

临床研究中的常见偏倚——病例错分

作者：曾琳，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/8264.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！

临床研究中常见偏倚——病例错分

。临床研究中常常需要进行疾病的诊断，但是我们都知道，在研究中使用金标准对研究对象进行诊断在很多情况下并不可行。比如研究吸烟和肺癌的关系，肺癌的诊断就很难用病理学的方法来判断研究对象是否罹患肺癌，因为我们总不能不管三七二十一就取研究对象的肺组织做蜡块切片吧。所以我们常常会用影像学手段来替代金标准来判断研究对象是否罹患肺癌。

也就是说由于可行性的限制咱们在临床研究中常常会用其它诊断手段来替代金标准来进行研究对象的诊断。看到这儿大家可能会觉得奇怪，这种情况在临床研究中司空见惯，大家都接受了，有什么可说的呢？小编要说，存在、常见并不就一定就正确。使用替代标准来对研究者进行诊断是会给研究结果带来偏差的。下面咱们用一个例子来看看疾病的错分会如何影响我们的研究结果。

还是说吸烟和肺癌关系的例子，假设我们开展了一项队列研究，随访观察900名吸烟者和900名不吸烟者，他们20年后肺癌的发生真实情况如以下四格表：

金标准诊断的四格表

暴露	肺癌	无肺癌
吸烟	150	750
不吸烟	50	850

那么我们真实的研究结果应该是：

$$RR = \frac{150 / (150 + 750)}{50 / (50 + 850)} = 3$$

吸烟是肺癌的危险因素，且吸烟者罹患肺癌的风险是不吸烟者的3倍。

因为病理诊断的可行性不好，在这个研究中我们只能使用影像学方法来进行诊断。虽然，影像学方法诊断的准确度很高，但毕竟不是金标准，一定会存在偏差。假设影像学诊断的灵敏度和特异度分别是0.9和0.8。即：sen=0.9，sep=0.8，那么我们在研究中会得到一个怎样的四格表呢？首先我们来看看如何通过真实情况来推算影像学诊断结果。

病理诊断和影像学诊断的关系

		病理诊断		合计
		+	-	
影像学 诊断	+	a	b	<u>a+b</u>
	-	c	d	<u>c+d</u>
合计		<u>a+c</u>	<u>b+d</u>	N

因为 $Sen = a / (a + c)$ $Sep = d / (b + d)$ 那么

$$D_{\text{影像}} = a + b = D_{\text{病理}} \times Sen + \bar{D}_{\text{病理}} \times (1 - Sep)$$

$$\bar{D}_{\text{影像}} = c + d = D_{\text{病理}} \times (1 - Sen) + \bar{D}_{\text{病理}} \times Sep$$

那么从真实的四格表我们可以得到以下数据：

$$D_{\text{吸烟}} = D_{\text{吸烟} \cdot \text{病理}} \times \text{Sen} + \bar{D}_{\text{吸烟} \cdot \text{病理}} \times (1 - \text{Sep}) = 150 * 0.9 + 750 * (1 - 0.8) = 285$$

$$\bar{D}_{\text{吸烟}} = D_{\text{吸烟} \cdot \text{病理}} \times (1 - \text{Sen}) + \bar{D}_{\text{吸烟} \cdot \text{病理}} \times \text{Sep} = 150 * (1 - 0.9) + 750 * 0.8 = 615$$

$$D_{\text{不吸烟}} = D_{\text{不吸烟} \cdot \text{病理}} \times \text{Sen} + \bar{D}_{\text{不吸烟} \cdot \text{病理}} \times (1 - \text{Sep}) = 50 * 0.9 + 850 * (1 - 0.8) = 215$$

$$\bar{D}_{\text{不吸烟}} = D_{\text{不吸烟} \cdot \text{病理}} \times (1 - \text{Sen}) + \bar{D}_{\text{不吸烟} \cdot \text{病理}} \times \text{Sep} = 50 * (1 - 0.9) + 850 * 0.8 = 685$$

影像学诊断结果四格表

暴露	肺癌	无肺癌
吸烟	285	615
不吸烟	215	685

这时由于我们的诊断方法用的是影像学方法而非金标准，那么我们得到的结果是：

$$RR' = \frac{285 / (285 + 615)}{215 / (215 + 685)} = 1.33$$

即吸烟是肺癌的危险因素，且吸烟者罹患肺癌的风险是不吸烟者的1.33倍。150虽然同样得到的结果方向和真实结果一样，那么我们得到的危险因素的风险估计不到真实风险的一半。

下面我们来多做几次模拟：

Sen	Sep	RR
0.8	0.7	1.17
0.9	0.7	1.2
0.8	0.8	1.29
0.95	0.9	1.64
0.99	0.95	2.02
0.999	0.995	2.83

可见，由于替代的诊断方法与金标准存在差距，无论我们怎样提高替代诊断方法的准确性都会低估我们的研究中危险因素的致病风险。替代诊断方法的准确性越高，得到的研究结果与真实值之间的偏差越小。也就是说，在可能情况下，我们应该进行在研究中选择金标准对研究对象进行诊断，如果实在不可行也要尽量选择目前准确度最高的方法进行诊断，减少病例错分带来的偏倚。

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](http://www.iikx.com)转发