

AME科研时间系列医学图书007

傻瓜统计学

主编：胡志德 周支瑞



中南大学出版社
www.csupress.com.cn



AME
Publishing Company



丁香园
WWW.DXY.CN

AME科研时间系列医学图书007

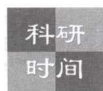
傻瓜统计学



主编：胡志德 周支瑞



中南大学出版社
www.csupress.com.cn



图书在版编目 (CIP) 数据

傻瓜统计学 / 胡志德, 周支瑞主编. —长沙: 中南大学出版社,
2015. 6

ISBN 978 - 7 - 5487 - 1596 - 2

I. ①傻 II. ①胡... ②周... III. ①统计学 - 通俗 - 读物
IV. ① C8-49

中国版本图书馆 CIP 数据核字 (2015) 第 123155 号

AME 科研时间系列医学图书 007

傻瓜统计学

胡志德 周支瑞 主编

☐ 丛书策划	汪道远	昌 兰
☐ 责任编辑	彭敏宁	李 媚
☐ 责任校对	石曼婷	
☐ 责任印制	易红卫	潘飘飘
☐ 版式设计	朱三萍	林子钰
☐ 出版发行	中南大学出版社	
	社址: 长沙市麓山南路	邮编: 410083
	发行科电话: 0731-88876770	传真: 0731-88710482
☐ 印 装	湖南印美彩印有限公司	

☐ 开 本	720×1000 B5	☐ 印张 10.5	☐ 字数 205 千字	☐ 插页
☐ 版 次	2015 年 6 月第 1 版	☐ 2015 年 6 月第 1 次印刷		
☐ 书 号	ISBN 978 - 7 - 5487 - 1596 - 2			
☐ 定 价	39.00 元			

图书出现印装问题, 请与经销商调换

主编：

胡志德 第二军医大学临床检验诊断学在读博士

周支瑞 复旦大学附属肿瘤医院放疗科在读博士

编委：

Daniel J. Sargent

Department of Biostatistics, Mayo Clinic, Rochester, MN, USA

Shi Qian

Department of Health Science Research, Mayo Clinic, Rochester, MN, USA

Kevin Phan

The Collaborative Research (CORE) Group, Macquarie University, Sydney, Australia

David H. Tian

The Collaborative Research (CORE) Group, Macquarie University, Sydney, Australia

Christopher Cao

The Collaborative Research (CORE) Group, Macquarie University, Sydney, Australia

Deborah Black

Faculty of Health Sciences, University of Sydney, Sydney, Australia

Tristan D. Yan

The Collaborative Research (CORE) Group, Macquarie University, Sydney, Australia

邓喜成 湖南省儿童医院心胸外科

朱廷准 沈阳军区总医院神经外科

毛 智 解放军总医院骨科

谷万杰 南京鼓楼医院麻醉科

丛书介绍

很高兴，由AME出版社、中南大学出版社和丁香园网站联合策划的“AME科研时间系列医学图书”，如期与大家见面！

虽然学了四年零三个月医科，但是，仅仅做了三个月实习医生，就选择弃医了，不务正业，直到现在在做医学学术出版和传播这份工作。2015年，毕业十周年。想当医生的那份情结依旧有那么一点，有时候不经意间会触碰到心底深处……

2011年4月，我和丁香园的创始人李天天一起去美国费城出差，参观了一家医学博物馆——马特博物馆(Mütter Museum)。该博物馆隶属于费城医学院，创建于1858年，如今这里已经成为一个展出各种疾病、伤势、畸形案例，以及古代医疗器械和生物学发展的大展厅，展品逾20 000件，其中包括战争中伤者的照片、连体人的遗体、侏儒的骸骨以及人体病变结肠等。此外还有世界上独一无二的收藏，比如一个酷似肥皂的女性尸体、一个长有两个脑袋的儿童的颅骨等。该博物馆号称“The Birth of American Medicine”。走进一个礼堂，博物馆的解说员介绍宾夕法尼亚大学医学院开学典礼都会在这个礼堂举行。当时，我忍不住问了李天天一个问题：如果当初你学医的时候，开学典礼在这样的礼堂召开的话，你会放弃做医生吗？他的回答是：不会。

