临床研究中匹配还是不匹配？这是一个问题

作者：张华 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/1554.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！

在临床研究中混杂因素的控制是一个永恒的话题，匹配作为处理混杂因素的一种方式，在临床研究设计和分析阶段都可以使用，也被
广泛应用。最近有人问

相信不同的人对此用不同的答案，我首先说一下我的见解，那就是做匹配，如果匹配不成功再做多因素分析。原因如下：

1.在设计阶段匹配，**匹配可以减少工作量**
。如果数据还未收集已经预估到对照组的病例数量比暴露组多很多，我们可以先收集所有病例的一部分重要的混杂因素，根据暴露组的特点选取一组相像的对照组，再收集匹配成功的病例收集更多的资料，而匹配不成功的病例的资料不需要再收集，可以提交工作效率，减少研究的工作量
；

2.从统计效率上讲，
**两组病例比例差别较大时，统计效能并不
能有太大提高**
。从统计模拟的结果来说，当病例组数量固定时，对照组可以是病例组数量的1-4倍，当高于4倍时，对照组病例数量再增加，统计效能增加很少很少，可以说是事倍功半，因此即使数据已经完全收集完成，我们使用1：4匹配的方法放弃一部分病例组的病例，对研究的把握度影响不大。

3.
匹配可以去掉一些极端病例，提高两组的可比性
。匹配的过程是寻找相像的匹配对象的过程，可以将一些极端病例剔除。极端的病例可能是真实的数据，但极端的病例数据出错的可能性比一般病历数据的可能性要大，有可能因为各种原因出错。即使是数据信息是真实的，如果不剔除也会影响回归模型的稳定性和准确性。而匹配后的两组病例基线资料更加相似，可比性提高。

匹配的弊端也是有的，第一就是**损失了外部真实性**
，匹配后的病例数量减少，对照组的代表
性下降。所以
做不做匹配还是要结合研究类型、研究的阶段和研究目的去选择。我个人认为研究首先要关注两组可比性，只有两组可比时，在研究对象中结论可靠，再去探讨结论是否可以外推到设计人群和目标人群

。第二，当病例组本身样本量不大时(如10-50例)，如果再进行匹配，会导致总体样本量较少，损失的统计效能不能忽视，匹配后研究质量不高。而在这种情况下，即使不匹配，使用回归的方法做分析，研究质量也不高，结论也不可靠。此时匹配时可以增加对照的比例(如1：6匹配)可适当提高统计效能。

以上是我个人见解，大家有不同意见可直接拍砖，共同讨论。

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转发