

序 Ben Goldacre——《Bad Science》一书的作者

验证疗效的科学试验

高质量研究促进高质量卫生保健

第二版



Imogen Evans Hazel Thornton Iain Chalmers Paul Glasziou

验证疗效的科学试验

欲购买英文版《验证疗效的科学试验》平装本，
可登陆 Pinter & Martin 网站 (www.pinterandmartin.com)。

付款时输入密码“TT25”，
Pinter & Martin 网的所有书籍
可享受 25%的折扣并免邮费及包装费（限英国）。

验证疗效的科学试验

高质量研究促进高质量卫生保健

第二版

**Imogen Evans, Hazel Thornton,
Iain Chalmers, Paul Glasziou**

序 Ben Goldacre



谨以此书献给 **William Silverman (1917-2004)**，一位不断鼓励我们挑战权威的长者。

验证疗效的科学试验
高质量研究促进高质量卫生保健

第一版于2006年由大英图书馆首次出版
第二版于2011年由Pinter & Martin有限公司首次出版

版权©2011 Imogen Evans, Hazel Thornton, Iain Chalmers 和 Paul Glasziou
序©2011 Ben Goldacre
第一版序© 2006 Nick Ross

本书所有作者的著作权参照了《1988年版权、设计和专利法》

版权所有

大英图书馆图书在版编目 (CIP) 数据核字
本书的目录记录可通过大英图书馆获取

ISBN 978-1-905177-48-6

本书的出售需符合以下条件：未经出版商事先许可，不得通过买卖及其他方式以任何形式或出版形式以外的装订和封面，借出、转售、出租本书或进行传播；且不得通过类似的形式欺骗消费者。

由TJ国际有限公司（英国，康沃尔郡帕德斯托）印刷装订

本书出版用纸来源于可持续森林并经FSC认证。

Pinter & Martin有限公司
埃弗拉 帕若德 6号
伦敦 SW2 1PS

www.pinterandmartin.com

验证疗效交互平台：www.testingtreatments.org

目录

关于作者

致谢

序言—Ben Goldacre

序言—Nick Ross

前言

我们为什么会写《验证疗效》这本书

引言

- 1 新的不一定更好
- 2 预期的效果没有变为现实
- 3 越多未必越好
- 4 越早未必越好
- 5 应对疗效的不确定性
- 6 正确的验证疗效的试验
- 7 考虑偶然因素
- 8 评价所有相关的可靠证据
- 9 对正确疗法的试验监管：是推动还是阻碍？
- 10 高质量、低质量和不必要的研究
- 11 做高质量研究，人人有责
- 12 那么，什么让卫生保健变得更好？
- 13 合理的研究：未来更美好的蓝图

参考文献

附加资源

插图列表

要点列表

索引

作者简介

Imogen Evans 曾在加拿大和英国行医、讲学，后加入《柳叶刀》(The Lancet) 杂志从事医药新闻行业。1996-2005 年，她就职于英国医学研究理事会(Medical Research Council, MRC)，后从事伦理学研究，曾任欧洲生物医学伦理委员会理事会(the Council of Europe Biomedical Ethics Committee) 的英方代表。

Hazel Thornton 曾在接受常规乳房 X 线检查后，应邀参与了一项临床试验，但因没有得到足够的知情同意而拒绝参加。然而，此事激励她倡导公众参与研究，从而获得与患者相关的研究结果。她经常以书面和口头形式发表对此话题的看法。

Iain Chalmers 曾在英国和巴勒斯坦行医，后成为一名卫生服务研究者，曾负责英国牛津大学国家围产期流行病学研究所及英国 Cochrane 中心。自 2003 年以来，他领导开展了“詹姆斯·林德计划 (James Lind Initiative)”，提倡用更高质量的对照研究实现更好的卫生保健，特别是通过更广泛的公众参与。

Paul Glasziou 既是一名医学研究者，也是一名兼职的全科医生。在意识到医学研究和实践之间的差距后，他便开始致力于寻找并解决在日常临床实践中应用高质量研究的障碍。他曾任《英国医学杂志》(BMJ) 旗下《循证医学杂志》(Journal of Evidence-Based Medicine) 编辑，并于 2003~2010 年担任牛津大学循证医学中心主任。他曾参与编写数本有关循证实践的专著。自 2010 年 7 月起，他就职于澳大利亚邦德大学 (Bond University)，并入选澳大利亚国家卫生医疗研究委员会会员 (National Health and Medical Research Council Australia Fellowship)。

致谢

我们感谢以下人员对本书（Testing Treatments）第二版的编写提出的宝贵意见和所做的其他贡献：

Claire Allen, Doug Altman, Patricia Atkinson, Alexandra Barratt, Paul Barrow, Ben Bauer, Michael Baum, Sarah Boseley, Joan Box, Anne Brice, Rebecca Brice, Amanda Burls, Hamish Chalmers, Jan Chalmers, Yao-long Chen, Olivia Clarke, Catrin Comeau, Rhiannon Comeau, Katherine Cowan, John Critchlow, Sally Crowe, Philipp Dahm, Chris Del Mar, Jenny Doust, Mary Dixon-Woods, Ben Djulbegovic, Iain Donaldson, George Ebers, Diana Elbourne, Murray Enkin, Chrissy Erueti, Curt Furberg, Mark Fenton, Lester Firkins, Peter Gøtzsche, Muir Gray, Sally Green, Susan Green, Ben Goldacre, Metin Gülmezoğlu, Andrew Herxheimer, Jini Hetherington, Julian Higgins, Jenny Hirst, Jeremy Howick, Les Irwig, Ray Jobling, Bethan Jones, Karsten Juhl Jørgensen, Bridget Kenner, Debbie Kennett, Gill Lever, Alessandro Liberati, Howard Mann, Tom Marshall, Robert Matthews, Margaret McCartney, Dominic McDonald, Scott Metcalfe, Iain Milne, Martin McKee, Sarah Moore, Daniel Nicolae, Andy Oxman, Kay Pattison, Angela Raffle, June Raine, Jake Ranson, James Read, Kiley Richmond, Ian Roberts, Nick Ross, Peter Rothwell, Karen Sandler, Emily Savage-Smith, Marion Savage-Smith, John Scadding, Lisa Schwartz, Haleema Shakur, Ruth Silverman, Ann Southwell, Pete Spain, Mark Starr, Melissa Sweet, Tilli Tansey, Tom Treasure, Ulrich Tröhler, Liz Trotman, Liz Wager, Renee Watson, James Watt, Hywel Williams, Norman Williams, Steven Woloshin, Eleanor Woods, and Ke-hu Yang.

Iain Chalmers 和 Paul Glasziou 对英国国立卫生研究院（National Institute for Health Research, NIHR）的支持表示感谢。Paul Glasziou 对澳大利亚国家卫生医疗研究委员会的帮助表示感谢。

此外，特别要感谢我们的出版人——Pinter & Martin 的 Martin Wagner，感谢他一直以来的宽容大度、乐观的鼓励和自始至终冷静的头脑。

序

医学不应该被权威所左右，每个人都可以在这点上提出一个简单却至关重要的问题：“你是怎么知道的？”本书的内容就回答了这个问题。

医务工作者与患者的交流方式已经发生了巨大改变。很久以前，类似于“医患沟通技能的培训”中就包括了医生如何隐瞒他们的患者将死于癌症的事实。而今天，我们教育医学生——让我们引用培训材料上的原话——如何“与患者紧密协作，以达最佳健康结局”。现在，如果患者愿意，那么最佳的医疗实践模式应做到让患者参与讨论和选择他们自己的治疗方法。

要做到这一点，关键在于让每个人都要明白，我们是如何知道某种疗法有效的或是有害的，我们如何权衡利弊以防范医疗风险。遗憾的是，医生可能也像其他大多数人一样对此知之甚少。更糟的是，这其中陷阱重重，诱导我们误入歧途。

这一系列陷阱中最先最可怕的是自我误导。大多数疾病都有自然史，周期性或随机地好转或加重。所以如果你正好在症状最严重的时候去治疗它，那么就会感觉是治疗发挥了作用，因为此时无论你做什么，症状都会好转。

同样，安慰剂效应能让我们都上当：在一些实例中，仅仅服用一片不含任何有效成分的药片的确会缓解病情，因为患者相信他们接受的治疗措施是有效的。正如 Robert M Pirsig 在《万里任禅游》（*Zen and the Art of Motorcycle Maintenance*）中所言：“科学方法的真正目的，在于防止让自然现象误导你理解了其实自己不理解的事情”。

然而总有那么一些人拿科学研究当幌子。如果这本书要向你传达一个关键信息，那一定是“正确试验”的概念，这也是我自己不断借用和使用的一个短语。并非所有的试验都是一样的，因为一项科学研究可能受到的偏倚太多了，而这往往导致人们错误地将自己的主观理解作为“正确”答案。

有时候证据可能会因疏忽大意或单纯的动机（任何动机都有可能产生影响）而被歪曲。医生、患者、专家、护士、专业理疗师以及管理者都应将这个想法深埋心中：只有真正的治疗才值得他们投入巨大的精力，同时它也是弥足珍贵的。

有时候证据会由于其他原因被歪曲。我们不应陷入肤浅的“药厂阴谋论”中：毕竟他们带来了许多重大和救死扶伤的进步。然而有些研究涉及巨大的经济利益，您将会在本书中得知具体原因，90%的试验都是由药厂资助。当药厂资助的研究与具有独立资助的研究相比，前者得到药物有效的结果的概率是后者四倍时，这就有问题了。一种新药上市的成本高达八亿美元（其中大部分用于新药上市前的阶段），如果某种药物最终被证明无效，那么花出去的钱就是覆水难收了。正因为风险如此之高，所以有时候无法达到正确试验的理想状态。¹

同样，证据在传播的过程也有可能出现偏颇，产生误导。有时这种情况反映在事实与数据的展示方式上，只展示部分事实，粉饰假象，通过对科学证据“精挑细选”，让一种治疗方法得以粉墨登场。

但在主流文化里，我们还可以看到更有趣的现象在上演。我们可以理解人们对“灵丹妙药”的期望，但通常情况下，医学研究对提高疗效、降低风险和解决争议的作用一般。但对于媒体宣传而已，这种理性的认识通常会被其以“治愈”、“奇迹”、“希望”、“突破”和“受害者”等诸如此类的词汇给挡在门外。²

当人们如此强烈地希望为自己的生命做主并参与自己的卫生保健决策时，我们却很遗憾地发现存在着大量被歪曲的信息，这只会是对患者权利的剥夺。有时候这种歪曲体现在某一具体的药物上：英国媒体将赫赛汀描述为治疗乳腺癌的奇迹或许就是一个最引人注目的近期实例。³

有时候，为推广他们的疗法并挑战那些反对他们的证据，媒体界的狂热分子及其盟友会造成更大的危害：想方设法影响公众对如何知道事物对自身有益还是有害的正确理解。

大量设计科学的高质量研究比较了顺势疗法糖丸与模拟糖丸，前者疗效并不比后者好。可当面对这一证据时，顺势疗法医生狡辩说，上述试验的整个理念是有问题的，他们还找出一些复杂的理由来说明为什么他们的药丸在所有药丸中是独一无二的，因此无法通过试验验证。而当有证据表明政治家们引以为豪的防止青少年怀孕的教育方案是失败的时，他们也将做出相同的举动。事实上，正如本书所讲，任何宣称有效的干预都能通过透明的科学研究进行验证。⁴

有时候，这些对事实的歪曲甚至会对公众的理解认知产生更深层次的危害。最近一项基于最科学、偏倚最小的试验制作的“系统评价”显示，没有证据支持服用抗氧化维生素片能延长寿命（实际上，他们甚至可能缩短寿命）。这本书中精彩演绎的就是根据类似方法，让人们遵循明确的规范，知道应该去哪里寻找证据，应该选择哪些证据，如何评价这些证据的质量。但是，当系统评价的结果跟抗氧化补充剂药厂的声明相悖时，报纸及杂志中往往会充斥着大量批评言论，认为该系统评价中的个别研究都是出于政治需要或是赤裸裸的腐败原因而被“精挑细选”出来的，真正有用的证据被故意忽略等等。⁵

这真令人遗憾。系统评价的理念——搜集此前所有的证据——可以说是过去30年里医学领域最重要的创新性成果之一。但那些记者和药厂为了维护自己的蝇头小利，阻挠公众触及这些理念，严重损害了大众的利益。

问题就出现在这里。你有很多理由来阅读本书。从最简单的层面说，它会帮助你以更加明智的方式为自身健康做决策。如果你是医务工作者，本书随后的内容很可能会超越你所接受过的任何循证医学教育。从公众的层面上说，如果越来越多的公众开始理解如何作出客观的比较，并能辨别一种疗法是否真的优于其他时，那么就会如作者所言，公众不再畏惧研究，相反，他们有可能成为其中活跃的一份子，更多地参与其中，降低那些与自己切身相关的治疗的不确定性。

而阅读本书还有最后一个原因——读者可以了解药厂的一些小把戏，不过这基本上没有什么实用性。事实上，这些把戏只是有点搞笑、花哨和颇有些小聪明而已。本书比我所见过的所有同类书籍更为卓越，因为它的作者阅历丰富、知识渊博且怀有一颗慈悲之心。

《验证疗效》一书将人类的思考带入到现实的问题中。医学不仅和人类的苦难和死亡有关，还和决策者及研究人员的人性弱点相关——这在本书中体现得尤为淋漓尽致，研究者自己的故事和疑惑，他们的动机、想法以及观念的转换。这种大众能够如此接触科学的情况实属罕见。作者也显得游刃有余，时而说到严肃的学术论文，时而谈及转瞬即逝的医学文献，并从学术论文、述评、自传及随笔的讨论中抽丝剥茧，寻找遗落的珍宝。

这本书应配发给所有学校和患者候诊室。而目前，它在您的手中。往下读吧！

Ben Goldacre

2011年8月

第一版序

本书内容有益于我们的健康。它揭开事关生死的决策是如何作出的谜底，指出这些决策常存在的严重缺陷，并引发全球医生修正其临床实践方式的挑战。

本书不会引起不必要的恐慌；它非常肯定现代医学所取得的诸多成果。其目的始终是改善医疗实践，而非诋毁。

我本人最初洞察到医学中根深蒂固的草率是在 20 世纪 80 年代。当时我应邀作为一名业外人士加入评价乳腺癌最佳疗法的共识委员会。我感到非常震惊（当您阅读了本书的第 2 章[现在是第 3 章]后，您也许会同样震惊）。我们应用的证据来自一流的研究者和临床医生，但我们发现一些最有名的专家往往仅凭直觉或彻头彻尾的偏见来决定治疗方案。女性患者生存的机会及因手术毁容的可能性，很大程度上取决于治疗她的医生及其偏见。例如，一位外科医生喜欢大刀阔斧的切除，另一位倾向于简单的肿块切除，而第三位则选择具有侵袭性的放疗等。似乎这个科学评价的时代与他们相距甚远。

事实上，此类情况过去常常存在，很多医生一直这样行医。尽管现在已有所改善，但许多有才华、坦诚且技艺精湛的医生竟然不知道什么才是好的科学证据。他们各行其事，因为那是他们在医学院学到的，或因为其他医生也这么做，或因为这在他们既往经验中行之有效。但个人经验，尽管看似正确，却往往是可怕的误导——正如本书一针见血揭示的一样。

