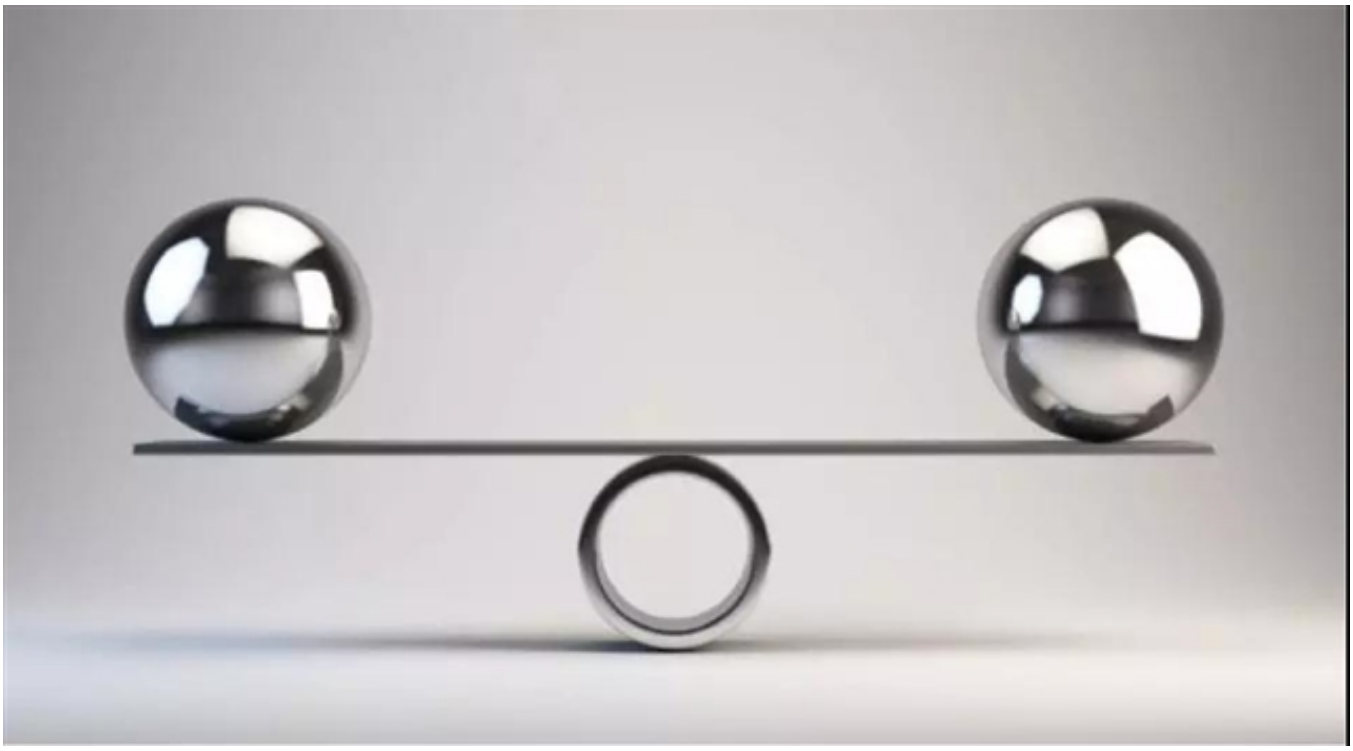
沿用还是弃用P值：还是让它锦上添花吧！ BMC Journal

作者：writer 来源：科学网

本文原地址：https://www.iikx.com/news/progress/6691.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！

沿用还是弃用P值：还是让它锦上添花吧！ BMC Journal。p值已经走到的尽头，没有其他花样了吗?在接下来的分享中，Philosophy, Ethics, and Humanities in Medicine 的主编Christina Cantrell和James Giordano通过最近关于p值的讨论，为我们讲述了p值的价值。



© Sashkin / stock.adobe.com

是不是该权衡下P值的价值了?

最近，Ioannidis等人提出p≤0.05这一阈值可能不具备足够的统计学严谨性来证明结果的有效性。提出这一观点是有些道理的。将p值的严格性从≤0.05提高到≤0.005或0.001，将会为统计学显著性提供一个更加细化的方法和标准。

但在绝大部分情况下，凡事并非完美无缺，即使没有明显的毛病或问题，也会面临一些负面影响。因此，我们应适当地识别和权衡这种标准转变可能或将产生的相对利弊，这一步是很有意义的。这项任务应由研究企业中尽可能多的利益相关者和股东共同承担，因为这其中的利弊对于双方都至关重要。

首先，一个重要的问题是生物医学(和其他领域)中有哪些学科和领域最能受益于采用更严格的p值。科学界(或者一些重要的公众人群)有必要对这一问题达成共识或广泛认可，这一点是非常重要的。例如，对于一些受试者数量较少和/或结果和效应的差异相对较小的研究领域(例如，新兴技术领域)，会面临阻碍，使用p ≤0.001、0.001等阈值是不错的选择。对于补充医学领域，在某些情况下也是如此，例如当先前的研究存在瑕疵，或者没有很好地获取结果时。在这两种情况下，更严格的p值阈值和对(阳性)结果的更严格的审查会使结果的接受度更高。

然后，也有一种情况是，使用更严格的p值会使安慰剂效应的发生率更高。这可以是正面的，也可以是负面的，很大程度上取决于所用的安慰剂及其含义。诚然，将安慰剂反应视为仅仅缺乏效应，则安慰剂效应为负面效应。但这些反应真的缺乏效应吗?或者更确切地说，它们是否反映了由所评估的药物或技术以外的某些因素引起的生理效应呢?

如果我们想通过采用更严格的p值来让研究方法更符合潮流，那么我们是否也应该对安慰剂效应有个符合潮流的认识呢?即认为安慰剂反应是由多个过程和机制所引起的生物功能变化。

在研究安慰剂(和其他)反应底物的研究中采用更严格的p值可以揭示是什么在起作用，什么不起作用，在谁的体内起作用以及在什么条件下起作用。通过更严格的标准(例如p≤0.005～0.001)得出，这些发现可能确实具有显著的成本效益。同样地，采用更严格的p值可能会提示需要对先前的研究重新进行检查和评估。

为实现这一目标，我们建议首先从研究(健康、疾病和损伤)机制以及最有可能产生潜在获益或损害的工具和技术的研究开始，以便重新评估产生的结局可能带来的相对益处和/或损害。

或者，是否应该像一些人建议的那样，放弃采用p值?p值是否不再具有价值了，对此我们并不认同，我们认为p值(至少部分程度上)仍然是有用的。对于某些类型的结果，将p值与其他结果结合起来应用，可以提供更准确的度量指标来评估结果的质量、意义和价值。例如，所评估的干预措施产生了哪种效应?这些发现对临床医生和患者有意义吗?可信区间是多少?是否可以使用其他统计方法(例如贝叶斯法)来评估正在测试的内容?

鉴于此，我们建议创建标准，以确定在特定类型和特定研究背景下给定p值的相对可接受度和要求。可以肯定的是，这可以为我们提供一个有用的工具包。因此，所谓的一锤在手，打遍天下无敌手这种使用单一工具的时代已经过去，现在是时候克服困难并抓住机遇开发新的工具和度量标准了。我们认为，只有拥有具有不同尺寸和重量的锤子，并且可以与其他工具配合使用才能破除这一局限性，并有助于开发出更适合当前和未来工作的研究方法。

更多 科学进展 请访问 https://www.iikx.com/news/progress/

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