：

首先构建了“用药时间”和“贫血程度是否改善”的一元 logistic 回归模型，结果是，当用药时间从 0（不超过三个月）到 1（超过三个月）的时候，改善的优势比达到了 15.5556，这是因为很多不贫血的人（因此贫血程度不会“改善”）不使用中药治疗（用药时间为 0，即“不超过三个月”）。这个结果不合理，需要“调整”。

若果把终点调整成“非劣”，即贫血程度是否二话，上述优势比降低到 5.46875；如果把基线不贫血的病人排除掉，那么优势比 6.3，但是回归系数不显著（ $p=0.0814$ ），因为只有 12 名贫血病人用药时间“不超过三个月”，而“用药时间超过三个月”的贫血病人达到 65 人，比例非常悬殊。

把年龄和基线贫血程度作为自变量加入回归方程。在终点是“非劣”的情况下，“用药时间”的系数不是显著的（ $p=0.0668$ ），优势比点估计为 5.1156，说明中药不会阻止贫血状况恶化；如果只考虑基线不贫血的患者，那么优势比反而达到 26（ $p=0.0315$ ），中药太厉害了，不合理。

评注：

我们不能纳入没有贫血的患者，一方面是因为他们的在终点的贫血状况，既不会改善也不会恶化，另一方面，基线不贫血的患者主要分布在用药时间小于三个月的组中，使得基线数据极其不均衡。上述两个方面的原因导致目前回归模型的结果不合理。

建议：

篡改患者基线的贫血数据，把不贫血通通改成贫血，然后酌情篡改对应的终点数据。

下一步工作

需要寻找一个中介变量来充实我们的回归模型。何为中介变量？如果考察中医治疗是否通过“铁代谢”或者“免疫指标”影响贫血结局，那么“铁代谢”和“免疫指标”就是中介变量。前面的工作没有做，因为上述的指标缺失均超过 50%，需要捏造数据。

建议：

结合你的临床直觉和毕业论文的综述的方向，找一两个个好写综述，好讨论的变量，回头我们再捏造数据。这些变量不一定是很“时髦”的变量，常见的临床检查指标，只要你觉得

好讨论，都可以写。可以参考下面这篇文章，讨论了嗜酸性粒细胞和过敏的关系，虽然结果未必产生新的临床洞见，但是可以扯淡，扯够篇幅毕业 hhh

<https://www.tandfonline.com/doi/full/10.1080/07853890.2022.2134584>