有些医生认为将科学的严谨性应用到个体患者的治疗中相当幼稚。他们声称，医学既是一门科学也是一门艺术，医学似乎很崇高，其实自相矛盾。医学知识有限，而任何个体的复杂性几乎无限，因此始终存在不确定因素。在临床实践中，好的治疗通常需要好的推测。但过去，许多医生常常混淆了推测与良好证据的区别。有时尽管存在很大的怀疑，他们仍然轻言“确定”。他们避开使用可靠数据，因为他们不知道应该如何对此评价。

本书介绍了更复杂但能更好地区分治疗有效或无效、安全或危险的方法，也阐述了该方法与个人经验的差别。本书尽量避免使用专业术语，采用通俗易懂的表达方式，如“正确试验”等。该书旨在提醒人们，科学像人类一切其他事物一样，容易出现误差和偏倚（因为错误、自负或尤其是医学中的毒瘤——商业需求）；但同时我们也应认识到，尽管如此，在人类知识领域中，正是严谨的科学方法创造几乎所有最突出进步。医生（包括像我这样的“媒体医生”）应该停止诋毁临床研究，不要再视其为“人类豚鼠试验”；相反，应从道义上要求所有医务工作者推动为患者而进行的科学试验，并鼓励患者参与其中。

对那些关注自身或家人健康或卫生政策的人而言，这是一本很重要的书。患者不仅是卫生保健的接受者，也是参与者。普通公众和医生及研究人员一样，都肩负着同样的使命，因为我们既是卫生保健的接受者，也是卫生保健的支付者。如果我们只是医学被动的消费者，将永远无法提高医疗服务的标准；如果我们只喜欢最简单的答案，得到的将是伪科学；如果我们不促进对疗法的严格验证，得到的将是无效有时甚至是有害的治疗，而我们将与真正有效的治疗擦肩而过。

本书以患者为中心，以改善医疗实践为目的，但对医生、医学生以及研究人员也同样重要，他们都可从中有所收获。最理想的是，能够强制每个新闻工作者阅读，并提供给每一位

患者，因为如果医生不充分权衡科学证据，一切都会变得更糟，要知道我们的生死与这些科学证据休戚相关。

我可以保证：如果科学试验对您还是一个全新的话题，一旦读完本书，您对医生的建议的感受会和以往大不相同。

Nick Ross
广播电视节目主持人和记者
2005年11月16日

前言

2006 年第一版《验证疗效》的出版，是受到这样一个问题的启发：“你如何保证医疗中的研究最大程度上满足了患者的需求？”我们共同的经验发现研究常常无法解决这个关键的问题，当时我们团队的成员包括 Imogen Evans（临床医生，曾是一名研究人员及新闻工作者）、Hazel Thornton（患者，提倡研究和卫生保健质量的独立非专业人员）以及 Iain Chalmers（卫生服务研究人员）。此外，我们敏锐地觉察到许多疗法，无论新旧，均未基于良好的证据。所以我们开始着手编写本书，期望通过鼓励患者与专业人员的对话，促使对治疗效果进行更加严格的大众评审。

我们为本书的受关注程度而倍感鼓舞，无论是在大英图书馆内印有版权标记的书稿，还是在网站（www.jameslindlibrary.org）供大家免费下载的电子版，都吸引着普通读者与专业人员。许多国家都已经将第一版《验证疗效》作为教学参考书，而一些语种的全译本也可以从网站（www.testingtreatments.org）免费下载。

从一开始，我们就认为《验证疗效》是一项不断进行的工作。治疗效果，无论新旧，几乎总有不不确定性，因此有必要对所有治疗方法持续进行恰当验证。为满足上述要求，非常有必要查找和再查找证据，并在开始新研究之前严格、系统地评价现有证据，当然，也应将新的研究结果通过不断更新的系统评价解读并公之于众。

在着手编写第二版《验证疗效》时，有了 Paul Glasziou 的加入，我们由原先的三个人变为四个人。Paul 是一名全科医生及研究人员，致力于将高质量的研究证据应用到日常的临床实践中。我们新的出版商——Pinter & Martin，曾于 2010 年将第一版的《验证疗效》再版发行，而且新书的电子版和以前一样可以从网站（www.testingtreatments.org）免费下载获得。

尽管两版的基本前提仍然一致，但我们对第一版进行了大刀阔斧的修改与更新。例如，我们将有关筛查利弊的篇幅扩展成单独一章（第 4 章），题为“越早未必越好”。而在“对正确疗法的试验监管：是推动还是阻碍”（第 9 章）这一章里，我们介绍了过度监管的研究是如何危害患者的。当看到倒数第二章（第 12 章）的时候，我们不禁要问：“怎样改善卫生保健？”，这一章还告诉我们如何将系列证据以真正有益于所有人的方式整合在一起。在全书最后我们提出了未来更美好的蓝图和行动计划（第 13 章）。

我们希望本书会让更广泛的群体了解到应该如何正确验证治疗方法，以及每个人应该如何参与到整个过程。本书并非针对个体治疗效果的“最佳治疗指导”，而是强调一些根本的问题，即如何确保研究在开展之前应有充分依据，能够被正确实施，能够区分出疗效的利弊，并且能够用来回答患者、公众和卫生专业人员真正关心的问题。

**Imogen Evans, Hazel Thornton,
Iain Chalmers, Paul Glasziou
2011 年 8 月**

引言

“我们无法知道何时可以完整全面地观察到自然界的复杂事物。Karl Popper 强调，知识有限，无知无涯。在医学领域，我们永远无法确定治疗效果，只能尽量减少治疗的不确定性。但这种认识并非像它听起来那样悲观：经得起多次有力挑战的结论后来常常被证明是十分可靠的。这种‘起作用的真理’是有着相当坚固结构的建筑模块，支持着我们每天的床旁实践。”

William A. Silverman. *Where's the evidence?*
Oxford: Oxford University Press, 1998, p165

现代医学已经取得了巨大的成功。我们很难想象如果没有抗生素，生活会变成什么样。其他有效药物的研发为心脏病和高血压的治疗带来了革命，还使精神分裂症患者生活发生了转变。儿童疫苗接种让脊髓灰质炎和白喉在大多数国家中成了遥远的记忆，人工关节让不计其数的人免于病痛及残疾的侵扰。现代影像学技术，如超声、计算机断层扫描（CT）以及核磁共振成像（MRI）等，能够确保患者得到精确的诊断并接受正确的治疗。以前许多类型的癌症一旦确诊就意味着宣判死刑，但如今患者能够在得了癌症后继续生存，而非死亡。另外，艾滋病已在很大程度上从一种急性杀手转变为慢性（长期）疾病。

当然，医疗卫生领域取得的丰硕成果是因为社会及公共卫生的进步，如清洁的自来水，更好的卫生设施及居住环境。但即便是怀疑论者也无法忽视现代医疗保健所带来的深刻影响。在过去的半个多世纪里，更好的卫生保健为延长寿命作出了主要的贡献，并改善了生活质量，特别是慢性疾病患者的生活质量。^{1,2}

但现代医学所取得的辉煌成就往往会导致我们忽视许多其持续存在的问题。即便在今天，许多医疗决策仍是基于低质量的证据。依旧有太多的治疗方法对患者有害，还有一些治疗几乎或是根本未被证实有益，而其他一些值得推广的治疗却很少被使用。既然每年关于疗效的研究会产出堆积如山的结果，那么为什么还会出现上述情况呢？令人遗憾的是，这些证据往往是不可靠的，而且大量的研究并没有解决患者想要知道的问题。

部分问题在于医学中很少有治疗后效果极其显著或突出的。而新疗法能起到多大作用常常充满不确定性，甚至不确定它们是否利大于弊。因此，有必要通过精心设计的正确科学试验——即那些在设计之初就减少偏倚且将偶然因素考虑在内的试验（见第 6 章）——来可靠地确证疗效。

18 世纪美国政治家本杰明·富兰克林（Benjamin Franklin）提出了以下著名论断：“这个世界上除了死亡和税收外，没有什么确定的。”³而当一个人患上某种疾病或者接受某种治疗时，同样无法精确预测结果，这有时被称为“富兰克林法则”。然而，富兰克林法则并非这个社会的第二天性。不确定性的必然性并没有在学校教育里被充分强调，其他的基本概念也是如此，例如如何获得并解读证据，如何理解可能性和风险的信息等。正如一名评论家指出：“在学校里，你学到的是试管里的化学、描述运动的公式和光合作用的相关知识。但你十有八九不会学到有关死亡、风险、统计和那些能够杀死或治愈你的科学的东西。”⁴然而，尽管那些基于可靠的科学证据的医疗实践拯救了无数人的生命，但你仍然无法在任何一家科学博物馆里找到哪怕一件能够解释科学研究关键原则的展品。

不要过于确定

“通过探索我们可以更好地了解事物。但对于确定的真相却无人知晓，因为这一切都只是由推测编织成的网。”

Xenophanes, 6th century BCE

“我一直坚信所有事物都存在着不同的观点。”

Charlie ('Peanuts') Brown, 20th century CE

“我们的许多错误表明因果推理这种实践方法……仍然是一种艺术。尽管通过掌握分析技巧、统计学方法和规则以及逻辑标准可以帮助我们，但最终我们的结论仍然逃不出主观判断。”

Susser M. Causal thinking in the health sciences.

Oxford: Oxford University Press, 1983.

不确定性和风险的概念的确相当重要。例如，“证明为阴性结果”意味着逻辑上的不可能性，指证明某个事物不存在，或某种治疗方法没效果。这不仅仅是一种哲学论证，它还有着重要的实践价值。此处以一种名为盐酸双环胺的复合药物（有效成分是抗敏安和吡哆醇或维生素 B6）为例。盐酸双环胺（商品名是 **Debendox** 或 **Diclectin**）曾被广泛用于减轻女性妊娠早期的恶心症状。而后有报道称盐酸双环胺会导致婴儿先天缺陷，这立即引发了雪崩似的诉讼。1983 年，迫于这些法庭诉讼的压力，盐酸双环胺生产商撤回了所有上市的药物。但随后几篇对所有证据的评价结果显示，没有证据支持该药物与先天缺陷有关。虽然我们并不能对药物没有不良反应下定结论，但的确没有证据表明它会对人体造成伤害。讽刺的是，在撤回盐酸双环胺后，人们对于当时市场上唯一能被用于治疗孕妇晨吐的药物的潜在致畸作用却是一无所知。⁵

通常科学研究最多能做的是减少不确定性。治疗可能会有利有弊。一项精心开展的高质量研究会将某一健康问题的一种治疗和其他治疗或不治疗进行比较，从而反映出该治疗有益或有害的概率（可能性）。鉴于不确定性一直都存在，那么尝试着不去把事物看得那样绝对将会对我们有所帮助。而且从概率的角度思考会让这一切变得易于理解。⁶人们只需要了解在一种情况下出现一种特定结局的可能性，以高血压患者发生卒中为例，则是影响卒中发生几率的因素及某种治疗改变卒中发生几率的概率。在掌握了足够可靠的信息后，患者和医疗卫生专家就可以共同努力，对治疗的利弊进行评价。他们能够根据患者个人情况及意愿采取可能最恰当的治疗方案。⁷

本书旨在促进医患交流并提高信心，而非破坏患者对医生的信任。而这一切的发生有赖于患者能够参与，帮助自己和医生一起严格评价治疗方案。

第 1 章，我们简要介绍了有关治疗的科学试验的必要性，以及一些新的疗法是如何产生预料之外的有害影响的。第 2 章，我们描述了其他疗法的预期效果为什么未能实现，并强调许多常用的治疗并未经过充分评估这一事实。第 3 章则说明了为何治疗越多未必越好，而第 4 章解释了为什么筛查健康人群对于早期发现疾病有利也有弊。第 5 章，我们强调了遍布于医疗卫生方方面面的诸多不确定性中的一部分以及如何处理它们。

第 6、7、8 三章以通俗易懂的方式对一些“技术细节”给予了说明。在第 6 章，我们概述了有关治疗的科学试验的基础，并强调确保同类比较的重要性。第 7 章则强调了为何将偶然因素考虑在内是必须的。第 8 章则阐述为何系统地评价所有相关的可靠证据如此重要。

第 9 章概述了为何疗效研究的监管系统（通过研究伦理委员会或其他类似的实体机构）会对高质量研究的实施产生障碍，并解释了为什么监管系统可能会无法保证患者的利益。第 10 章则对高质量、低质量和不必要的三类疗效性研究的主要区别进行对比，并指出商业因素和学术权威如何歪曲研究，最终导致那些真正关乎患者利益的问题得不到解决。

第 11 章指出患者及公众应该如何做以确保更好的验证疗效。我们可以在第 12 章看到来自疗效研究的有力证据确实能为患者个体带来更好的卫生保健。最后在第 13 章，我们不仅向大家展现了未来更美好的蓝图，也向大家提供了切实的行动计划。

本书每一章节的参考文献都精选于重要的参考资料，本书末尾（见 184 页）还有单独的“附加资源”模块。如果您还想更深入地探索某个问题，那么詹姆斯·林德图书馆（www.jameslindlibrary.org）是个不错的起点。读者也可通过一个新网站“验证疗效交互平台”（www.testingtreatments.org）免费获得第二版《验证疗效》一书的电子版，在接下来几年里还会陆续补充其他语种的译本和资料。

本书的所有作者都坚信以下原则：基于人们的需要且行之有效的卫生保健应做到人人享有。而这种社会责任依赖于可靠且可获得的信息，这些信息是高质量研究验证出的疗效。考虑到目前卫生保健资源有限，如果要让所有人都享受到医学进步带来的巨大成果，那么治疗方法就必须基于高质量证据，并且被公平、有效地应用。将宝贵的资源浪费在疗效甚微的手段上，或毫无根据地摒弃本可评价未知疗法的机会，都是极端不负责任的做法。因此，有关疗效的科学试验能够从根本上确保我们拥有平等的治疗选择权。

我们希望广大的读者不要被作者在《验证疗效》一书中，对这一话题所分享的一些观点所束缚，而是应该继续对治疗方法提出千奇百怪的问题，找出医学知识存在的漏洞，并参与到医学研究中来寻找答案，既能利于自身，又可惠于众生。

第一章 新的不一定更好

为什么科学验证治疗方案是必需的

未经科学的（无偏倚的）评价，无效甚至有害的疗法可能会被认为有效，而被医生采用，相反，真正有效的疗法则可能会被认为无效而不被采用。因此，所有的治疗方法都应进行科学试验，不论该项治疗是由谁提出，或是当前公认的常规疗法抑或补充（替代）疗法。那些未被验证的关于疗效的理论，尽管听起来很有说服力，却远不足以令人完全信服。例如，有些理论称某些治疗会有效，而科学试验却显示其无效；有些理论似乎很有把握地预测某些治疗无效，但试验结果却显示其确实有效。

尽管人们想当然的认为“新的”就意味着是“更好的”，就像那些洗衣机专用洗涤剂广告一样，当用科学试验予以验证时：新疗法相比于已有的治疗方法，效果更差与更好的可能性差不多。同样，人们想当然认为那些已经用了很久的疗法肯定是既安全又有效的。这种情况下，我们必须意识到，卫生保健行业中基于习惯或固有观念而非科学证据的疗法比比皆是，这些疗法常常是无效的，有时甚至会造成重大伤害。