2013年5月，参加BMJ的一个会议，会议之后，有一个晚宴，BMJ对英国一些优秀的医疗团队颁奖，BMJ的主编和BBC电台的著名节目主持人共同主持这个年度颁奖晚宴。令我惊讶的是，BMJ给每个获奖团队的颁奖词，从未提及该团队在过去几年在什么大牛杂志上发表过什么大牛论文，而是，关注这些团队在某个领域提高医疗服务质量，减轻病患痛苦，降低医疗费用等方面所做出的贡献。

很多朋友好奇地问我，AME是什么意思？

AME的意思就是，Academic Made Easy, Excellent and Enthusiastic。2014年9月3日，我在朋友圈贴出3张图片，请大家帮忙一起从3个版本的AME宣传彩页中选出一个喜欢的。最后，上海中山医院胸外科的沈亚星医生竟然给出一个AME的“神翻译”：欲穷千里目，快乐搞学术。

AME是一个年轻的公司，拥有自己的梦想。我们的核心价值观第一条是：Patients Come First！以“科研(Research)”为主线。于是，2014年4月24日，我们的微信公众号上线，取名为“科研时间”。“爱临床，爱科研，也爱听故事。我是科研时间，这里提供最新科研资讯，一线报道学术活动，分享科研背

后的故事。用国际化视野，共同关注临床科研，相约科研时间。”希望我们的AME平台，能够推动医学学术向前进步，哪怕是一小步！

如果说酒品如人品，那么，书品更似人品。希望我们“AME 科研时间系列医学图书”丛书能将临床、科研、人文三者有机结合到一起，像西餐一样，烹调出丰富的味道，搭配出一道精美的佳肴，一一呈现给各位。

汪道远
AME出版社社长

序一

人类已经进入了“大数据”时代！

医学也已经进入了“大数据”时代！

面对铺天盖地的数据，面对呈指数增长的信息量，医学统计学的作用日益凸显！正确地运用统计学方法，从杂乱无章的数据中提炼出一般性规律，并将其作为重要的循证医学的证据指导疾病的预防、诊断和治疗，是每一名科研工作者和临床医生不容推卸的责任。在“大数据”背景下，当今的医学科研，尤其是临床科研，对统计学的要求已经越来越高，甚至可以毫不夸张地说，到达了前所未有的高度。作为杂志主编和审稿人，我在过去几年为国际杂志审稿的过程中，对此深有体会。很多来自国内的稿件虽然聚焦了一个十分重要的临床问题，也具有较高的原创性，但是由于存在致命的统计学问题，因此不得不忍痛拒稿；作为一名作者，在论文投稿和修改的过程中，我也深切地感觉到不论是编辑还是审稿人，都会十分重视统计学处理过程，几乎不容许作者的统计学处理过程有任何瑕疵。作为一名学者，我也注意到很多论文之所以能“超水平发挥”而被刊登在高水平的杂志上，往往与作者巧妙且合理的统计学处理密不可分。

我曾私下与很多国内同行进行过交流，他们普遍能意识到医学统计学对于医学科研的重要作用，但总抱怨缺乏通俗易懂的统计学书籍，掌握不了学习医学统计学的要领。的确，专业的统计学书籍多出自流行病学和医学统计学人士之手，侧重于理论方面的介绍，而较少关注如何选择统计学方法、如何在统计软件中进行操作、如何解读统计学结果等问题，因此难免出现曲高和寡的情况。鉴于此，我特向广大同行推荐这本《傻瓜统计学》，这本书最大的特点是坚持用最通俗易懂的语言并结合一些实例讲解医学统计学。全书仅仅围绕为什么要选择这种统计学方法而不能选择其他统计方法、如何在软件中进行相关操作、如何正确解读统计结果并结合科研设计和专业知识进行解释等问题。全书文风朴实但不乏深刻、幽默但不失严谨，让读者在欣然一笑之间加深了对医学统计学的认识。

