

---

# 回答一个问题时变量很多怎么办？

作者：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/2166.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！**

A问题：“在论文中回答一个问题可能包括20+变量，其中16个都是有意义的，结果非常多，罗列在论文中是不是堆数据？如果想改变，如何改变？”这是一个很好的问题，过去我也被困扰过很长时间，后来想通了，找到了解决办法，在这里和大家交流分享。

这位朋友提出的问题实质是，当我们

在研究和回答某一问题时发现许多因素与结局相关，即“多因一果”时，如果将所有结果都放在论文中有堆数据的嫌疑；而全部不放在论文中会导致结果不完整，同样会受到质疑，是一个两难选择，怎么办？

这种情况常见于病因学研究和预后研究，许多危险因素/预后因素与结局相关，尤其是比较传统的研究领域，已经有大量研究证实了多个危险因素/预后因素，在研究中必然是阳性结果，是对已知规律的验证，是有意义的，说明研究的方法和过程符合科学研究的要求，因而得出了与以往研究类似的结果。这些结果不放在论文中读者一定会质疑研究工作做得好不好，为什么不验证已知规律。但仅仅验证已知规律的论文没有发表的价值，论文发表的关键是必须有创新点，只要有一个创新点就足以支持论文发表。如在研究中发现了一个新的危险因素/预后因素，这就是论文总结发表最有价值的“点”，论文的组织撰写及结果展示都应围绕这个创新点展开。以“想要好好学习”提出的问题为例，16个变量都与结局相关，其中是否有一个危险因素/预后因素是新的，没有文献报告？如果有，可以以这个变量为核心组织数据分析和结果展示。在统计分析时16个变量的地位是相同的，但在论文中属于创新发现的变量是“老大”，其它变量为“老大”服务。单因素分析结果的展示往往把“老大”放在最重要的位置，如最早出现，甚至给“老大”专门做一个表格，提示读者：这是这篇论文中最重要的发现！其它15个变量可简化处理，只要提供关键信息验证已知规律即可，不需要展开。多因素分析仍以“老大”为核心，分析“老大”是不是独立的危险因素/预后因素，“老大”对结局的贡献及其与其它变量对结局贡献的相对比例等。显然，由16个变量同等对待转变区别对待，就可以避免堆数据，就可以突出创新点，就可以点石成金。

创新点可以是发现一个新的危险因素/预后因素，可以是在已知某一危险因素/预后因素的基础上进一步做剂量-反应关系分析，可以是多个已知危险因素/预后因素之间的交互作用……。数据分析和论文撰写中有许多这类套路，不妨在读论文时留心学习，作为预案，有可能在今后的研究中用上。

我们讲

临床研究应以临床

问题和科学问题为导向，是说研究应

该有想法，不是堆数据

---

。A问题是在没有想法时出现的，拿到的16个阳性变量就没有重点了，堆在论文中不合适，不堆也不合适。统计分析只是工具和手段，不能解决堆数据问题，没有想法只能堆，有了想法就不堆，关键是想法。我在美国与生物统计专家讨论时发现一个现象，这些专家从来不跟我讨论统计问题，而是问我研究的具体问题，问题搞清楚了，想法提出来了，统计分析的思路和方法自然就出来了。统计分析方法都是现成的，拿过来正确使用就可以了，尤其有了计算机和软件，计算是非常容易的，所以没有必要在这方面纠缠，不清楚看书找人就能解决。

临床研究可以“小题大做”，当可用的变量很多时“小题大做”大有可为。通过深入分析可以挖掘出隐藏在现象背后的未知规律，是非常有意思的挑战。在统计分析和评价结果时要有“平常心”，我始终认为如果是规律自然会在分析过程中蹦出来，如果不是规律，强行按事先设想一定要这样或那样是不可能的，除非造假。既然规律自己会蹦出来，就不要急于求成，按要求一步一步做分析，自然会水到渠成。我还有这样的经验，在“平常心”状态下找到的规律在后续别人的研究中得到反复证实，证明确实是客观规律。

当我们在研究中遇到困难时，纠结时，要想到这是修炼的机会，是创新的机会，不要放弃，要想方设法克服困难。研究是痛苦的过程，写论文是痛苦的过程，这种痛苦并快乐的特殊体验只有身在其中的人才能体会到。

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](https://www.iikx.com)转发