对科学试验的需求并非新事物：早在 18 世纪，詹姆斯·林德（James Lind）医生曾用一项科学试验比较了当时用于治疗坏血病（一种在长途航海中夺去大量海员生命的疾病）的六种方法。结果表明，橘子和柠檬（现在我们知道是因为含有维生素 C）是非常有效的疗法。

轶事仅仅只是轶事

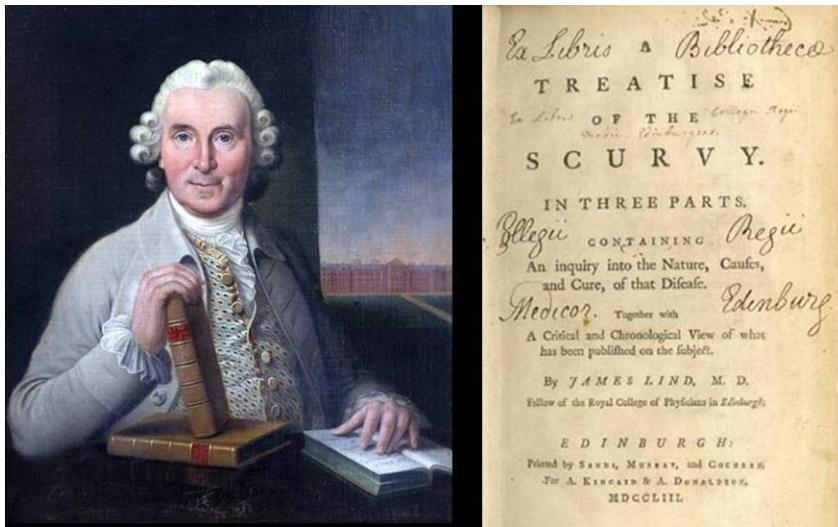
“对于各种奇闻轶事，我们的大脑就像一个存储硬盘，我们很容易通过那些引人入胜的故事获取信息；但我着实惊讶于那么多人，包括我相当多的朋友，没有发现其中的“陷阱”。科学告诉我们，轶事和个人经验很可能具有严重的误导性。而科学要求的结果是能够被验证的同时具有可重复性。另一方面，医学只能是一门科学。但当面对的是个体的患者时，由于人类存在太多的个体差异，我们很难对任何事情下定论，因而也留有很大猜测的空间。但我们要明确这一界限，如果一意孤行，我们很快就会违背科学的本质：投机取巧，混淆事实与推测，直到我们难以分辨二者。”

Nick Ross. Foreword. In: Ernst E, ed. *Healing, hype, or harm? A critical analysis of complementary or alternative medicine*. Exeter: Societas, 2008:vi-vii

1747 年，作为时任英国皇家海军舰艇萨利斯伯里号（HMS Salisbury）随船外科医生的詹姆斯·林德，召集了 12 名处于该病相似阶段的患者，将他们安置在船体的同一位置，并确保他们的基本饮食摄入相同。以上三点至关重要，因为其建立了“具有可比性的环境”（见第 6 章和第 3 章第 26 页方框中的内容）。林德将每两名海员分为一组，随后让每组分别接受当时六种坏血病疗法中的一种——苹果酒、稀硫酸、醋、海水、肉豆蔻，或两个橘子和一个柠檬。最终，水果组轻松“摘得桂冠”。此后，英国海军部下令所有船只都供应柠檬汁，到 18 世纪末，这种致命的疾病便从英国皇家海军中消失。

在林德所比较的不同疗法中，英国皇家医学院推荐的是稀硫酸，而海军部则倾向于醋，

但林德医生的科学试验证明这两个权威机构的推荐都是错误的。而令人惊讶的是，医学权威出错的情况并不少见。在卫生保健领域，过分信赖专家观点、个人习惯或先例而非科学试验的结果仍在继续导致着严重的问题（见下文及第 2 章）。



詹姆斯·林德（1716-1794），苏格兰皇家海军外科医生，附图是他的论著和其中最著名的一本书的扉页，该书记录了他在 1747 年开展的一项对照试验，试验表明柑橘和柠檬用于治疗坏血病比当时所用的其他五种方法更有效（见 www.jameslindlibrary.org）。

如今，当医生和其他医务工作者就某种特定疾病情况的最佳疗法持不同意见时，疗效的不确定性常会备受关注（见第 5 章）。患者、公众和医生对于消除这种不确定性均起着重要的作用。为符合患者和医务人员的最大利益，对治疗方案的研究必须严谨。医学专家必须确保自己给出的治疗推荐意见是基于高质量的证据，患者也应该要求确保其所获得的治疗基于高质量的证据。只有形成这种重要的合作关系，公众才能有信心去接受现代医学所提供的一切（见第 11、12 和 13 章）。

预期外的有害作用

沙利度胺

沙利度胺（反应停，译者著）是关于一项新治疗弊大于利的骇人听闻的案例。¹ 20 世纪 50 年代后期，作为当时常规处方药巴比妥类药物的替代药，沙利度胺这种安眠药显得更加安全，且不同于巴比妥类药物，过量的沙利度胺不会导致昏迷。沙利度胺还特别被推荐给孕妇使用，因为它可以用于减轻晨吐。

随后，在 20 世纪 60 年代早期，产科医生开始注意到新生儿出现严重四肢畸形的情况大幅增加。这种医学史上罕见的疾病会导致患儿肢体极其短小，好像手掌和脚掌是直接长出来的一样。德国和澳大利亚的医生认为这种婴儿畸形与其母亲在妊娠早期服用了沙利度胺有关。²

1961 年底，厂家召回沙利度胺。许多年后，经过公众游行和法律诉讼，受害者们开始得到赔偿。人们为这种毁灭性畸形所付出的代价是惊人的：在沙利度胺作为处方药销售（有的国家甚至将其作为非处方药销售）的大约 46 个国家中，成千上万名婴儿惨遭不幸。沙利

度胺惨剧震惊了医生、药厂和患者，引发了世界范围内对药物研发和审批过程的重大变革。

3

婴儿流行性失明悲剧

“第二次世界大战刚结束时，引入了许多改善早产儿预后的新疗法。随后几年，人们不幸发现，许多治疗婴儿的新方法产生了完全意想不到的有害作用。这些在临床上发生的悲剧中，最轰动的当属 1942 至 1954 年间出现的“流行性”失明——晶状体后纤维增生症。后来发现该病和用于治疗新生儿发育不全的输氧方式有关。人们用了 12 年的时间平息这场悲剧，并从中清醒地认识到，所有的医疗变革在推广之前必须经过严格周密的评价。”

Sliverman WA. Human experimentation: a guided step into the unknown.

Oxford: Oxford University Press, 1985: Vii-Viii

万络

尽管药品试验监管已得到一定程度的重视，但即使经过最好的药物试验也无法保证该药绝对安全。非甾体抗炎药（NSAIDs）的例子就很好地说明了对药物提高警惕的必要性。非甾体抗炎药在多种情况（例如，关节炎）下都被用于缓解疼痛和炎症，同时还能降温退烧。“经典”的非甾体抗炎药包括许多像阿司匹林和布洛芬这样可以在药店柜台上买到的药物。而众所周知，它们常见的副作用为胃肠道刺激，进而引起消化不良，有时还会引起出血甚至胃溃疡。因此，药厂会理所当然地试图研制一种不会引起上述副作用的非甾体抗炎药。

罗非昔布，或罗非考昔（最为人所熟悉的是其商品名“万络”）于 1999 年作为一种号称比之前抗炎药更加安全的替代物问世，很快被推广使用。约五年后，市场上的万络被药厂自主召回，原因是它能增加心脏病和卒中等心血管疾病的风险。那么究竟发生了什么事呢？

万络在 1999 年被美国食品和药物管理局（Food and Drug Administration, FDA）批准为“缓解骨关节炎症状、成人急性疼痛以及痛经（月经期的疼痛）”的药物。之后其被证实能缓解成人和儿童风湿性关节炎的症状。在该药研发阶段，药厂的科研人员就已经了解到它对机体凝血机制有潜在的有害作用——该药能够增加血栓形成的风险。然而为求批准，药厂向 FDA 提交的仅仅是所有研究中那些关注万络抗炎作用而未涉及其潜在副作用的一小部分研究。

在 FDA 批准该药之前，药厂就已经以风湿性关节炎患者为对象，开展了一项主要比较万络和萘普生（另外一种非甾体抗炎药）对肠道副作用的大规模试验。这次试验再次故意没有对该药心血管方面的副作用进行研究。因此，随后有人对研究数据和安全监察委员会成员（该委员会负责监察研究的累计结果，以审查是否需要停止该项研究）的利益冲突提出了质疑。

然而，在证明万络较萘普生会减少胃溃疡和胃肠出血的试验中，确实发现万络组中有更多的人心脏病发作。结果这份发表在重要医学期刊上的研究报告遭到了严厉的批评。认为该研究的缺陷在于通过以淡化其发生心血管疾病风险严重性的方式来分析和展示结果。后来，该杂志的编辑抱怨称研究者没有提供有关副作用的关键数据。这一结果于 2000 年提交 FDA，2001 年经过关节炎顾问委员会讨论，最终 FDA 在 2002 年修正了万络的安全信息：会增加

心脏病和卒中的风险。

药厂随后继续发掘万络的其他用途，并于 2000 年开展了一项证实其能否预防结肠（下消化道）息肉（可能发展为结肠癌的良性肿瘤）的研究。这项研究草草结束，原因是它的阶段性结果显示万络与心血管并发症风险增高有关，从而使药厂在 2004 年自主召回了市场上的万络。在既没有受药厂雇用也没有收取公司咨询费的前提下，有作者在报告中称心血管并发症只会在使用万络 18 个月后才出现。该声明基于一项有缺陷的分析，而随后被发表该报告的期刊正式纠正。⁴ 面对众多患者事后的法律诉讼，药厂坚称从该药的批准前研究到万络上市后的安全监察，他们一直很负责任，同时重申：“证据”会显示导致不良后果的原因是由于患者已有的心血管疾病危险因素，而非万络本身。⁵

万络事件告诉我们，在沙利度胺事件发生半个世纪后的今天，为确保疗法试验做到公平、过程透明、证据有力，我们需要做的还有很多。正如某个评价小组指出：“我们的系统将患者的利益放在第一位。专家、临床医生、药厂和期刊间的合作对于拓展专业知识及提高对患者的服务水平是必不可少的。信任是这一合作关系的必要因素，但最近的事件使建立一套专门维护患者利益的体制显得很有必要。让这些体系中所有相关人员及机构给出新的承诺是唯一能让整个不幸事件带来些有用借鉴的方法。”⁴

文迪雅

2010 年另一种新药——罗格列酮，人们更熟悉的是其商品名“文迪雅”，成为了新闻头条，原因是其与心血管系统相关的预期外的副作用。10 年前文迪雅作为一种治疗 2 型糖尿病的新方法由欧美的药品管理机构批准销售。2 型糖尿病是由机体胰岛素分泌不足或机体细胞对胰岛素不敏感所引起，相对于 1 型糖尿病（机体完全不产生胰岛素）更常见。2 型糖尿病常和肥胖有关，通常在合理安排饮食、运动和口服降糖药几个月之后可获得理想的疗效，无需注射胰岛素。2 型糖尿病的长期并发症为增加心脏病和卒中的发病风险，所以治疗的主要目的是减少这些并发症的风险。文迪雅因其能通过一种新方式使自身产生的胰岛素作用更充分而被推广，并声称它控制血糖水平的效果优于已有药物。可见，人们关注的是该药的血糖控制效果而非其导致痛苦甚至死亡的严重并发症。

当文迪雅被许可销售时，证明它有效的证据很少，同时没有证据说明它对心脏病和卒中风险的影响。药品监管机构曾要求药厂做进一步研究，然而当时该药已经在世界上被大面积作为处方药推广了。随后，有关文迪雅对心血管系统不良影响的报道开始出现并逐渐增多；直到 2004 年世界卫生组织（WHO）着手干预此事，要求药厂再次研究有关并发症的证据。药厂开展了相关研究，并证实了该药会增加心脏病和卒中的风险。⁶

此后又过了六年，药品监管机构才真正认真调查了相关证据并付诸行动。2010 年 9 月，美国 FDA 宣布将文迪雅的使用严格限制于其他药物治疗无效的 2 型糖尿病患者；同月，欧洲药品管理局宣布文迪雅将在随后的两个月内被召回。两大药品监管机构都将该药增加心脏病和卒中风险作为他们这一举措的原因。期间，一些有独立见解的调查者发现当时错过了可以采取进一步行动的计划，应该如一个医学专家小组所指出的，药品监管机构和医生有必要“在大范围地给那些向我们寻求建议和治疗的群体用药之前，掌握更可靠的证据。”⁷

人工心脏瓣膜

不仅药物治疗可产生预期外的有害作用，非药物治疗也可以带来严重的危害。人工心脏瓣膜是目前严重心脏瓣膜病患者的标准疗法，并且近年来人工心脏瓣膜的设计有很大的改善。然而，一个特殊类型人工心脏瓣膜的事例说明了试图改进设计的尝试如何造成灾难性的后果。20 世纪 70 年代初，一个名叫 Björk-Shiley 的心脏瓣膜装置面世，但是早期的样式容易形成血栓（血凝块形成），因而限制了它的使用。为克服此缺点，在 20 世纪 70 年代末该装置被进一步完善以减少形成血凝块的可能性。

此新型装置主要为一个由两个金属瓣轴（起支持作用）固定的圆盘，当时全世界有数千个该装置已被应用。但不幸的是，其后发现该瓣膜的结构存在严重缺陷：其中的一个瓣轴有断裂的倾向——该缺陷称为瓣轴断裂——这将导致严重甚至致命的瓣膜功能障碍。

该事件发生后，瓣轴断裂被认为是该装置上市前检测的问题，是由焊接不全导致，但是其原因未经充分调查。美国 FDA 接受了如上解释，同时也接受了制造商关于降低瓣膜血栓风险的保证而没有勒令他们对瓣轴断裂造成的风险进行任何补偿。FDA 在 1986 年强制该瓣膜装置下市，是等到有非常明显的严重瓣膜断裂的证据出现之后，而不是在造成数百名患者不必要死亡之前。虽然现在药品管理系统得到了改进，包括更好的上市后患者监测和全面的患者登记，但仍迫切地需要在新装置面世时提供更大的透明度。⁸

好的难以置信

赫赛汀

不仅商业公司会避开其缺点而鼓吹新疗法的优势，专业的炒作和媒体的积极宣传都会进一步夸大药物的益处而掩盖其潜在风险。而且，就乳腺癌治疗药物群司珠单抗——大家所熟知的商品名为赫赛汀（见第 3 章）事件来看，这些风险可能不仅仅包括有害的副作用，还包括诊断难点。

2006 年上半年，患者、专业人员的强烈要求加上制药企业和广大媒体的推动，促使英国国家卫生署（National Health Service, NHS）为早期乳腺癌患者批准了赫赛汀。“患者心声的力量”取得了成功——赫赛汀成为了一种特效药（见第 11 章）。