据悉，本书主编和出版社正在策划出版续集，也推荐各位同行关注。

季加孚 教授、主任医师、博士生导师

北京大学肿瘤医院 / 北京肿瘤医院 / 北京大学临床肿瘤学院院长

北京市肿瘤防治研究所所长

教育部恶性肿瘤发病机制及转化重点实验室胃癌研究室主任

中国抗癌协会胃癌专业委员会主任委员

2015年4月27日

序二

这个世界上存在三种谎言：大谎言、小谎言和统计学。

志德兄是一个有趣的段子手，在AME沙龙讲座上他讲了个做学术要善于“吹”的小故事，每每引得广大听众捧腹。现在，他又拉上支瑞，写出了一本怎么教人撒谎的书，并且宣称傻瓜也能学会，这本书应该畅销。

杜克的同事说：当你娶一个中国女孩，其实你娶了两个，因为你的人生中从此多了一个陶华碧。其实，统计学之于医学研究，正仿佛老干妈之于白米饭——或多或少总得来点，但大多数人不知道这货怎么做出来的。对于临床工作者而言，每天看到的许多病例，经过统计学的加工，最后变成了论文里的几个方块、几个矩阵甚至一棵树，中间存在很多神秘的区间。很多同道或许会觉得：知道个大概就行了，凯撒的归凯撒，统计的归统计。当志德和支瑞决定让这本《傻瓜统计学》公之于世的时候，除了希望神作大卖以外，我最大的心愿就是将志德的头像加上一个椭圆形的框加印在封面上，让大家在喜欢上志德的同时，也能够喜欢上统计学，从而改变统计学带给学术研究的一些固有麻烦。

AME的固有观念是“快乐搞学术”，相信有了这本书以后，广大青年学者可以做到“快乐搞统计”。为了做到这一点，作者选择围绕应用实例来说明统计问题，以期还原统计的真面目。我们可以很容易地将《傻瓜统计学》一书与其他“正常人统计学”进行区别，那就是：全书几乎不含数学公式！对于长期身处临床的广大读者而言，简化的中间描述环节给大家带来的好处就是：1)为直截了当地解决统计问题提供了可能；2)控制了本书的字数从而产生更为宽容合理的书价。与此同时，本书还不吝添加了作者处于人生低谷时的片段感悟以鼓励读者，增加了本书的趣味性。显然，这是一本与众不同的书，也许您已经通过不同途径阅读到其中的一部分章节，所以我们希望她的全貌能够在最短的时间内和大家见面：和老干妈一样，Hot and Spicy!

《傻瓜统计学》包含了多位青年学者的学术梦想，而AME则为实现这些梦想提供了宽广的舞台。对于AME而言，可以说我们所贩卖的正是梦想：这本《傻瓜统计学》，正是一本给人“学术开光”，教人如何“圆梦”的书，相信会让您受益匪浅。

Yaxing Shen M.D. Graham fellow

Department of Thoracic Surgery

Zhongshan Hospital, Fudan University

2015年6月6日，于杜克

序三

地球上的医学统计学也就这么简单

笔者在攻读博士期间，学校为阻止学生逃课，不仅在课前课后进行抽查点名，而且还明文规定，逃课者取消考试资格。即便如此，仍有不少玩世不恭的同学以种种不是理由的理由逃课。有的同学虽然迫于“取消考试资格”的压力来到课堂听课，也会拿出事先准备好的专业文献，在课堂上一本正经地学习起来，全然不顾及讲台上早已口干舌燥的老师的感受。

逃课，在英语、政治和一些专业课上俨然是司空见惯的事情了。与之形成鲜明对比的是，医学统计学的课堂总是场场爆满，全然不用老师点名。一般而言，课堂上的场景大致是这样的：老师在讲台上讲得唾沫横飞，下面的学生则把脖子伸得跟长颈鹿似的，目不转睛地盯着老师的课件，生怕漏掉一丁点内容。下课铃声一响，讲台总是被围得水泄不通，各位同学都忍不住拿出事先准备好的笔和纸，在老师面前不停地比划着，请教自己课题中的统计学问题。

即便如此，每年的毕业季，总有不少的急得火烧眉毛的研究生，手拿着自己的毕业论文，不辞辛苦地在校园内外四处寻找所谓的统计学高手帮助自己分析数据。此时，多数研究生总免不了仰天长叹：这地球上的统计学咋就这么难呢？的确，对于医学统计学，很多临床医生和研究生始终无法走出“一学就会，一会就用，一用就错，一错就懵”的怪圈。究其原因，主要是因为很多临床医生在学习统计学时总抱着一副依葫芦画瓢的态度，试图通过套用统计学方法的方式去解决自己遇见的问题，而不去深究这些统计学方法的来龙去脉。

