诊断试验中如何处理中间状态？

作者：李楠，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/6255.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！

诊断试验中如何处理中间状态？经典的诊断试验通常是用一个指标区分两种状态。比如我们通过检查羊穿得到的细胞，看21对染色体是否存在三条(存在为阳性，不存在为阴性)，从而诊断胎儿是否患有21三体综合征;或是通过认知功能筛查量表得分，判断患者是否存在认知功能障碍(低于某分值认为异常，否则认为正常)。上面两个例子不同的地方是分别用了二分类变量(阳性/阴性)和连续变量(量表得分)作为诊断指标，连续变量时需要实现确定一个判断界值。但是他们都有一个重要的共同点，就是**待区分的真实状态只有两个**。

其实很多时候，我们在对患者进行临床诊断的时候，并不是仅有两种待区分状态的。比如用筛查量表判断患者认知功能的时候，实际上患者存在正常、认知功能受损、痴呆这三种状态;又或是通过某种实验判断患者韧带状态时，患者存在正常、部分撕裂、完全撕裂三种状态。这时候我们该如何对待呢?
在评价某一种诊断方法或指标的时候，是否还能计算灵敏度、特异度等这些指标呢?

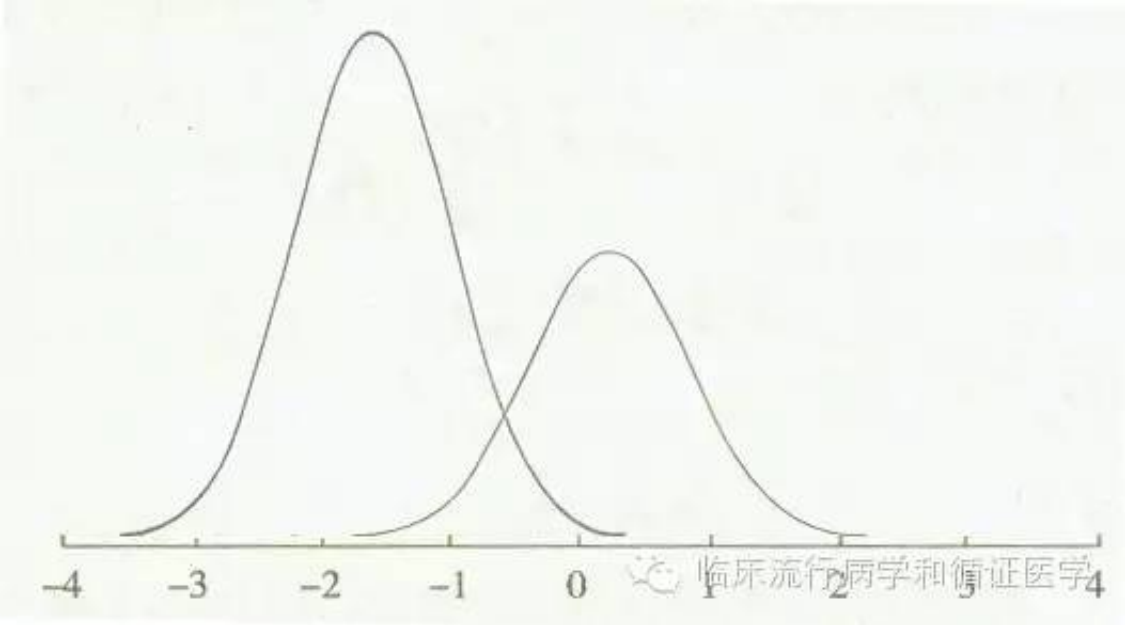
让我们先来看看两个最常用的定义。

灵敏度：又称为**真阳性率(true positive)**，即实际有病而按照筛检试验的标准被正确地判为有病的百分比。

特异度：又称为**真阴性率(true positive)**，即实际无病而按照筛检试验的标准被正确地判为无病的百分比。

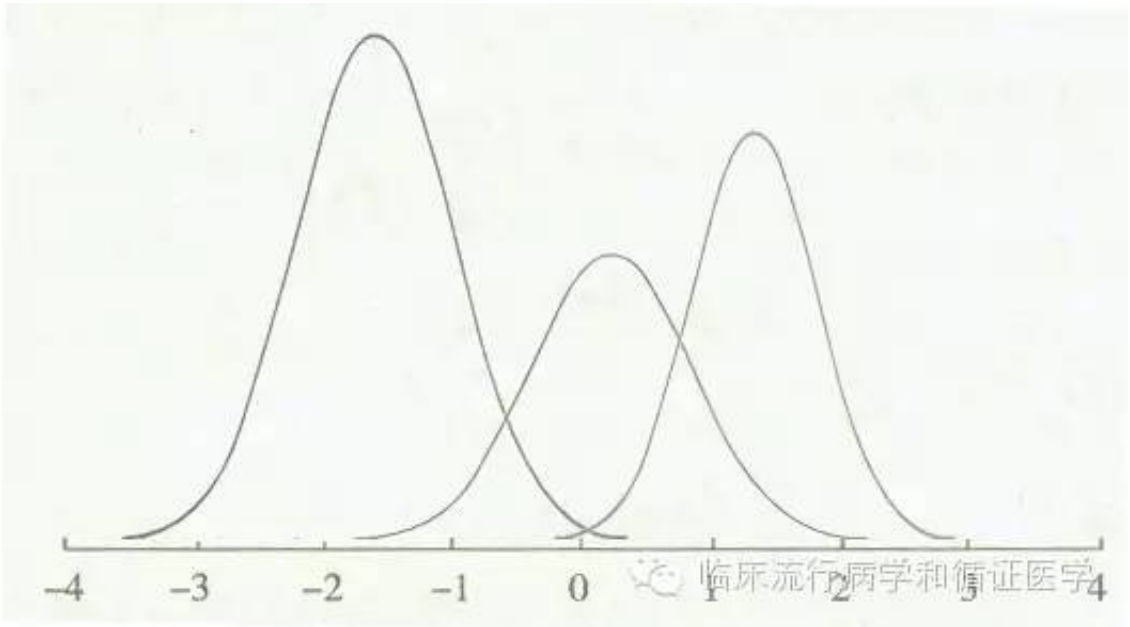
当然还有很多其他指标，但是讨论今天的问题，先说这两个指标就够了，其他指标在存在多个待诊断状态时遇到的问题与这两个指标相似。