但当时赫赛汀仅被允许用于治疗转移性（已扩散的）乳腺癌，而对早期乳腺癌的作用并未进行充分验证。事实上，该药的生产商只是申请将其用于治疗处于癌症早期的一小部分特定类型的女性——人类表皮生长因子受体-2（HER2）蛋白检测阳性的女性，只有五分之一的具有此基因检查记录。精确评估患者 HER2 是否阳性的难度和费用、被误诊（即出现“假阳性”）从而被误治的可能性却很少被热情过度但缺乏批判头脑的媒体所报道。同样，至少五分之四乳腺癌患者的 HER2 为阴性这个事实也未被强调^{9,10,11,12}。

直到第二年，英国国家卫生与临床优化研究所（National Institute for Health and Clinical Excellence, NICE）——一个负责科学验证疗效并提出推荐意见的机构，才推荐赫赛汀作为 HER2 阳性的早期乳腺癌患者的治疗选择。但同时也提出一个严重警告：因为越来越多的证据证明赫赛汀可能对心脏有副作用，NICE 推荐医生在开药之前先评估病人的心功能，不能给有从心绞痛到心律失常等各种心脏疾病的患者服用该药。考虑到有关其副作用（有些还很严重）的短期研究数据，NICE 强调用药时应谨慎。而长期结果，不论有害还是有利，都需要经过一定的时间才能出现。¹³

其他国家也存在强烈要求使用赫赛汀的情况。例如在新西兰，患者支持小组、新闻媒体、药厂和政府官员都要求早期乳腺癌患者应该使用赫赛汀。新西兰药品管理局（Pharmaceutical Management Agency, PHARMAC）——和英国 NICE 的职能类似，也评价了赫赛汀治疗早期乳腺癌患者的证据。2007 年 6 月，基于上述评价结果，PHARMAC 确定早期乳腺癌患者适合服用 9 周的赫赛汀，并且在此期间同时服用其他的抗癌药物而非先后服用。这种 9 周疗程是当时全球都在使用的三种用药疗程之一，并且 PHARMAC 决定设立基金，用于一项关于赫赛汀治疗最佳疗程的国际研究。然而在 2008 年 11 月，新选举出来的政府官员忽视了 PHARMAC 的循证决策，并宣布为该药物的 12 月疗程设立基金。¹⁴

卷入赫赛汀热潮

2006 年，有一位恰好接受过医学培训的英国患者，发现自己也被卷入了赫赛汀热潮。她是在 2005 年被诊断为 HER2 阳性的乳腺癌患者。

“在确诊之前，我对目前乳腺癌的治疗了解很少，和其他患者一样，就在网上查一些资料。乳腺癌治疗网站当时正开展一项活动，旨在让所有 HER2 阳性的乳腺癌患者都可以获得赫赛汀，我也注册了。当时看到网站和媒体的数据，我只是不明白为何如此有效的药物竟没有用于那些癌症复发的女性，而她们肯定是会接受的啊！我开始觉得如果我不接受此药，我将几乎不可能活下来！同时，我受到了太阳报的访问，该报支持赫赛汀运动并对于作为医生但同样是‘癌症受害者’的我的故事很感兴趣。

在完成化疗后，我和我的肿瘤医生讨论了赫赛汀治疗。他对研究新报道的赫赛汀的长期心脏副作用表示担忧，特别是想到该药将会被用于其他方面都健康的女性，但是这并没有受到网站和媒体的关注。再者，仔细分析之前广为流传且已深入脑海的‘50%有益’的数据，如今对我而言已经变为 4~5%有益了，也就是说和发生心脏风险的比例是同等的！因此我决定不再接受该药物的治疗，即使我的肿瘤复发了我也会因作出这个决定而庆幸。

这个故事阐述了一个接受过医学培训并且通常比较理性的女性，当被诊断为患有可能致命的疾病时受到的伤害。关于赫赛汀治疗早期乳腺癌的大量信息，源于媒体和相关行业不恰当的大肆宣传，并因像我这样的患者而愈演愈烈。”

Cooper J. Herceptin (rapid response). BMJ . Posted 29 November 2006 at www.bmj.com.

关于赫赛汀的使用还存在很多的不确定性——例如，服用该药的时间、应服用多久、该药对某些患者的长期作用是否弊大于利、该药是否可以延迟或者防止癌症的复发。除此之外的担忧也接踵而来，该药和诸如蒽环类药物、环磷酰胺的抗乳腺癌药物一起服用，可能会使患者产生心脏病副作用的概率从 4%增加至 27%。¹⁵

要点

- 验证新的治疗方案是必要的，因为其比现有治疗更好或更差的可能性是一样的
- 对治疗方案有偏倚（不科学）的试验可导致患者遭受痛苦甚至死亡
- 事实上，被批准的治疗并不能保证就是安全的
- 治疗的副作用通常需要经过一段时间才会出现
- 治疗的好处常常会被夸大，而坏处则会被淡化

第 2 章 预期的效果没有变为现实

一些治疗在被意识到弊大于利之前就已经使用很长一段时间了。而预期效果可能不会变为现实，在这一章中我们会解释为什么会这样。

对婴儿睡姿的建议

不要以为只有药物会产生伤害，医学专家的建议同样可以致命。许多人都听说过美国儿童保健专家本杰明·斯波克医生（Benjamin Spock），他的畅销书《婴儿与儿童保健》（Baby and Child Care）几十年来一直被父母和专家们奉为育儿圣经，尤其是在美国和英国。但斯波克医生却因他给出的一条善意的建议而铸成大错。这条建议从逻辑上看无可辩驳，且明显带有权威性。从 1956 年著作的出版直到 20 世纪 70 年代，他一直提倡：“婴儿躺着睡有两大坏处：一是如果婴儿呕吐，躺着睡使他更可能被呕吐物呛噎；二是他倾向于将头一直偏向同一边……这可能使一侧头部扁平……我认为最好一开始就让婴儿习惯趴着睡。”

让婴儿趴着睡不仅成了医院的标准规范，也被数百万父母在家中尽职尽责地实行着。但我们现在才发现，这种从未经过严格评价的做法导致了数以万计的婴儿猝死，而这些悲剧本可避免。¹虽然并非所有的婴儿猝死事件都归咎于这条令人遗憾的建议，但当摒弃了这一做法并提出让婴儿躺着睡的建议后，婴儿猝死事件的发生明显减少。当 20 世纪 80 年代关于趴着睡的有害证据清晰地出现时，医生和媒体开始警告该睡姿的危险，之后婴儿猝死事件的数量开始大幅度下降。随后这一消息演化成了一项名为“躺着睡（back to sleep）”的运动，尽可能消除斯波克医生那条令人遗憾的建议的不良影响。



用于心脏病发作后患者抗心律失常的药物

尽管斯波克医生的建议看上去符合逻辑，但其基于未经验证的理论。这样做引起危险的其他例子不难找到。一些患者在心脏病发作后会出现心律失常——也就是心律失常。这些患者比起未出现心律失常的患者更容易死亡。由于有药物能够抑制心律失常，便假设这些药物也能降低心脏病发作后死亡的风险，这似乎符合逻辑。实际上这类药物恰好有相反的效果。但在临床试验中，这类药物仅被着眼于验证它们是否能减少心律失常。1983 年研究人员系统评价了试验所累积的证据，发现没有证据表明这些药物能降低病死率。²

然而，这类药物在接近 10 年的时间里继续被使用，继续导致患者死亡。在 20 世纪 80 年代后期，在抗心律失常药物使用的高峰期，据估计仅在美国该药每年导致数万人早死。每年由于该药致死的美国人数比越南战争期间阵亡的还要多。³ 随后人们发现，某些说明该药物可致死的试验结果出于商业目的从未被报告（见第 8 章，第 97 页）。⁴

己烯雌酚

曾经有段时间，医生对有流产和死产史的孕妇服用一种名为己烯雌酚（DES）的合成（非天然）雌激素是否有益不确定。一些医生将其作为处方药，另一些医生则没有。20 世纪 50 年初期，DES 开始盛行，被认为可以改善胎盘功能障碍，而这种障碍有可能是引起流产或死产的原因。报告称有流产和死产史的妇女经 DES 治疗后产下活婴，这使得那些使用过 DES 的医生备受鼓舞。

例如，一名英国产科医生给一位有过两次死产经历的女性患者从妊娠早期开始便开了 DES 这种药物。这位孕妇最后产下一名活婴。这位产科医生推断该妇女成功分娩的“自然”能力在这段时间可能已经得到改善，因此在她第四次妊娠期间停用了 DES；但婴儿因“胎盘功能不全”在子宫内死亡。所以，在她第五次和第六次妊娠期间，这位产科医生和这位妇女都毫不犹豫地认为应再次使用 DES，结果这两次妊娠都产下了活婴。这位产科医生和妇女由此得出 DES 是有用的药物。不幸的是，根据这一事件得出的结论从未经过严格的科学试验验证是正确的。在该妇女接受这种治疗的同一段时间，实际上无偏倚的研究正在进行，其报告没有发现 DES 有益的证据。⁵

即使科学试验没有得出 DES 有助于预防死产的证据，但 DES 的故事并没有结束。20 年后，有害的副作用证据开始出现：有位母亲对患了罕见阴道癌的女儿的患病原因进行了仔细观察，发现女儿的患病很可能跟她自己在妊娠期间曾服用 DES 有关。⁶ 这是次正确的观察，而最重要的是它后来被验证是正确的。此后，大量研究显示无论是男是女，凡出生前接触过 DES，便会有各种各样的严重副作用。这些副作用不仅包括罕见癌症的发生率增加，还包括生殖系统的其他异常。

当被正式宣布 DES 不能用于妊娠期的时候，已有数百万人用过此药。若当时我们了解现在所知事实的真相，如果在 20 世纪 50 年代医生们就利用可得的关于 DES 最可靠的研究证据，那么很少有医生会再给患者开这种药了，因为从一开始被作为处方药使用时，DES 就从来没有被证明是有效的。可悲的是，疗效证据的缺乏被人们普遍忽视了。⁷

激素替代疗法（HRT）

对于更年期妇女常见的烦人的热潮红，激素替代疗法（hormone replacement therapy, HRT）疗效显著，且有证据表明这种疗法可能有助于预防骨质疏松。HRT 的益处逐渐被越来越多地挖掘，包括预防心脏病及卒中。数百万妇女在医生的建议下开始长期使用 HRT。然而，这些主张的根据却非常不可靠。

难怪她会感到困惑

2004 年 1 月，一名切除了子宫的患者在给《柳叶刀》（The Lancet）的一封信中写到：

“我于 1986 年因子宫肌瘤切除了子宫。医生还发现我患有子宫内膜异位症，也切除了我的卵巢。因为当时我才 45 岁并即将绝经，于是我接受了激素替代疗法（HRT）。第一年，我使用结合雌激素（倍美力）治疗，但从 1988 年至 2001 年间，给我做过手术的外科医生私下为我又做了手术，每 6 个月进行一次雌激素皮下植入。我对这种疗法总是心存疑虑，因为一旦雌激素皮下植入完成，我感觉自己无法对其进行控制，而且术后多年我经常头痛。除此之外，我觉得自己很健康。

然而，我的外科医生向我保证，HRT 有许多优点并且适合我，对此我也表示认同。随着时间的推移，HRT 已不像前些年那样仅被作为美容药物使用，它有越来越多的益处被报告出来。它不仅对心脏和骨质疏松有好处，对于卒中也有部分预防作用。每次我去看医生，他似乎有更多的关于使用 HRT 有益的证据。

我的外科医生在 2001 年退休了，我去找 NHS 的医生。他的话让我大吃一惊！他说的与我私人外科医生刚好相反——停用 HRT 会是个不错的选择：它可以增加心脏病、卒中和乳腺癌的风险，也是导致头痛的原因。我再次做了一次雌激素皮下植入，然后用了很短时间的倍美力，从那时起，我已有 8 个月没有使用 HRT。医生告诉我是否继续使用 HRT 由我来拿主意，为此我感到困惑……

我不明白为什么在如此短的时间内 HRT 及其众多神奇的优点被全盘否定了。像我这样的外行该如何选择？尽管到目前为止我还未受到 HRT 的不良影响，但我还是花了很多时间来论证和思考我是否应该继续使用 HRT。我对整件事感到非常困惑，我肯定其他女性患者也有相同的感受。”

Huntingford CA. Confusion over benefits of hormone replacement therapy.

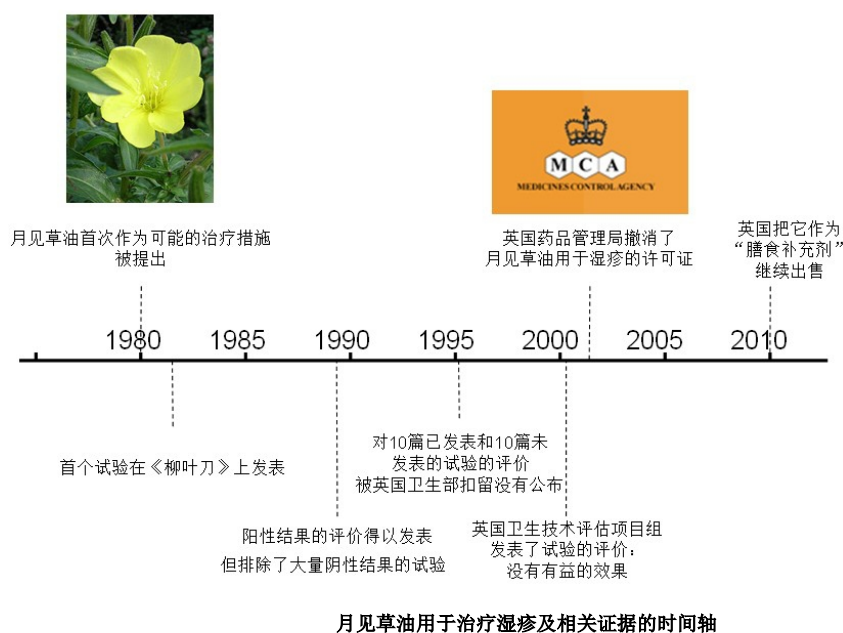
Lancet, 2004; 363: 332

单就心脏病而言，妇女们在过去 20 年间一直被告知 HRT 能降低这种严重疾病的风险，实际上这一建议是基于有偏倚（不科学）的研究得出的（见第 1 章和第 6 章）。1997 年，芬兰和英国研究者系统地评价了高质量研究的结果后发出警告：这一建议可能是错误的。⁸ 他们发现，HRT 不仅不能减少心脏病的发生，反而可能增加发生的风险。一些著名的评论家否认上述结论，但两项大样本的高质量试验验证了系统评价的结果。如果 HRT 的效果在其问世时就被正确评估，那么妇女们就不会被误导，当中许多人就不会早死。更糟的是，我们现在发现 HRT 可增加卒中和罹患乳腺癌的风险。⁹

总的来说，HRT 仍是妇女更年期症状的有效治疗方法。¹⁰ 但可悲的是过去曾明确将其作为降低心脏病和卒中的方法来大力提倡。虽然 HRT 增加这类严重疾病的风险是不大的，但由于曾广泛应用，故受此影响的妇女实际上非常多。