也许有的同行会说：您真是站着说话不腰疼！深究统计学方法的来龙去脉，对我等非统计学专业人士而言，谈何容易？确实，目前市面上有很多统计学专著，多为统计和流行病学专业人士撰写，其内容十分丰富和全面，阐述问题的角度也十分专业和深刻。但问题在于，这类专著多半仅仅适合统计学专业人士之间的交流，对于统计学功底十分薄弱的临床医生而言，却过于深奥。据笔者观察，多数临床医生连标准误和标准差的区别都搞不清楚，哪有什么心思去研究多元线性回归、Bland-Altman分析之类的统计学方法。临床医生更希望得到的是一本通俗易懂、简单明了的统计学参考书。他们不想知道 t 检验背后的统计学原理，但是却十分渴望了解“什么情况下该用 t 检验？如何在统计软件中进行 t 检验？如何解读 t 检验的结果？”等充满实战意义的问题。

长期以来，在临床医生和医学统计学专家之间本身就存在一道“鸿沟”。临床医生深知统计学对于临床研究来说具有不言而喻的重要性，但问题在于，当他们向统计学专家请教统计学问题时，总是不知道统计学专家在说什么？而

统计学专家则认为，自己阐述问题的方式已经简单得不能再简单了。笔者认为，造成这种沟通不畅的原因主要是统计学家过高地估计了临床医生的统计学功底，认为只要稍加点拨，问题便会迎刃而解。而实际情况是，临床医生的统计学功底极其薄弱。

鉴于此，在AME出版社的策划下，两位主编鼓起勇气撰写此书。本书内容不一定全面和丰富，阐述问题的角度也不一定深刻而专业，但是我们在阐述问题时坚持了通俗易懂、简单明了的原则，力争用大家喜闻乐见的方式来描述一些医学统计学问题。本书中的题材来源多为“AME科研时间”（微信号：amegroups）此前曾推送过的内容，以及AME旗下一些学术杂志上刊登的统计学文章。我们从不期望本书能在医学统计领域掀起多大的波澜，毕竟两位主编均不是医学统计学专业出身，我们只是希望本书能对尚未走出“一学就会，一会就用，一用就错，一错就懵”怪圈的同行人有所裨益，让大家看完后，在开怀一笑的同时会若有所思，然后自信地说一句：地球上的医学统计学原来就这么简单啊！

胡志德

2015年4月10日

目 录

丛书介绍/汪道远	II
序一/季加孚	IV
序二/沈亚星	V
序三/胡志德	VI

第一部分：统计杂想篇

第一章：几则有趣的医学统计学故事	1
第二章：有多少 t 检验曾被滥(乱)用	6
第三章：常见的滥(乱)用卡方检验的情况	9
第四章：如何正确认识“ p 值”	14
第五章：浅谈医学科研中计数资料的统计学处理原则	18
第六章：手把手教你解读生存曲线	25
第七章：肿瘤分期中的统计学陷阱	30
第八章：受试者工作特征曲线分析法	43
第九章：临床医生谈统计学应用	53
第十章：生物统计学家：临床研究的幕后高手	56

第二部分：循证医学篇

第十一章：走出循证医学的误区	60
第十二章：正确认识系统评价/Meta分析在医学科研和临床实践中的作用	64

第十三章：学术型外科医生制作系统评价和Meta分析的技术和指南·····	69
第十四章：临床医生该如何高效的检索？·····	87
第十五章：系统评价哪家强？·····	92
第十六章：随机对照试验中那些真假难辨的亚组分析·····	99
第十七章：随机对照试验的批判性阅读·····	104
第十八章：系统评价的idea从哪里来？·····	110
第十九章：顶级医学杂志发表什么样的Meta分析？·····	115
第二十章：亚组分析在Meta分析中的应用·····	120
第二十一章：什么样的文章是一篇好文章？·····	124