通过定义不难看出，灵敏度、特异度及其他指标都有一个关键的核心假设，也就是患者只存在两种待诊断状态。即“无病”与“有病”。我们举个当筛查变量是连续变量时的例子，我们待区分的两类人，他们的筛查指标分布往往是下面这样的：



此时我们需要的是把右侧的“有病”受检者，与“左侧”的无病受检者区分开。可以简单直接的计算该方法在某一界值时候的特异度和灵敏度等指标。

但如果我们的待区分状态有三个的时候，似乎就变的比较复杂了，让我们先来看看带筛检患者的分布会变成什么样子：

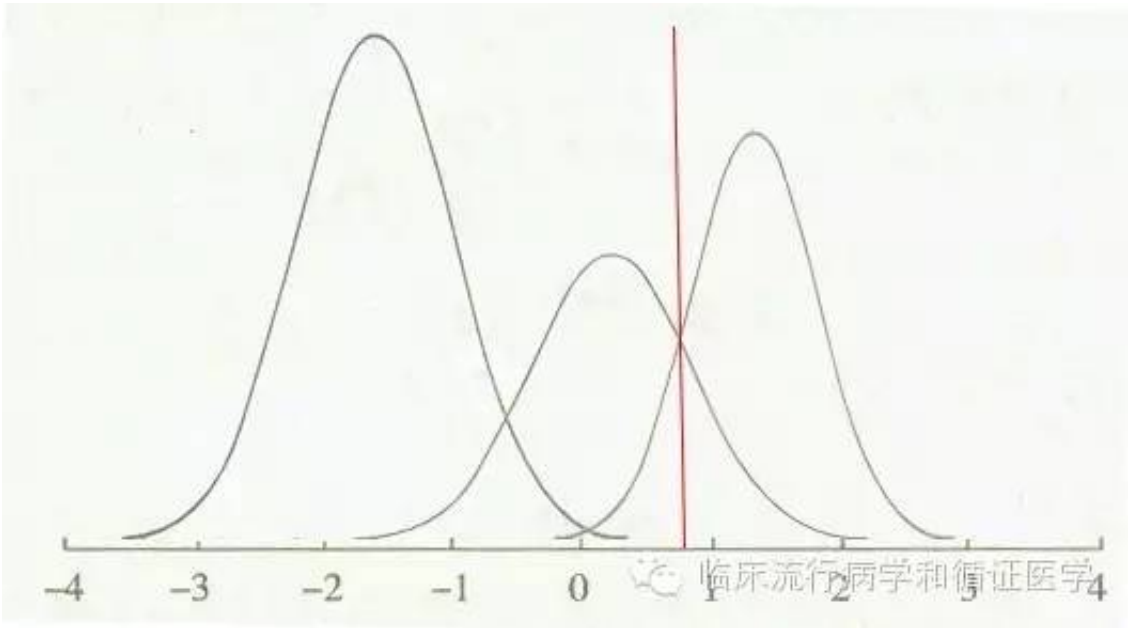


从左到右分别是“无病”、“轻度异常”、“有病”三个状态，此时似乎按照灵敏度、特异度的定义计算时就无从下手了。
这时如果需要探讨该筛查指标的效能，就**需要将三个状态降低为两个状态才行**。

我们可以把中间状态“轻度异常”归为“有病”计算相关指标(如下图)，此时灵敏度、特异度的意义在于评价该指标从待诊断人群中区分出异常(包括轻度异常)人的能力：



也可以把中间状态“ 轻度异常 ”归为“ 无病 ”计算相关指标，此时灵敏度、特异度的意义在于评价该指标从待诊断人群中区分出“ 有病 ”人的能力：



好了，这是我们对待区分三类待诊断人群时候的标准方法。肯定不少医生会问：那如果我只关心这种筛查方法区分“ 轻度异常 ”对象的能力，是否可以但提出“ 轻度异常 ”这一组中间状态来计算灵敏度、特异度呢?这时候我就要多问您一句了，真实情况下可能出现仅从“ 正常 ”、“ 轻度异常 ”两类人中鉴别“ 轻度异常 ”的情况么?比如一些疾病一单发病，症状极其明显，不用任何指标肉眼就能看出来，这时其实诊断指标的主要目的就是从未发病的潜在患者中诊断出轻度异常的人，我们当然可以去掉患者来计算特异度和灵敏度。

但是如果某一疾病起病隐匿，需要更复杂的金标准才能将真正有病的人鉴别出来，此时我们在对人群进行筛查时，不可能事先排除掉“ 有病 ”的人，而是将三类人混在一起诊断，这时我们就无法仅仅计算出该方法诊断“ 轻度异常 ”的能力。即便根据金标准排除有病的人之后计算出来，也

没有任何临床价值。

此外，还有不少医生犯过这样的错误，就是把所有“轻度异常”的人数作为分母，把正确判断为阳性的“轻度异常”人数作为分子，试图计算敏感度。这是最致命的错误在于，您把“无病”和“有病”的人归为一类了!认为他们是相同的状态。这样计算出来的指标怎么能符合逻辑呢?

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