月见草油治疗湿疹

未经充分评价的治疗即便不会导致死亡或危害患者，但会浪费金钱。湿疹是一种令人痛苦的皮肤疾病，儿童和成人皆可感染。皮损处会又丑又痒。尽管使用皮质激素乳膏对此有效，但令人担心的是这些治疗的副作用，如皮肤变薄。20世纪80年代初期，一种天然植物提取物月见草油作为可能的一种副作用较少的替代药物问世。¹¹月见草油富含一种名为 γ -亚麻酸（GLA）的必需脂肪酸，故貌似其可治疗湿疹。例如有人认为，GLA在湿疹患者体内的转化（新陈代谢）途径受损，故理论上补充GLA对患者有益。琉璃苣油，又名七瓣莲油，GLA的含量更高，也被推荐用于治疗湿疹。



GLA是安全的，但它是否有效？为此开展了大量研究，但得出的结果相互矛盾。供应商赞助的研究导致已发表的证据受到严重影响。1995年，英国卫生部要求与月见草油制造商无关的研究者评价20个已发表和未发表的研究，没有发现有益的证据。因遭到药厂反对，卫生部从未将这份报告公之于众。但五年后，同一批研究者发表了另一篇关于月见草油和琉璃苣油的系统评价，结果显示，即使是大样本、高质量的研究也没有得出令人信服的证据表明这些治疗有效。¹²

还有一种说法未被验证——可能只有大剂量GLA才起作用。2003年，一项高质量的科学试验把这种假设推翻了。¹³可笑的是，直到这些研究结果发表以后，英国药物管理局（MCA，后来更名为药品和保健品管理局MHRA）在2002年10月才最终因无证据表明药物有效而撤销了两个主要的月见草油制剂的生产许可证。

然而，由于没有发现有关月见草油安全性的问题，它作为多种疾病的“膳食补充剂”仍在很多地方被出售。就其对湿疹的治疗作用而言，说明书上采用了模糊的描述，如“湿疹患者可能得到缓解”，“可能有用”，“其某些药物特性对湿疹等疾病可能具有抗炎作用”。

要点

- 医学理论或专家意见对安全有效的治疗都不是可靠的指导
- “已被使用”的治疗并不意味着利大于弊

- 即使患者没有遭受未经充分验证的治疗带来的伤害,使用它们仍会对个人和社会的资源造成浪费

第3章 越多未必越好

许多人误认为如果一种治疗确实有效，那么肯定是越多越好。这种观点是完全错误的，实际上，越多的治疗可能效果越差。治疗的最适剂量应该兼备最佳的效果和最少的不良反应，而确定最佳剂量是所有的治疗都普遍存在的挑战。如果增大剂量，患者的获益将进入一个平台期（即疗效不再增加），但通常不良反应会加重。所以说，“越多”可能在实际上降低了治疗效果，或者甚至导致伤害。

利尿剂的使用就是如此：在低剂量水平，它可以有效降低血压，同时不良反应较少，而在高剂量水平，它不仅不能继续降压，反而会导致一些预期之外的副作用，如排尿过多、阳痿、血糖升高等。与之相似的是，阿司匹林在低剂量水平，即每天标准剂量的 1/4 到 1/2 之间，可以有效预防卒中且很少有不良反应，然而服用几倍于标准剂量的阿司匹林，不但不会有预防卒中的作用，还会导致头痛和增加胃溃疡的风险。

“最佳剂量”的概念与原理不仅适用于药物治疗，而且可扩展至其他类型的治疗，包括手术治疗。

乳腺癌的过度治疗

新闻报道中常常鼓吹的乳腺癌治疗为我们提供了一些颇有价值的反面教材，即认为越多治疗必然有利。

我们这样做，是因为……

“我们（医生）这样做，是因为其他医生是这样做的，而我们不想和其他人不一样，所以我们也照样做；或是因为老师、同事和住院医师就是这样教我们的；或是因为老师、管理者、监督者、指南制定者要求我们这样做，并且认为我们必须这样做；或是因为患者希望我们这样做而我自己也认为应该这样做；或是因为更多的其他因素，如不必要的检查（特别是医生主导的）和不必要的就医；或是因为对法律体系及对临床实践评估的担心，我们认为应该这样做（所谓的“掩盖自己”）；或是因为我们需要时间（让一切顺其自然），所以我们这样做；最后但最常见的，我们必须做点什么（为自己辩解），和我们无法套用常识，所以我们这样做。”

Parmar MS. We do things because (rapid response).

BMJ. Posted 1 March 2004 at www.bmj.com.

整个 20 世纪甚至直到 21 世纪，乳腺癌患者一直被要求使用并忍受一些极其残忍和痛苦的治疗。一些治疗，无论手术还是药物，带给患者的伤害远远超过治疗疾病本身所需要承受的。然而这些却被患者和医生毫无疑问地接受了，患者深信越是激进的或毒性越大的治疗，越有可能战胜疾病。勇敢的医生和率直的患者代言人花费了数年时间才动摇了这种错误的观点。他们不仅要寻找可靠的证据来推翻“越多越好”的无稽之谈，同时还要遭受同行的嘲讽和医学权威的阻挠。

直到今天，即使有明确证据表明某种治疗没有任何益处，或者会导致明显伤害甚至死亡的可能性，但是人们对疾病的恐惧和“越多一定越好”的想法依旧左右着治疗决策。例如，

这种心态仍然促使一些患者和他们的医生选择传统的破坏性手术治疗。还有些人宁愿选择有明确痛苦的副作用的高剂量化疗，或者是可能导致严重心脏疾病的赫塞汀（见第 1 章），而不去考虑那些疗效相当的简单的疗法。怎么会这样呢？

激进的治疗措施并非总是最佳

“我们很容易认为，癌症治疗后疗效好是因为采用了激进的治疗措施。为了保护患者免受不必要的风险和侵袭性治疗早期或晚期会出现的副作用，采用随机对照试验对激进治疗和非激进治疗进行比较非常重要。这种比较是符合伦理的，因为患者拒绝可能的获益同时意味着免受可能不必要的危害——没有人知道最终会是什么结果。”

Brewin T in Rees G, ed. *The friendly professional: selected writings of Thurstan Brewin*. Bognor Regis: Eurocommunica, 1996.

破坏性手术

直到 20 世纪中期，手术仍然是乳腺癌的主要疗法，这是因为人们坚信乳腺癌进展缓慢且有规律，如先从乳腺肿瘤部位侵袭至周边腋窝淋巴结。由此推论，手术范围越大，手术时间越早，越有可能阻断肿瘤扩散。治疗基本上就是广泛的“局部”手术，即乳腺或乳腺周围组织的切除。虽然称为局部手术，但根治性乳房切除术要切除乳房本身、大面积的胸肌并广泛清扫大量的腋窝淋巴结组织。

传统的根治性乳房切除术

在 20 世纪中后期之前，William Halsted 于 19 世纪后期提出的根治性乳房切除术一直在乳腺癌手术中颇为盛行。施术者不但要切除整个乳房，还要切除覆盖胸壁的胸大肌。为了便于手术医师进入腋窝清除淋巴结和周围脂肪组织，甚至连胸小肌也会被切除。

范围更大的根治性乳房切除术

在当时，“越多越好”的观点促使实施根治性手术的医生进一步扩大手术范围，锁骨下淋巴结和胸骨下乳内淋巴结都被切除。为了触及乳内淋巴结，还会取掉几根肋骨，并用凿刀切开胸骨。一些外科医生觉得这样还不够，甚至还要切除患侧手臂和体内各个部位的腺体（肾上腺、垂体、卵巢），以此来抑制激素的生成，因为激素被视为肿瘤扩散的“动力”。

如果某位患者历经这样的手术而幸存下来，她剩下的将是一副任何外衣都难以遮掩的受到了极度摧残的胸廓。如果手术部位在左侧，那么只会剩下一层薄薄的皮肤盖在心脏上。

修改于 Lerner BH, *The breast cancer wars: hope, fear and the pursuit of a cure in twentieth-century America*. New York; Oxford University Press, 2003.

Adapted from Lerner BH, *The breast cancer wars: hope, fear and the pursuit of a cure in twentieth-century America*. New York; Oxford University Press, 2003.

然而，一些善于思考和观察的乳腺癌专家注意到，这些日益剧增的破坏性手术似乎并未对乳腺癌的病死率产生任何影响。于是，他们提出了一种新的理论：乳腺癌实际上一开始就是一种系统性（即广泛扩散）疾病，而不是通过周围淋巴结有规律地从乳腺开始扩散。也就是说，他们认为在发现乳房包块时（见下文）癌细胞就已经存在于身体其他部位了。如果真是这样，他们建议切除肿瘤及周围适当范围内的正常组织，再加上一个疗程的局部放疗，会更有利于患者，并且可能与根治术同样有效。此时，“系统疗法”也被引入，该疗法用于对

抗身体其他部位癌细胞的生长或繁殖，也是基于这个乳腺癌扩散的新理论。

这一新思路导致的直接结果是，医生开始主张采用范围更小的手术治疗，如乳房肿瘤切除术，即切除肿块及周围正常组织边缘，然后术后行放疗，部分患者采用化疗。但在比较新方法与旧的根治性手术时，乳房肿瘤切除术的推动者们遇到了巨大的阻力。一些医生坚信非旧即新的疗法，而患者也会随之强烈地要求非旧即新的治疗。其结果是大大延误了收集关于新旧疗法优缺点的重要证据。

尽管面对这些困难，过度手术最终还是受到了来自不愿再继续施行疗效不确定的根治性手术的外科医生，以及不愿再忍受破坏性手术的直言不讳的女性患者的挑战。

20世纪50年代中期，美国外科医生 George Crile 开始公开指出他对“越多越好”观点的质疑。Crile 医生认为没有其他办法可以激起医生的批判性思考，因此他在《生活》(Life)¹ 杂志中撰文对此进行呼吁，并一针见血地指出：医学界内的争论现在应公之于众，而非仅局限于学术圈。另一名美国外科医生 Bernard Fisher 与其他专业的同事共同设计了一系列严谨的研究癌症生物学的实验，结果表明即使是在发现原发癌之前，癌细胞的确能够通过血液循环广泛转移。因此，如果癌细胞已经存在于身体其他部位，侵袭性手术意义不大。

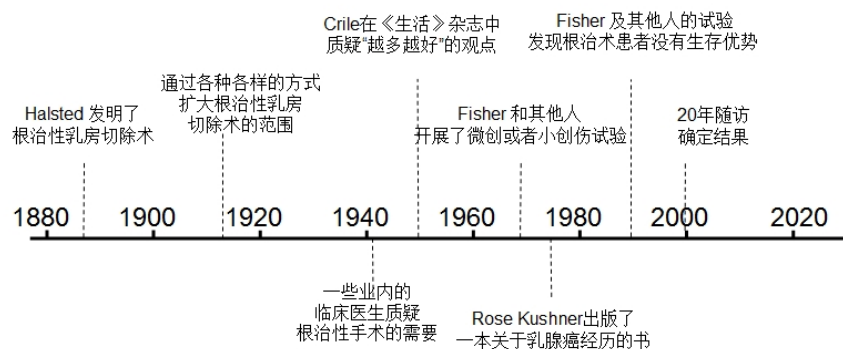
Crile 医生根据其临床判断，提倡采用小范围的局部根治术，而 Fisher 医生与一个日渐壮大的研究小组则以更规范、严谨的方式展开了合作。他们试图通过众所周知的无偏倚（科学）的方法——随机对照试验（见第6章）来肯定或否定根治性手术的价值。他们认为通过这样的研究可以让医学界和公众信服。1971年，直言不讳的 Fisher 医生指出，所有外科医生都应出于伦理和道德责任，通过开展这样的试验去验证其各自持有的理论。对 Fisher 医生的试验长达20年的随访结果显示，在早死的风险这一结局指标上，乳房肿瘤切除术加放疗与乳房全切术的疗效相当。²

随机分配——一个简单的解释

“随机化就是为了最大限度减少偏倚并确保各组患者在所有已知和未知的影响因素上尽可能相近。这样才能确保结局指标的组间差异是由于治疗本身引起的差异，而不是接受每种治疗的患者的差异造成的。

这就避免了医生有意或无意地将某种治疗分配给某类特殊患者，而将另一种治疗分配给另一患者，同时也避免了某一类患者选用某种治疗，而另一类患者选用另一种治疗。”

Harrison J. Presentation to Consumers Advisory Group for Clinical Trials, 1995



对乳腺癌手术“越多越好”方式的挑战

其他国家的研究者们也开展了随机对照试验（见第 6 章）来比较乳房保留疗法与根治性乳房切除术，如 20 世纪 60 年代英国医生 Hedley Atkins 和随后意大利医生 Veronesi 及其同事的研究。试验的结果验证了 Fisher 的结论：即使进行了 20 年的随访，也没有证据表明行根治性乳房切除术的患者有更长的生存期。³ 与英美两国一样，瑞典和意大利也开展了其他多种疗法的随机试验，如手术加放疗与单纯手术的比较，短期化疗与长期化疗的比较。

总之，来自早期试验的结果和详细的实验室研究均支持这一理论：乳腺癌是系统性疾病，早在乳腺肿块出现之前癌细胞已经通过血液在全身扩散了。⁴ 随后，世界上越来越多的医生根据大量的证据确信根治性乳房切除术弊大于利。在 20 世纪最后数十年里，患者和公众的态度也开始转变。在美国和其他地方，在像 Rose Kushner（见第 11 章）一样积极的患者工作带头下，由于得到更加充分有效的信息，来自全世界的患者团体们开始挑战手术中“越多越好”的理念及与之相伴的医学中的家长主义。

患者和卫生保健专家开展的世界范围内的行动卓有成效地挑战了过去随处可见的过度手术。然而令人难以置信的是，依然有开展不必要和破坏性乳腺癌手术的报道，例如，在 2003 年，日本有超过 150 人次实施了这类手术。⁵

截至 1985 年，乳腺癌试验数量的不断增加使人们很难及时跟进所有试验结果。为解决这一问题，牛津大学的 Richard Peto 及其同事将所有试验结果汇总，首次完成了一系列包括了多个试验全部受试者所有信息的系统评价（见第 8 章）。⁶ 现在，乳腺癌治疗的系统评价正在定期更新并发表。^{7,8}

骨髓移植术

然而，破坏性手术的消失并未说明“越多越好”理念的终结，事实上远非如此。20 世纪的最后 20 年期间，一种新的治疗方法——大剂量化疗加骨髓移植术或干细胞移植术被引入。1999 年，《纽约时报》（New York Times）的一篇评论文章总结了这种疗法背后的原理：

“医生从患者身上抽取一些骨髓或红细胞，然后给予患者足以破坏骨髓的大剂量毒性药物。希望借此清除癌细胞，把存储的骨髓回输到体内后，将快速生长，让患者免于死于感染。通过骨髓捐献的方式对血癌患者进行骨髓移植，其疗效早已确立，但这是因为癌细胞仅存在于患者骨髓中。而要用这种方法来治疗乳腺癌却是基于一种全新的、从未经过验证的原理。”⁹