第三部分：软件操作篇

第二十二章：如何用GraphPad Prism进行两组连续变量的比较·····	127
第二十三章：手把手教你用GraphPad Prism绘制生存曲线·····	135
第二十四章：如何用Sigmaplot绘制ROC曲线？·····	143
第二十五章：如何画出人生的第一张森林图？·····	151

第一章：几则有趣的医学统计学故事

- 1 两个指标诊断疾病的问题
- 2 降糖药的研究与学生自杀事件
- 3 如何看待统计学结果

医学统计学是一门很奇妙的科学。要说它简单吧，其实也挺简单的，常见的统计方法也就十余种，在教科书上都能找到，只要熟练掌握了，虽不敢夸下海口说“以秋风扫落叶的气势横扫四海之内的杂志”，但足以轻车熟路地应付99%的科学研究。要说它复杂吧，也挺复杂的，毫不夸张地说，在绝大部分国内期刊，甚至在很多低分SCI杂志上，乱用统计学的现象多如牛毛。本文拟谈几则与医学统计学相关的故事，希望能帮助大家从宏观上正确认识医学统计学这门科学。

1 两个指标诊断疾病的问题

路人甲做了一个研究，旨在比较两个指标(A和B)对肝癌的诊断价值。路人甲以A和B的参考范围上限作为诊断界值，得出了A和B在该界值下对应的诊断敏感性和特异性。结果表明：A的诊断敏感性为0.80，特异性为0.90；B的诊断敏感性为0.85，特异性为0.87。路人甲很快撰写论文报道了自己的研究成果，指出B诊断肝癌的敏感性高于A，而特异性低于A。

路人乙是这篇文章的审稿人，当他看见这个结论后，脸色铁青，毫不犹豫地在审稿意见中写道：就敏感性而言，B高于A；就特异性而言，A高于B。诊断敏感性和特异性与所采用的界值密切相关，作者得出的敏感性和特异性仅仅代表了一个诊断界点下面的诊断效能，无法反映A和B的总体诊断价值。文章的结论到底是想说明A优秀还是B优秀呢？拒稿！

这个故事说明：统计指标选错了，统计出来的东西往往难以“自圆其说”。

稿件被退了，路人甲有些许郁闷。经过认真学习科研设计与统计学知识后，路人甲终于明白了一个问题：两个指标诊断性能的比较是不能比较敏感性和特异性的，而应该比较受试者工作特征(ROC)曲线的曲线下面积，因为曲线下面积才是衡量整体诊断效率的最佳指标。路人甲很快绘制了ROC曲线，统计结果表明：A的曲线下面积为0.80，B的曲线下面积为0.82。路人甲欣喜若狂，赶紧动笔写论文，并且理直气壮地给文章定了一个结论：B的总体诊断效率是优于A的，其理由就是因为B的曲线下面积大于A。

路人丙是这篇文章的审稿人，当他看见这个结论后，脸色铁青，毫不犹豫地 在审稿意见中写道：从表面上看，B的曲线下面积高于A，但导致这种差异的原因有抽样误差和试验效应两种。你怎么能确定这不是抽样误差呢？在统计学上，要确定0.82是否高于0.80，就一定要经过统计学检验。拒稿！

这个故事说明：在医学科研中，没有经过统计学检验的结论多半是不科学的。

稿件被退了，路人甲很是郁闷。他吸取了经验教训，自学了很多统计学理论，终于弄清楚了采用何种方法去比较曲线下面积。接下来的事情就是改稿，然后另选杂志继续投稿。路人甲在文稿中特别注明了，曲线下面积是经过统计学检验的，B的曲线下面积(0.82)与A的曲线下面积(0.80)之间的差异是有统计学意义的，而且还大摇大摆地在后面加了个括号，写明 $p=0.01$ 。路人甲仰天长叹了一口气，很郑重地给自己的研究下了结论：本研究表明B的诊断效率是优于A的。

路人丁是这篇文章的审稿人，当他看见这个结论后，脸色铁青，毫不犹豫地 在审稿意见写道：B是常见的诊断指标，其检测结果并不对临床医生设盲，在很大程度上可以影响临床医生对疾病的诊断。A是新近发现的诊断指标，其结果完全对临床医生设盲，不可能影响医生的诊断。在此条件下比较A和B的诊断效率显然是不“公平”的。作者的结论(B比A优秀)是不可靠的。再说得通俗点，如果把A和B分别理解成法庭上的原告和被告，那B无疑既充当了当事人，又充当了法官的角色。在这种情况下，A输掉了官司是十分正常的。如果换一个公平的、独立的法官来断案，B能否胜出就不好说了。拒稿！

