



吕军：如何做好临床研究 这锅“饭”



Q1 吕军教授您好，我们非常荣幸能够邀请您来参加第四届 ISMRP 暨上海肿瘤会议！您作为暨南大学附属第一医院临床研究部主任，深耕临床研究多年，在大数据应用方向更是成就卓越。现在很多医生也都有做临床研究的想法，但苦于很难进行实操，这方面您有什么好的建议吗？

A1 这确实是一个很好的问题，也是许多年轻医务人员的困惑。大家都希望做科研，而临床研究正处于热门阶段。我们团队专注于临床研究方法学，在医院也进行了一些相关工作，研究生们也都从事这类工作，那么针对这个问题我来说一下我的心得。

在现阶段，临床医生进行临床研究其实面临许多挑战，包括缺乏时间、数据和研究方法等。所以这个时间我们就要挤一挤了，例如每周末留出一天专门用于研究，这一年下来就会有 52 天的时间，这种方法非常有效率。在方法的学习上，最重要的是练习，并且要有实际的数据进行操作。虽然我们可以从医院收集数据，但这种方法往往效率较低。例如，一个横断面研究可能需要至少 500 例的数据，这需要很长时间才能收集齐。因此，我建议大家充分利用学术界现有的大量公共数据库，这些数据库提供了高质量的原始数据。如果能善加利用这些公共数据库，我们相当于就有原料了，有了原料我们就可以把临床研究慢慢做起来。

我把临床研究比作“做饭”，做饭我们要有锅，还要有米，两个缺一不可。没有米，巧妇难为无米之

DOI: 10.14218/MRP.2024.12102
通讯作者：吕军 Email: lyujun2020@jnu.edu.cn

炊；没有锅，压根就搭不起来台子。先说“锅”，临床研究主要依赖于临床流行病学和统计学。在医学领域，公共卫生与预防医学下设的一个重要二级学科便是流行病与统计学，它涵盖了各种研究方法。作为一名流行病统计学方向的博士生导师，我对此领域较为熟悉。流行病学和统计学的核心内容略有不同。统计学主要研究统计方法，而流行病学更注重临床研究的设计，包括病例对照研究、队列研究和横断面研究等。统计方法可以细分为初级、中级和高级方法。中级统计方法主要是三大检验，用于比较病例组和对照组之间的数值差异，但其功能有限。临床研究更常依赖于高级统计方法，如线性回归、逻辑回归等，三大回归方法是临床研究的主要支柱。在医学教育中，统计学常被视为很难学的科目，尤其对于医学生来说，学习过程常伴随着困难与挫折。部分原因在于传统学习中注重公式推导，而忽视略微实际应用。

对于医务人员，统计学的学习应注重以下三方面：

一是方法的应用场景，避免用错方法；二是实际操作，即如何处理数据；三是结果解读，即统计结果对临床研究的意义。掌握这三点对于医务人员已基本足够，学习时可优先关注这些方面，满足研究的需求。关于研究设计，使用现有数据进行回顾性研究是常见做法，包括横断面研究、病例对照研究和队列研究。过去的学习中往往偏重理论，缺少医院病例实践的机会。

有效的教学应基于实际经验。在我的标准里，什么样的人有资格讲这三种研究设计？必须是做过这样研究的人，需具备相关的实战经验，并在该领域发表过至少三篇论文。理想情况下，讲授三种研究设计的讲师应累计发表至少九篇相关论文。实战经验能帮助研究人员更快地判断适用的研究设计并形成论文框架，从而解决临床问题。统计和设计加起来就是我们的“锅”，搭好了“锅”，还需要“米”。

前面有提到，我们自己收集数据可能效率不够。我之前试过一个比较好的研究，但收集数据要3个月，以这个速度我一年只能做4篇研究——太少了。但是国际上有大量经过多年积累的公共数据库，如美国、

英国和欧洲的数据库，提供高质量的数据供研究使用。我们团队曾发表30篇论文介绍各类数据库，共涵盖30余个数据库。实际上公共数据库比30个要多得多，翻番甚至乘以10可能都是有的。当前我们团队集中使用约六至七个数据库。对于初学者，推荐使用SEER（肿瘤数据库）和MIMIC（重症数据库），其中MIMIC与医院数据库结构相似，能帮助研究者更好地理解医院数据。这些数据库的优势在于提供了高质量和大规模的数据，用于技术锻炼和研究方案的开发。通过对这些数据的使用，可以形成多种研究方案。

有人说你这是国外的数据对我们国家有什么用？实际上使用这些数据库，不仅能磨练技术，还能将来处理自有数据打下基础。所以我觉得医务人员研究之初可以用一下公共数据库，不要觉得它是什么洪水猛兽这样不好的东西。当然我们自己也要收集数据。现在比较好的方式是既用国外的数据，又用我们自己的数据，中外对比，高分论文很多都是这个策略。关键在于掌握这些方法，尤其要注重实操练习。

很多医务人员问：“我需要多久才能学会？”要实现这样的能力，需要系统化的学习过程。这也是我们在医院内部培训的方式：用一个月的时间让学员掌握理论和实际操作技巧。理论期结束后，进入为期五个月的实战期。这段时间用于将所学应用于实际研究并发表论文，以此检验学习成果。论文发表是检验学习效果的重要标志，只有发表论文才能证明学习的实效。尽管这半年的学习过程会很辛苦，但付出绝不会徒劳。

一旦完成这一周期，系统性地掌握了方法并成功发表论文，基本就能实现SCI论文的自由发表。普通临床研究论文的发表并不难，挑战在于发表中科院一区和影响因子10分以上的高水平论文。要达到这一高度，需要从简单的开始，比如先发表影响因子较低的论文，从1分或2分逐步提升，不可能一开始就发顶刊论文。成功就如同建造金字塔，必须夯实基础才能攀登到顶端。要实现高水平论文的发表，可能需要两到三年甚至更长的时间。因此，

第一个半年是奠定基础的重要阶段，而未来的 2-5 年则是实现顶级刊物发表的计划。

Q2 感谢您的分享！那最后还有一个小小的疑问，不是您医院的医生有机会跟您学习吗？

A2 我们最初的培训主要面向内部人员，但随着领导提出更高的要求，我们开始考虑让更多人参与学习。因此，我们开设了一些对外的培训班。原本内部的课程周期是一个月，但

对于对外培训，显然无法让学员停留一个月。因此，我们将一个月的课程内容尽力压缩为 5 天，并且只提供线下课程。课程涵盖了从统计学到研究设计、数据库使用到实操训练等多个方面，还包括教授 R 语言等实用技能，以及更深入的操作。我在授课时，非常重视与学员的互动，通过观察学员的眼神和反应来判断他们是否掌握了知识，并在必要时进行及时指导。面对面的授课方式可以确保学习效果，尤其是对于一些难度较大的内容，还会安排大量的实践练习。只有在现场观察学员的表现，我才能确认他们是否真正掌握了技能。