尤其在美国，即使每 100 位患者中会有 5 位死于这种治疗，成千上万的处于绝望中的妇女仍然要求医生和医院为其施行这种非常痛苦的治疗。该治疗通常要花费掉上万美元，并且部分需要自费。最终，保险公司迫于压力只好赔付了部分患者的治疗费，尽管并无证据支持这种疗法是否有效。许多医院和诊所却从中牟取了利益。1998 年，一家医院的年收入为 1.28 亿美元，其中大部分都来自于施行骨髓移植的癌症中心。对于美国医生而言，这不仅是一份名利双收的肥差，还是一个生产论文的富矿。执迷不悟的患者需求更是刺激了这个市场。美国的私立医院争相开展这项治疗，竞争相当激烈，甚至打出了降价广告。20 世纪 90 年代，在美国，即便是招募患者开展临床试验的学术医疗中心也开展了这项治疗。这些值得质疑的疗法竟然成了癌症治疗机构的“摇钱树”。

无限制地开展这种未经证实的治疗还导致了另一个严重后果：找不到足够的患者参加比较该疗法与标准疗法的临床试验。因此耗费了比预期更长的时间才得到了可靠的答案。

努力获取无偏倚的证据

在两项研究中，研究者预期三年招募 1000 名妇女，可实际用了七年……这并不奇怪……参与临床试验的这些患者必须签订一份知情同意书，上面说明了可能的糟糕预后以及并无证据来说明骨髓移植比标准治疗好。进入试验前，你不得不面对这些事实，那绝不是一件容易的事。但是假如没有对照的患者，即不进行随机对照试验，而对患者进行骨髓移植，热心的医生可能会告诉她骨髓移植可以救她的命。尽管患者有权知道真相，但可以理解，她们绝不会去找那些带走她们希望的医生。

摘自 Kolata G, Eichenwald K. Health business thrives on unproven treatment, leaving science behind. *New York Times Special Report*, 1992, 2 October

尽管在这种压力之下要获得无偏倚证据困难重重，部分临床试验仍得以实施，其他证据也得以严格评价。2004 年，对作为乳腺癌常规治疗的大剂量化疗加骨髓移植的累积结果进行了系统评价，发现尚无任何可靠证据证明大剂量疗法有效。^{10,11}

敢于挑战

因此，越多未必越好，这信息仍然很重要。今天，患有转移性乳腺癌的妇女仍然热衷于使用赫塞汀这样的治疗（见上文及第 1 章）。然而，即使在最好的情况下，赫塞汀只能支撑这些患者较短的生存时间，一般在数天与数周之间，而这样做的代价却是严重的不良反应，甚至是治疗本身导致的死亡。^{12,13} 在乳腺癌谱的另一边，过度治疗的趋势也很明显。激进的和不必要的治疗常常被应用于一些癌前病变的情况中，例如乳腺导管原位癌（DCIS），经常在乳腺癌筛查中被检测出（见第 4 章），即使不进行任何治疗，DCIS 终身也不会有任何问题。常规乳腺癌手术需要清扫腋窝淋巴结，而这会引起如淋巴水肿（见第 5 章）这样的影响手臂功能的并发症风险，值得庆幸的是，这种做法越来越被质疑，因为相比于其他疗法并不能有效提高患者的生存率。¹⁴

要点

- 治疗越多未必越好，有些情况可能弊大于利。

第4章 越早未必越好

前三章介绍了未充分验证的治疗是如何造成严重危害的。接下来，我们要讲述用于早期发现疾病的健康人群筛查。筛查听起来似乎很明智，但其在帮助人们避免疾病严重后果和保健方面到底有多大效果呢？虽然筛查对某些疾病的确大有裨益，但有时也同样会带来危害。

本章我们将列举多个实例阐述为什么早期诊断未必一定会更好？为什么许多筛查无益处或效果不确定？筛查的好处是如何被过分推销，筛查之害是如何被回避或忽略的。

从普通人到患者

筛查难免会使部分检测结果呈“阳性”的人成为患者，但不能就此轻易下定论。“如果患者向医生求助，医生应竭尽所能。医学知识本身的缺陷不应该让医生来负责，但如果医生要对患者展开筛查项目则另当别论。在我们看来，除非医生持有结论性证据证明筛查能够显著改变被筛疾病的自然病程，否则不应该开展筛查。”

Cochrane AL, Holland WW. Validation of screen procedures.

British Medical Bulletin 1971; 27: 3-8.

对健康人群的筛查决不能草率，我们应该警惕筛查本身存在的重大弊端。筛查是一种医疗干预，而且提供筛查本身就是一项干预。因为即使拒绝筛查，人们也不免会怀疑自己是否选择正确，这是人的天性。而如果一开始就不提供筛查，人们则不会受到影响。

筛查只有在具备以下有力证据时，才能用来筛查心有疑虑或将来需要治疗的健康人群：
(a) 在可负担的费用范围内其利大于弊；(b) 能够通过一个高质量且高效运行的项目提供（见下文）。¹

筛查远远不只是一个“一次性”的检测。受邀被筛查的人们需要足够无偏倚的相关信息，以决定是否接受筛查，也就是说他们需要知道自己到底被做了什么（见下文）。²

我们可以这样理解筛查：筛查=检测+有效的治疗策略

神经母细胞瘤筛查的教训

神经母细胞瘤（一种主要累及幼儿的罕见肿瘤）筛查的教训在许多方面都很有启示意义。神经母细胞瘤是一种影响身体多个部位神经细胞的肿瘤，患儿的存活率取决于如下因素：累及部位，确诊时的肿瘤扩散范围以及孩子的年龄。1~4岁的儿童患者，总体五年生存率约为55%。³该病有一奇特之处，即其中一种类型的神经母细胞瘤可不经治疗完全消失，此现象被称为“自然消退”。⁴

该肿瘤筛查项目总会让人跃跃欲试，理由如下：（1）一岁前诊断，患儿的预后将优于一岁后诊断；（2）疾病晚期的患儿病情进展比疾病早期的患儿快得多；（3）该筛查试验既简单又便宜，只需要从湿尿布中采集尿样进行检测即可；（4）该筛查试验可检测出90%的神经母细胞瘤患者。⁵

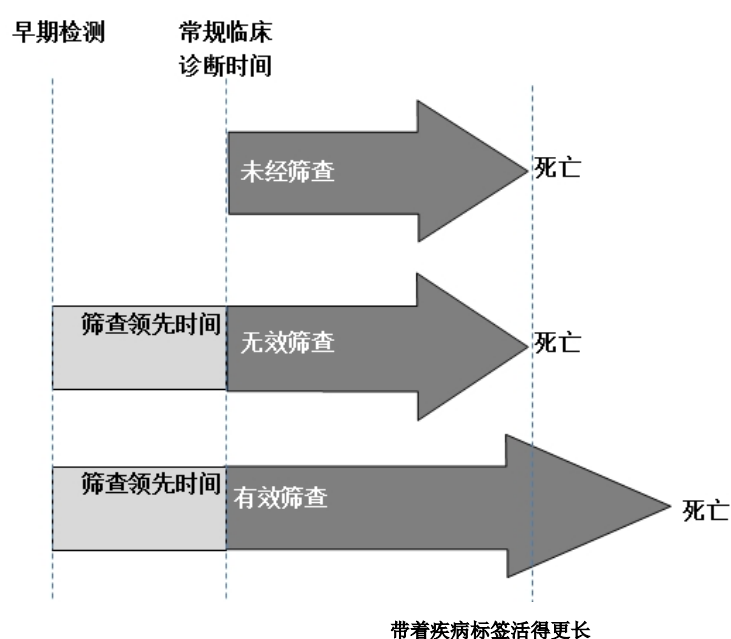
1985年在没有无偏倚（科学）临床试验证据证明筛查有益的情况下，日本首先开展了

针对六个月龄儿童的大规模神经母细胞瘤筛查。在前三年的全国筛查中有超过 337 个婴儿确诊，其中经过治疗后五年生存率为 97%。但 20 年过去了，没有证据表明神经母细胞瘤筛查能降低其病死率。这又作何解释？

在彻查当年日本引入并推广此筛查基于的证据后发现其存在严重缺陷——我们来分析下原因。97%的存活率的确令人惊叹不已，但其实只是“病程长短偏倚 (length-time bias)”的反映，即筛查那些进展缓慢的疾病效果最好（此例中则指的是生长缓慢的肿瘤）。相反，生长快的肿瘤通过筛查检测出来的可能性更小，因为婴儿会出现相应的临床症状（如腹部肿胀），很快便可引起医生的注意。快速进展型神经母细胞瘤可能远比缓慢型严重。但进展缓慢的神经母细胞瘤预后通常较好，包括自然消退（见上文）。⁶

所以这 337 个通过筛查确诊的病例本来的预后就会更好，也没有纳入那些结局会很差的婴幼儿。当然，一些本可自然消退的神经母细胞瘤也会被筛查出来。对于这部分病例，如果未经筛查，就没人知道他们得过此病；但筛查后，这种过度诊断会将得过此病的婴儿归为患者，从而遭受不必要的治疗所带来的伤害。

此外，那些鼓舞人心的小样本研究结果导致日本开展了全国范围内的筛查，而对筛查结果的分析是从肿瘤确诊的时间开始计算存活的长度，而非从婴儿刚出生开始计算。这一点很关键，因为更早的诊断本身不能使患者存活更长，他们仅仅使其带着疾病“标签”生活的时间更长了。换个说法，存活时间看起来更长是因为“疾病钟”开启更早而已。这叫做“领先时间偏倚”，可通过以出生时间代替诊断时间进行结果分析加以消除。



相反，通过加拿大和德国开展的一项纳入三百万名儿童的临床试验获得的无偏倚证据中，研究者并未发现该筛查有任何益处，反倒危害明显。⁷这些危害包括不必要的手术和化疗，两者都会产生严重的不良作用。因此，日本婴儿神经母细胞瘤筛查项目于 2004 年终止。

与此同时，澳大利亚新南威尔士州的婴儿却免受神经母细胞瘤筛查之害。在获知日本那条振奋人心的研究结果后，该州计划在 20 世纪 80 年代开展此项筛查，但如上所述，日本

的研究结果显示的是从诊断后开始存活的时间更长，而非从出生时间开始分析。因此，一位澳大利亚专家便重新分析了该结果，并摒弃先前的分析方式，转而以婴儿的出生时间为起点进行分析。结果显示，接受筛查与未接受筛查婴儿的存活率并无差异。该结果成功地说服了新南威尔士州的相关机构，使其放弃了此筛查项目，因而使婴儿们免于不必要的伤害并节省了不必要的医疗支出。

不要推测早期检测就是有价值的

“神经母细胞瘤的筛查说明，人们很容易陷入‘如果能够早期检测出某种疾病，筛查就一定有价值’的陷阱……这两项研究表明，神经母细胞瘤的筛查不但没有价值，还会导致‘过度诊断’，且会查出那些原本可能自然消退的肿瘤。两项研究也都提到筛查组儿童因接受治疗而出现了严重并发症……希望将来在考虑开展其他筛查项目时能汲取这些教训，如前列腺癌的筛查。”

Morris JK. Screening for neuroblastoma in children.

Journal of Medical Screening 2002;9:56.

权衡利弊

筛查有利的实例很多。在成人筛查中应用最广的可能要数初级保健中常规开展的心脏疾病和卒中危险因素检查。高质量证据显示，高血压、高胆固醇、吸烟可增加患这些疾病的风险，因此找出有上述危险因素的人群并给以建议和治疗，可预防心脏病和卒中。

苯丙酮尿症筛查：确实有益

新生儿要进行遗传性疾病苯丙酮尿症（PKU）的常规筛查。苯丙酮尿症患者不能够代谢苯丙氨酸——一种存在于日常食物（如牛奶、肉类、鱼类和蛋类）中的物质，若不治疗，血中苯丙氨酸就会累积而引起严重的、不可逆的脑损伤。苯丙酮尿症的检测需采集数滴婴幼儿的足跟血，用于实验室分析。若检测结果是阳性的，且经进一步检测后确诊，则需对患者进行特殊饮食疗法来帮助其健康成长。

腹主动脉瘤筛查：进一步治疗

对于老年人，腹主动脉瘤筛查也是有益的。主动脉是人体的主要血管，从心脏发出，行经胸腔再到腹腔。部分人群随着年龄的增长腹主动脉壁会变薄，并开始向外膨出，形成动脉瘤（常见于 65 岁及以上的男性，很少有明显症状）。较大的动脉瘤在毫无征兆的情况下常常突然破裂引起大出血，并导致死亡。⁸

老年人主动脉瘤发生率的相关证据可作为引入筛查项目的依据。例如，在英国，男性（而非女性）到 65 岁时就要进行筛查性的超声扫描，通过扫描可查出大的主动脉瘤，进而可接受专业建议并采取治疗（多为手术治疗）；主动脉瘤较小者则需通过进一步扫描进行监测；对于主动脉瘤未增大者，则无需再次筛查。筛查与手术的质量至关重要，主动脉瘤手术是主要治疗方法，但如果手术并发症发生率很高，则是弊大于利了。

乳腺癌筛查：虽已大量开展但仍存有争议

在许多国家，常规乳房 X 线摄影筛查得以广泛应用，人们也当然认为乳房 X 线摄影筛查一定是基于利大于弊的可靠证据了。然而正如一位美国公共卫生专家 2010 年所述：“没有哪项筛查有过比乳腺癌筛查更加细致的研究。在过去 50 年中，有超过 60 万名女性参与了 10 个随机对照试验，每个试验都做了近 10 年的随访。即使做了大量超常的研究努力，但讽刺的是，乳房 X 线摄影筛查依然是医学界最受争议的问题之一。”⁹

乳房 X 线摄影为何会饱受争议呢？一个根本原因在于它曾被筛查的人员和患者团体作为“明智选择”而“售予”女性。参与者得到的信息，强调了乳腺筛查的益处，而其伤害、局限性和不良结局则被忽略。¹⁰然而，乳房 X 线检查不仅可以实现早期诊断，而且与前列腺癌筛查（见下文）极为相似的是，还可诊断出那些终生都没有明显症状的癌症患者，同时也无法避免假阳性结果的出现。

最可靠的证据来自对随机对照试验（将参与试验的女性随机分配到临床试验的筛查组与未筛查组）结果的系统评价。有趣的是，系统评价的结果显示，如果有 2000 名女性进行了十年的定期筛查，有 1 位女性将能从中获益并可避免死于乳腺癌。但同时，筛查也可能使得 10 位健康女性成为“癌症患者”并接受不必要的治疗。乳房 X 线摄影实际上检测出的是这些不可能发展成为癌症的进展极缓的（甚至不会有任何进展）病变，这些被误诊的健康女性将进行乳房部分切除（更甚者为全切），还得经常接受放射治疗或化疗。¹¹

此外，这 2000 人中有 200 人会得到错误的警告，沉重的心理压力将持续至确诊，甚至确诊后也无法消除。在建议妇女进行乳腺自我检测或加强预防意识时，医生同时也会推荐乳房 X 线摄影，然而结果显示这些方式均弊大于利。¹²