这个故事说明：实(试)验设计有缺陷，再优秀、再正规的统计学方法也于事无补。

稿件又被退了，路人甲的心情极度郁闷。想来想去，决定把实验重做一遍，让A和B在一个公平的环境中比较(为便于描述，此处忽略医学伦理学问题)。在新开展的研究中，A和B都是对临床医生设盲的，不可能影响金标准。这下A和B的比较结果应该比较可靠了吧，路人甲又仰天长叹了一口气，感觉

自己如释重负了。科研太折腾人了，太不容易了！统计结果很快出来，A的曲线下面积是0.80，B的曲线下面积则变成了0.77，经过统计学检验后发现：A的诊断效能确实是高于B的($p=0.01$)。整个研究的试验设计滴水不漏，统计学过程天衣无缝，我就不信还有人敢拒这篇稿件，路人甲心中暗喜。

路人戊是这篇文章的审稿人，当他看见这个结论后，脸色铁青，毫不犹豫地审稿意见写道：A和B的检测并不矛盾，他们之间的关系不应该是竞争关系，而应该是合作关系。读者最关心的问题显然不是A和B“孰强孰弱”的问题，虽然这个问题有一定的专业价值。如果我是坐诊医生，我会说：A和B谁强谁弱关我什么事？总之来一个患者我就A和B都检测，我的患者都不差钱！作者的研究重点应该是明确A和B能否互补，联合使用是否有助于提高诊断准确性的问题，而非A和B“孰强孰弱”的问题。简单地说，就是明确1+1是否大于1的问题。

文章又被拒稿了。

这个故事说明：研究方向错了，即使是无懈可击的实(试)验设计和天衣无缝的统计方法，也是无济于事。

这四个故事说明：医学科研是很痛苦的，不重视统计学和科研设计，会走很多弯路的。

2 降糖药的研究与学生自杀事件

路人甲长期从事降糖药的研究，最近他发现了一种药物，可以降低患者的血糖。为了评价该药的降糖效果，路人甲费尽心机地设计了一个看似完美的随机对照试验(randomized controlled trial, RCT)，为了保证结果可靠，路人甲严格遵守RCT设计规范，包括随机、双盲、安慰剂对照等措施。研究结果表明：实验组和对照组在接受药物治疗前血糖浓度的均值都是10 mmol/L，差异无统计学意义，表明两组研究对象的基线血糖水平具有可比性。对照组未经任何药物治疗(为便于描述，此处忽略医学伦理学问题)，血糖浓度变化幅度为0 mmol/L，实验组经过药物干预后，血糖浓度平均下降了1 mmol/L。统计学检验结果表明，实验组和对照组血糖降低幅度的差异是有统计学意义的。($p<0.01$)。路人甲赶紧撰写论文，并毫不犹豫地给研究下了个结论：该药可以降低患者血糖。

路人乙是这篇文章的审稿人，当他看见这个结论后，脸色铁青，毫不犹豫地审稿意见写道：该药确实可以降低血糖，但是一个降糖幅度只能达到1 mmol/L的降糖药有何临床价值？拒稿！

这个故事说明：有统计学意义不一定有专业意义。

路人甲做了一个调查，同处一地的A和B两所中学，各有1 000名学生，过

去的一年，A校有5名学生自杀(自杀率为0.5%)，B校没有学生自杀(自杀率为0%)。统计学结果表明：两校自杀率的差异无统计学意义($p=0.07$ ，Fisher确切概率法，笔者进行了统计)。于是路人甲得出结论：A和B两校的自杀率的差异没有统计学意义，A校5名学生自杀纯属小概率事件。

路人乙是这篇文章的审稿人，当他看见这个结论后，脸色铁青，毫不犹豫地在审稿意见写道：5个鲜活的生命就这样没有了，5个家庭就这样毁了，你却告诉我这纯属小概率事件，你就不怕“人神共愤”吗？拒稿！