一位英国公共卫生专家提到，“患者个人从乳房 X 线摄影中获益的可能性很小，这种现象并未得到广泛理解。部分原因是由于乳房 X 线摄影筛查组织者的不恰当解释，他们认为着重强调阳性结果可确保受试者对筛查的依从性”。2010 年，在对现有证据进行评估后，他提出：“乳房 X 线摄影确实可以挽救生命，且在老年女性中更有效，但其危害也不容否认”，他提到的危害便是过度诊断和假阳性。此外，他还对乳房 X 线摄影筛查最新且尚未被客观评价的所有研究结果进行了严格评价。¹³在我们静候这一客观评估结果的同时，乳房 X 线摄影依然继续开展着。但最起码，受试女性应该被提供充分的有关利弊的信息以决定（如果愿意，也可同家人及医生一起决定）是否参加筛查。

前列腺癌筛查：好处不知道，坏处很明显

前列腺癌是全球第二大常见的男性恶性肿瘤，¹⁴大体分为两型：一种为侵袭型前列腺癌，此类型扩散迅速且病死率高；另一种为生长缓进型前列腺癌，该类型发病率更高，但不会危及生命。理想情况下，筛查试验应该检出危险程度更高的恶性肿瘤并使其接受治疗，而不是生长缓慢的前列腺癌。因为无论哪一种类型前列腺癌的治疗都会给患者带来痛苦的副作用（如小便失禁和阳痿），所以如果癌症一开始就不会导致健康问题的话，我们将付出昂贵的代价。¹⁵

多数前列腺癌患者血液中前列腺特异性抗原（PSA）的浓度会升高，但用于区分患或不患有前列腺癌的 PSA 浓度临界值尚不明确¹⁶，而且有近 1/5 临床确诊的患者 PSA 水平正常；此外，除了名称之外，PSA、诊断前列腺癌并无特异性，比如那些非前列腺癌的肿瘤、

感染甚至一些非处方止痛药都可引起 PSA 水平升高。基于以上理由，PSA 作为筛查检测有明显的局限性。

前列腺癌的过度诊断

“前列腺癌被视为过度诊断的典型实例。但这并不意味着不存在因早期诊断而免于过早死于前列腺癌的患者。但是……我们很难提前知道哪些男性可获益于筛查，哪些会接受伴有严重不良后果的不必要治疗。根本问题在于，通过前列腺癌筛查和检测，被发现的前列腺癌患者空前的多；而奇怪的是，这些癌症大多不会危及生命。在过去，这些人根本不会知道自己患有癌症，他们可能死于其他原因而非前列腺癌。那些无关痛痒的前列腺癌被查出后，被诊断为前列腺癌的患者较之前就多了很多，于是便有了“过度诊断”之说。这是所有考虑接受前列腺癌筛查的人面临的主要困境。

Chapman S, Barratt A, Stockler M. Let sleeping dogs lie? *What men should know before getting tested for prostate cancer*. Sydney: Sydney University Press, 2010: p25

然而在医务人员、患者及筛查用品销售公司的积极推动下，常规 PSA 测试用于健康男性前列腺癌筛查被许多国家广泛采用。在美国，提倡 PSA 筛查的呼声很高，据估计美国每年有 3000 万名男性接受筛查，并深信这是明智的做法。那么把 PSA 作为前列腺癌早期检测指标以改善患者结局的证据何在呢？它又会带来哪些危害呢？

而今，已有关于 PSA 利弊的高质量证据。2010 年，有研究对所有相关试验的结果进行了系统评价，发现尽管 PSA 筛查可提高前列腺癌的被检出率（如预期的那样），但并无证据说明其对该病的病死率或总死亡率有影响。¹⁷

PSA 发现者的话

“PSA 检测的普及给公共卫生体系带来了巨大的损失。这也成为了我最不堪回首的往事，因为正是在 1970 年发现了 PSA……”

美国人为检测前列腺癌耗资巨大，每年在 PSA 检测上至少花费三十亿美元，大部分由老年医疗保健（Medicare）和退伍军人管理局（Veterans Administration）支付。

可见，人们对前列腺癌重视有加。但我们不妨看看下面这些数字：美国男性一生患前列腺癌的几率为 16%，而因此死亡的仅有 3%，原因是大部分前列腺癌生长缓慢。换句话说，一个男性如果能活到老，又得了前列腺癌，那他死于这种病因的可能性远远小于其他病因。

即使如此，PSA 检测筛查出前列腺癌的概率犹如投掷硬币一般。几年来我不断试图让人们明白，PSA 检测不仅无法筛查前列腺癌，同时也不能对两种不同的前列腺癌做出区分（一种是致命的，另一种则是非致命的）。”

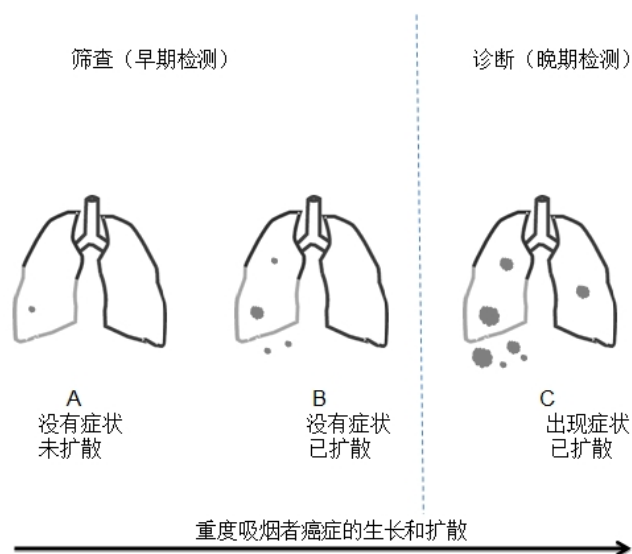
Ablin RJ. The great prostate mistake. *New York Times*, 10 March 2010.

那么这一证据能否逆转 PSA 筛查的现状呢？多年来，PSA 的发现者 Richard Ablin 一直致力于改变 PSA 检测的现状。他在 2010 年发表评论：“我从未想到自己四十年前的发现会驱使人们追求利益，给公共医疗带来如此大的灾难。如今的医学界必须面对现实，停止对 PSA 筛查的不正当使用。只有这样才能节省上百亿美元，让数百万病人不必接受痛苦的治疗。”至少每个人在进行 PSA 筛查之前，都应知道其局限性和可能的副作用。正如某个专家小组指出的：“对于 PSA 筛查，我们应该这样建议：该检测不能告诉他们是否患有致命性的癌症，但可能会导致他们接下来要进行一系列本来可以避免的检查和治疗。”¹⁸

肺癌筛查：早，但还不够早

筛查可早发现疾病，但未必够早，所以有时候并不能产生任何影响（见下图）。

有些癌症如肺癌，可在患者未出现任何症状或任何检查都没能检测出之前就扩散。下面用胸部 X 线检查来解释这一问题（见图 B 阶段）。20 世纪 70 年代，几项针对重度吸烟者的大型研究显示：尽管胸部 X 线检查可以更早地发现肺癌，但没有证据证明其可降低肺癌的病死率，而胸部 X 线检查发现的肺癌早已扩散至肺外组织。所以，虽然这些患者被确诊为癌症的时间提前了，也更早地接受了治疗，但没有证据表明这对他们的预期寿命有任何影响。



癌症的生长与扩散：如果我们可以检测和治理扩散之前的无症状病例（A 情况），筛查可能是有效的；如果仅能发现已经扩散的无症状病例（B 情况），则筛查是无效的。

“推销”筛查

“推销”筛查很容易：夸大疾病风险引起恐慌，再放大筛查效益（避谈其害处）给人希望。对于癌症筛查就更容易了：因为这是人们最害怕的疾病。大家都知道这样一句话：早期诊断是最好的预防措施。如果你质疑这个观点，也许有人会觉得你脑子有问题。

以前一个美国癌症协会的广告曾这样说道：“女性朋友们，如果你已经年满 35 岁，请一定要做乳房 X 线摄影检查。除非，你还不相信其重要性。如果是这样的话，你到时需要检查的就可能不只是乳房了。”

“推销”筛查的信息无处不在。新闻时不时地就会给大家讲述那些声称“做了早期诊断后幸存下来”的名人轶事，但却很少告诉大家因过度诊断和过度治疗而受伤害的例子。

大众杂志也会报道一些催人泪下但却完全不具代表性的故事，如关于某些年轻的乳腺癌患者及其对死亡和离开年幼孩子的恐惧。

医疗中心把筛查作为一种商业策略，通过免费检查吸引患者。和上述美国癌症协会标语一样，公共服务宣传实则皆为自己代言。

Woloshin S, Schwartz LM. Numbers needed to decide.

Journal of the National Cancer Institute 2009;101:1163-65.

最近，一个纳入了 5 万 3 千名目前或曾经的重度吸烟者的大型随机对照试验比较了胸

部 X 线筛查和螺旋 CT 扫描（一种特殊类型的 CT 扫描方式）的筛查效果，两组人员都被安排了每年三次的筛查。结果显示，螺旋 CT 扫描相对胸透能更早地诊断出肺癌，对于其中小部分患者此诊断够早（处于上图中的 A 阶段），从而能使患者及时接受治疗（螺旋 CT 扫描组的肺癌死亡人数为 346，胸部 X 线检查组为 425）。然而，这种有益结局的代价是大量人群被误诊为肺癌。整体来看，通过对上述人群 8 年以上的随访发现，通过筛查，每 1000 人中会有 3 人可避免死于肺癌，但有 13 人尽管早期检测出肺癌但仍会死于肺癌，还会出现 233 个有待进一步检查的假阳性结果。¹⁹

基因检测：有时有效，但大多数时候不靠谱

不久之前，“基因检测”已基本被限于罕见的单基因疾病，如儿童肌肉萎缩症——杜氏肌营养不良症或亨廷顿病（一种通常在中年时期才会发病的进行性神经系统疾病）。基因检测可用于诊断此类疾病，也可用于筛查有家族史的健康人群，如果家族史提示他们的疾病发生率高于平均水平，检测可指导这类人群制定家庭计划。

但多数疾病并非仅由单个基因缺陷所致。疾病通常取决于多个基因变异的相互作用，并且还会受到环境因素的影响。只有在基因变异“严格”组合并伴有环境诱导因素时，才会发展为疾病。¹

将众多疾病归因于异常的基因这一理论尽管复杂难懂，媒体和推销员们仍向消费者鼓吹遗传风险分析的优点和便捷。你需要做的只是采集唾液标本送到 DNA 分析公司并付费，他们便会将你的基因分析结果反馈给你。但那些信息未必能帮你（或你的医生）正确地预测发病风险，更别说可采取任何措施了。这种做法显然有悖于筛查的目的（见下文）。筛查结果反而可能使你忧虑不安，进退两难，甚至影响到你的家人。澳大利亚一位卫生新闻工作者报道：“对于任何关注‘生活蠕变医疗化（creeping medicalisation）’的人来说，基因检测市场无疑是最前沿的信息；这看似无害的技术，可将健康人群变成恐慌的患者，通过多基因遗传检测将他们列入了患病或早逝的范畴。”²⁰

别和你的基因打赌

相信单个（或少数）基因突变就采取行动，就好比看到一张牌后就把所有的钱都压在牌局上一样。你不知道哪种遗传因素会引起疾病，也不知道环境因素会有哪些影响；而人类的基因远不是 5 张牌的问题，而是关乎 2000 多个基因和千变万化的环境因素。另外，一个基因变异的作用可通过生活方式、家族史或其他保护性基因抵消。我们很多人都有缺陷基因，但却从未见其致病。

Sense About Science. *Making sense of testing: a guide to why scans and other health tests for well people aren't always a good idea*. London: Sense About Science 2008, p7. Available from www.senseaboutscience.org

筛查目的是什么？证据为什么重要？

上述例子告诉我们：在采用正被广泛使用的筛查之前，我们应该先想一想筛查的主要特征是什么，筛查目的何在。被筛查者既没有（或没注意到）所筛疾病的症状或体征，也没因相关不适就医。筛查的主要目的是通过检查找出可以从治疗中获益的人，以此来降低死亡和将来得某种疾病的风险。^{1,21} 筛查并不只是为了更早地诊断疾病，因为这样做可能无益，

甚至有害。

早在 1968 年，WHO 的一份报告就总结了评价筛查试验价值的基本标准。²² 针对现有卫生保健模式，这些标准已进一步完善。筛查目标人群需要了解检查的相关信息（包括可能的危害、后果、局限性和可能的益处），从而做出知情决策。WHO 的筛查标准总结如下：

- 筛查的疾病应是重大的公共卫生问题，如：该病极为严重和（或）关系到很多人
- 能够发现疾病的早期阶段
- 存在有效且可接受的治疗措施，筛查有助于改善其结局
- 存在有效可靠且能被目标人群接受的疾病检测手段
- 筛查项目质量高且在当地实施符合成本效益
- 目标人群所得信息应无偏倚，基于有力的证据，并明确利弊（如过度诊断带来的过度治疗）
- 参与筛查遵循自愿原则，即参与者可合理拒绝筛查
- 筛查造成生理或心理伤害的概率应小于其产生效益的概率
- 筛查发现异常后应有充分的医疗条件进行诊治

这一标准加强了本章开篇的观点：任何引入筛查项目的决定都应基于高质量的证据，并且应兼顺利弊。

一场筛查闹剧

2009 年，一位刚退休且长期对卒中预防很感兴趣的神病学教授得知邻居们收到了卒中和其他心血管疾病并发症筛查的邀请传单。该传单是一家血管筛查公司发的，邀请他们到当地一所教堂进行一系列检查（价格为 152 英镑/230 美元/170 欧元）。因为传单上的部分信息有误导作用，且出于好奇，他想亲自去探个究竟。

“首先进行的是主动脉瘤（输送血液出心的血管发生扩张）的超声筛查，由一位女性工作者操作，她并不愿意谈论查出主动脉瘤的意义。接下来是脚踝与手臂血压的测量，用于发现‘血液循环问题’……紧接着是非血管类疾病的附送检查：踝关节骨质疏松检查。然后又是检测“心房问题”的心电图检查……最后是颈动脉（颈部的动脉）超声检查“斑块沉积”。当问及这些检查项目的意义时，他们告诉我：“这会形成血栓并导致卒中。”在推荐治疗时，他们仅给出了血液稀释药物的含糊回答，而未提供任何手术相关的信息，直到我开门见山地询问是否可以手术，他们才回答“是”。“这有风险吗？”我继续问道。他们回答说：其风险取决于我的全科医生的后续工作，并需要与医生讨论此次检查中所有的异常。

所有操作没有任何隐私（主动脉瘤筛查例外）……现场看似没有医生出席，筛查人员也不打算（或不愿意）讨论有关假阳性与假阴性结果、异常情况的预后或治疗的风险与益处。

这纯粹是为了商业利益而进行的筛查。他们会在 21 个工作日内将结果（无论真假，甚至是未经检测的结果）“扔”给我，然后由我的全科医生来判断异常的结果……毫无疑问，在筛查者没有与被筛查者讨论任何异常的结果或对他们负一丁点责任的情况下，整个“筛查闹剧”极易引发那些神经脆弱人群的焦虑。