这个故事说明：有专业意义不一定有统计学意义。

这两个故事说明：做医学科研，不能死磕统计。

看完这两个故事，也许有的读者会有疑问：前面还强调“没有经过统计学检验的结论多半是不科学的”，为什么这里却淡化统计学的作用呢？对此，笔者认为：统计学仅仅是一种工具，用得好当然可以事半功倍，但是在某些情况下，工具往往就是个累赘，也许徒手干活才是最好的选择。

3 如何看待统计学结果

路人甲经历数十年的研究，动用了各种高精尖的研究手段，发现了一个新的蛋白(命名为蛋白A)。在肝癌患者中展开的研究表明：蛋白A和甲胎蛋白(α -fetoprotein, AFP)有很好的相关性，($r=0.99$, $p<0.0001$)。路人甲欣喜若狂，尽管蛋白A的检测过程还十分繁琐，检测费用还十分高，但是路人甲还是把持不住内心的激动，日夜兼程地撰写论文，宣称自己找到了一个新的肝癌标志物。

路人乙是这篇文章的审稿人，当他看见这个结论后，脸色铁青，毫不犹豫地在审稿意见写道：统计结果表明蛋白A和AFP的相关性十分明显。如果是这样，在临床实践中，通过检测AFP完全就可以得知蛋白A的浓度了，蛋白A在肝癌中的临床价值完全可以被AFP代替，还不说蛋白A的检测过程繁琐，费用太高的问题，你说蛋白A还有什么价值？拒稿！

这个故事说明：统计学阳性的结果未必是“好结果”。

路人甲发明了两套诊断肺癌的方案，分别命名为A和B。为了明确这两种方案到底谁“更胜一筹”，路人甲找了100个肺癌患者和100个疑似肺癌患者(结核、肺炎等)，分别用A、B两套方案去进行鉴别诊断。在200个研究对象中(100个肺癌和100个非肺癌)，方案A能准确地对其中的100个人进行鉴别。准确率50%，方案B仅仅能对其中的50个人进行准确的鉴别，准确率仅为25%。卡方检验表明：方案A和B准确率之间的差异有统计学意义($p<0.01$)。很明显，方案A的准确性要高于方案B。路人甲赶紧发表论文，指出：方案A诊断肺癌的准确

性优于方案B。

路人乙是这篇文章的审稿人，当他看见这个结论后，脸色铁青，毫不犹豫地在审稿意见写道：如果我(审稿人本人)是坐诊医生，他就会反着看方案B的结果，凡是方案B认为是肺癌的，他就认为患者不是肺癌；反之亦然。这样下来，在200个患者中，方案B应该能正确识别150个人，准确率为75%，统计学结果表明：方案B的准确率(75%)是高于方案A(50%)的，所以真实的情况是方案B优于方案A。实际上，当面对这200名患者的时候，随便到城隍庙找个盲人来“算命(猜患者是否患病)”，按照统计学理论，准确率也应该是50%，方案A可以说是“一无是处”。拒稿！

这个故事说明：统计学阴性的结果未必是“坏结果”。

还是那个AFP与蛋白A的例子。路人甲发现蛋白A和AFP之间有很好的相关性，也开始撰写论文，但是他的结论并不是“蛋白A是诊断肝癌的标志物”。他认为，既然蛋白A与AFP之间有很强的相关性，那提示AFP和蛋白A之间可能存在十分密切的“调节”或者“被调节”关系。最终，路人甲围绕“蛋白A与AFP之间的调节或者被调节关系”做了很多分子生物学试验，指出蛋白A是调节AFP表达的唯一(注意“唯一”这两个字)因子，因此二者之间才会呈现如此强烈的相关性。这是一项基础研究，虽然未能直截了当地指出蛋白A的临床价值，但是这个研究形象生动地讲述了一个完整的分子生物学事件，丰富了我们对于肝癌发生与发展分子机制的认识。最终，论文“堂而皇之”地被接受了。

这个故事说明：同一统计学结果，从不同的专业角度去解释，结论是完全不同的。

这三个故事说明：对统计学结果的解读一定要结合专业！从专业中来，到专业中去！

(胡志德)