Warlow C. The new religion: screening at your parish church. BMJ 2009;338:b1940

还有谁可以免于检查？

全身 CT 扫描

全身计算机断层（CT）扫描即进行头、颈、胸、腹部及盆腔 CT 扫描，是私营诊所中的检查项目之一。这些检查经常在未参考全科医生或初级保健医生意见的情况下，直接开展。全身扫描常被称为能将可能发生的疾病扼杀在摇篮中，或者提供正常的结果以图安心。在无症状体征的人群中开展此类检查不仅价格昂贵，而且没有证据表明其有任何益处。

此外，辐射暴露不容忽视，CT 扫描的辐射量比 X 线检查高 400 倍。鉴于此，2007 年英国环境辐射医学委员会（Committee on Medical Aspects of Radiation in the Environment, COMARE）强烈建议停止对无症状个体进行全身 CT 扫描。

2010 年，政府部门讨论之后，宣布对全身 CT 扫描的应用采取更严格的规范。同样，美国食品药品监督管理局也提醒公众，这些扫描对健康人没有确切的好处；而且指出“多数人并没意识到接受全身 CT 扫描检查并不一定能如其所愿，使其‘安心’，也未必能得到预防疾病的信息，异常的结果不一定有严重的影响，而正常的结果也可能不准确”。^{23,24,25}

取得平衡

在过度筛查和不筛查两者间取得平衡并非易事，也无法避免会引起部分不满。为了使整个人群获益，所有卫生保健系统都需要“精打细算”。这一基准不仅意味着筛查项目的引入必须基于有力的证据；同时随着越来越多的证据积累和环境的改变，还必须不断进行评估以确定其是否仍然有利。另外值得慎重考虑的一点是，筛查项目是应该提供给大范围人群还是只需针对疾病的高危人群。

要点

- 早期诊断未必能带来更好的结果；有时反而使问题更糟
- 只有在良好证据显示其有效的情况下，才能开展筛查项目
- 不开展筛查也许是最好的选择
- 受邀参加筛查的人群需获得全面的关于利弊的信息
- 筛查的益处往往被夸大
- 筛查的弊端常未予以重视或被忽视
- 讨论筛查的利弊及风险很必要

第 5 章 应对疗效的不确定性

本章我们将讨论各种各样宣称那些有效的治疗，无论新旧，都会面临的一个问题：疗效的不确定性。例如，几乎没有人会质疑心脏病患者常规吸氧的效果，但并没有有效证据支持这一疗效，甚至有些证据证明常规吸氧还是有害的。这种治疗效果不确定的问题从未彻底得到解决¹，其他许多治疗效果也存在争议。

效果明显的治疗：少见且容易发现

仅在极少数的情况下，对于治疗是否有效的证据是明确而毋庸置疑的。²通常，这种治疗效果显著且起效快。以心室纤维性颤动（心律失常疾病）为例，当发生此病时，心室（位于心脏下方的两个腔）的收缩会严重失调。该病属于急诊情况，可能在几分钟内导致死亡。将除颤器置于胸部，通过直流电击心脏可恢复心脏的正常节律。一旦成功，治疗立即起效。

其他显著效果的治疗（见第 6 章第 70 页）包括减轻脓肿疼痛的脓液引流法、严重出血性休克的输血疗法、糖尿病的胰岛素（胰腺分泌的一种激素）疗法等。20 世纪 20 年代之前，糖尿病患者寿命短且受尽折磨，常因血糖控制不佳而极度消瘦。但很快，基于动物实验结果，人们将胰岛素应用于患者，在治疗上取得了巨大的成功，当时胰岛素的疗效近乎神奇。发生于同一年代的还有肝脏制品治疗恶性贫血（后来发现此肝脏制品中含有维生素 B12）。当时这被视为一种致命性贫血疾病，红细胞会逐渐降低到非常危险的水平，使患者极度苍白和虚弱。应用肝脏提取物后，患者便快速恢复了。如今，维生素 B12 已成为此类贫血疾病的常规处方药。

本世纪初的一些例子也说明了类似的显著疗效。

鲜红斑痣的激光疗法

胎记也称为“鲜红斑痣”，是皮肤上永久性畸形扩张的血管引起的。这种疾病好发于面部，通常很顽固且随着年龄增长会越来越黑，极大地影响了面部美观。患者们尝试了大量的治疗如冷冻、手术、放射等，但是效果甚微，甚至还产生了许多副作用。激光疗法的引入带来了显著的效果：对于大多数类型的斑痣，一次激光治疗就可以见效，而且激光散发的热对周围组织造成的损害只是暂时的。^{2,3}

伊马替尼治疗慢性粒细胞性白血病

伊马替尼治疗慢性粒细胞性白血病的效果也十分显著。^{4,5}

20 世纪 90 年代末伊马替尼问世，在此之前，针对这类白血病的常规疗法效果极差。将这种新药给那些常规治疗后效果不佳的患者，疗效非常显著。伊马替尼可稳定患者的病情，相比于伊马替尼前的治疗，这种药还可以延长寿命，且副作用很轻。现在伊马替尼已作为慢性粒细胞性白血病的首选药。

妈妈的吻

一些低技术含量的方法也有显著的疗效。小孩子有时候会把一些塑料玩具、珍珠等小

东西塞进鼻子里，但他们自己很难将这些异物从鼻子里“吹”出来。“妈妈的吻”可以帮助取出异物，而且这个方法十分简单、有效。父母使用这个方法时要先捏紧孩子没有被堵住的鼻孔，然后再向孩子的嘴巴吹气。^{2,6}

草莓胎记的新疗法

有时一些效果显著的治疗是在无意中发现的。例如发生在婴儿身上的血管瘤很像鲜红斑痣，也是因血管发育不全或发育畸形所致。出现血管瘤时，小血管会聚集成团。血管瘤好发于头部和颈部的皮肤，在内脏器官如肝脏上也会有。在皮肤的病变因其颜色红亮，突出皮肤，常称为草莓胎记，婴儿出生时不易察觉，但会在出生一周左右显现。出生以后的前三个月内，血管瘤生长迅速，之后速度会变慢。大多数情况下，孩子长到五岁时草莓胎记会自行消失，留下淡淡的红色印记或略松弛的皮肤。

但是，有些血管瘤因部位特殊则需要治疗，如长在眼睛里会影响视力，长在鼻子里会堵塞鼻孔。此外，引起并发症的血管瘤也需接受治疗。例如，溃烂的血管瘤可能会发生感染，病变范围大的患者可引发心力衰竭，因为心脏须大量增加病变血管的供血。

前些年，类固醇仍然是血管瘤药物治疗的首选药。2008年一些医生偶然发现了另一种治疗方法，效果非常显著。对于病变累及整个脸部和眼部的血管瘤患儿，过去一直采用的是类固醇治疗。尽管采用了该疗法，婴儿病情还是会恶化为心力衰竭。因此，为了治疗心力衰竭，医生们开始给婴儿使用一种叫心得安的常规治疗心力衰竭的药物。然而令他们大吃一惊的是，用药24小时内血管瘤的症状得到了改善，用药一周后，肿瘤的面积大大缩小，且婴儿能够睁开眼睛了。心得安治疗六个月后，血管瘤居然消退了。第二年，医生们给十来个孩子使用心得安后也同样成功治愈。鉴于此，其他医生也给一些孩子使用了心得安，并获得同样的效果。目前人们正在制定对更多婴儿开展心得安研究的方案。^{7,8}

科学进展不应该成为新闻头条

科学从根本上讲就不适合作为新闻，因为科学本质上是一门揭示事物客观特征的学科，通常不会突然就取得重大进展或划时代的突破；科学的发展是循序渐进的，因为需要不断研究新的内容和理论，而且需要基于大量不同学科和不同层面的证据。然而，媒体仍迷信科学的“重大突破”。

Goldacre B. *Bad Science*. London: Fourth Estate, 2008, p219.

效果一般的治疗：常见且不明显

多数治疗都没有显著的效果，需要开展科学试验进行评价。有时，一种治疗可能只在某些情况下效果显著，但其他情况下却未必。

尽管维生素 B12 治疗恶性贫血疗效确切（见上文），但患者是否需要每季度或更频繁地接受注射，至今仍有争议。只有通过严格的对照试验才能回答这一问题。再如，髋关节置换术后的镇痛效果十分显著，但不同类型人工髋关节间的相对优势却很难把握，而这些细微的差别又非常重要，例如，某些类型的人工髋关节更易磨损。另外，对于前面提到鲜红斑痣激

光疗法，也仍存在一些问题。虽然目前这个疗法依然是“金标准”，但还需研究为什么几年后病变部位会再次变黑，不同激光的不同效果以及治疗是否可能需要结合皮肤冷却。^{9,10}

虽然治疗剂量的阿司匹林能够有效持续降低心脏病患者的死亡风险，但服用阿司匹林预防心脏病和卒中是否利大于弊还取决于患者是否有潜在的心血管疾病。因而需要权衡服用阿司匹林的利（如降低心脏病、卒中及死于心血管疾病的风险）和弊（如出血，尤其是脑出血引起的脑卒中、肠道出血等）。对于已患有心血管疾病的患者，服用阿司匹林明显利大于弊。若健康人群服用，还不确定（见第7章）。¹¹

当医生出现分歧时

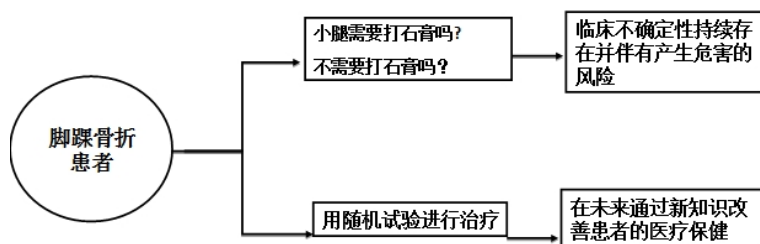
对于许多疾病和症状，治疗的效果如何、哪种方法对患者最恰当，还存在很大的不确定性。但这并不妨碍一些医生对一些疗法固执己见。即使与其他医生出现分歧，也不会改变自己的看法。因此，会导致对同一种疾病，医生的治疗却五花八门。

20世纪90年代，本书作者之一 Iain Chalmers 在美国度假时扭伤了脚踝，当时由一名骨科医生进行了诊治。这位医生先将患足用夹板暂时固定，并建议在肿胀消除后用石膏固定小腿六周。数日后，Iain Chalmers 回到国内，前往当地的骨科诊所就诊，一位英国骨科医生毫不犹豫地否定了这个建议。这位医生说，将患足用石膏固定是完全不合适的。这两种意见显然大相径庭，受此启发，Iain Chalmers 邀请该医生可否参与一项对照试验以找出到底哪种方案更好。这位英国医生却回答道，对照试验针对的是那些不确定自己判断是否正确的人，而他确信自己是正确的。

专家意见为何会出现如此大的分歧？患者又该如何选择？每位医生都坚信自己实施的治疗完全正确。然而，医生们的观点却存在明显差异，这清楚地暴露了整个业界对于普通骨折的最佳治疗方案尚存在不确定性。是否有证明哪种治疗措施更好的高质量证据？如果有，这两位医生是否知道这些证据？还是根本就没有人知道哪种治疗更好？（见下图）。

也许是这两位医生对治疗结局持有不同的价值观：美国医生可能更关注疼痛是否能缓解，所以建议用石膏固定，而英国医生可能更担心是否会出现肌肉萎缩（用石膏固定肢体易出现）。如果真是这样，为什么两位医生都没有询问 Iain Chalmers 哪种结局对他来说更重要？时隔二十年后，这种常见疾病依然存在治疗的不确定性。¹²

以上例子还存在几个问题。首先，是否已有比较两种不同治疗的可靠证据？若有，这些证据是否比较了关系 Iain Chalmers 或其他患者选择和价值观的重要结局（如缓解疼痛、减少肌肉萎缩等）？如果没有需要的证据又该怎么办？



医生应该做什么？

直面不确定性：一个生死攸关的问题

“不能直面疗效的不确定性会造成原本可避免的大伤亡。如果当初引进安定和苯妥英钠为抗惊厥药物时，能将其与已有几十年使用历史的硫酸镁进行比较，数以万计的妇女就能免受疾病之苦甚至死亡。同样，在用于治疗外伤性脑损伤的全身性类固醇激素疗法被广泛采用之前，如果能对这个治疗进行评估的话，也能使成百上千的患者免于不必要的死亡。以上两个例子说明为什么医生需要有一种职业责任感来解决疗效不确定性的问题，类似的例子还有很多。”

Chalmers I. Addressing uncertainties about the effects of treatments offered to NHS patients: whose responsibility?
Journal of the Royal Society of Medicine 2007; 100: 440.

有部分临床医生很清楚在没有可靠证据时应该怎么办，并且会同患者共同探讨各种治疗的不确定性。例如，有一位脑卒中治疗的专家评论到：“尽管现有研究证据表明患者在卒中单元接受治疗后恢复得更好，但对于很多类型的患者，是否应该使用溶栓药仍然不确定”（见 139 页第 11 章）。当他与患者讨论如何选择治疗方案时，他会向患者解释使用这些药物可能利大于弊，但对某些患者而言可能弊大于利。他进而解释了为什么需要与某些治疗到底会产生有利还是有弊的结果并不清楚的患者进行交谈，他只在有严格对照试验结果的前提下，才推荐使用某种药物，因为设计严格的对照试验可减少不确定性。¹³但是关于溶栓药，仍存在多种不确定性。¹⁴

解决不确定性是门专业

“医学专业性的重要体现在于能够鉴别和解决医学中存在的确定性。医学专业人员每天要应对各种疾病发病机理、诊断和治疗方面的不确定性。但是他们很难公开承认医学在以上领域一直都存在不确定性，一些专家对于承认不确定仍然会感觉到不舒服，尤其是在治疗患者时。不确定性同时极大地推动了医学研究，促进了人类健康，这也是医学研究理事会的主要任务。未来，医学专业人员获取自己领域内最新研究结果的能力将越来越重要，只有这样，他们才能知道目前仍存在哪些不确定性、用于解决不确定的正在开展的研究有哪些或需要怎样的研究。总之，未来专业化的特点将会是立足于患者利益的研究意识。相信将会有医学专家积极参与研究，但与此同时也应鼓励周围的人参加，如果合适的话，还可以让自己的患者积极参与医学研究计划的制定、并在日常实践中应用研究所得的结果。”

咖啡因治疗早产儿呼吸困难

相同的疾病，其疗法却五花八门，这清楚地说明了医务人员对这些不同疗法孰优孰劣的不确定。并且，由于这些疗法在医务人员当中根深蒂固，想通过科学试验解决不确定性，也许需要很长时间。咖啡因用于早产儿的治疗就是一个有力的例子。早产儿通常不能正常呼吸，有时还会出现呼吸暂停，称为早产儿呼吸暂停，最易发生于妊娠小于 34 周的早产儿。20 世纪 70 年代末，咖啡因疗法被证实能够减少疾病的发作，随后一些儿科医生便开始采用此方法。

然而咖啡因疗法的效果仍然存在争议。尽管科学试验显示咖啡因可以减少窒息的发作次数，但很多儿科医生认为儿童窒息的发作并不十分严重，不需要使用这类药，还有一些医生担心婴儿服用这种药可能不安全。这意味着，一部分早产儿会接受此治疗而另一部分则不会。直到咖啡因疗法问世 30 年后，一项大型研究才评估了该疗法的不确定性，结果发现，咖啡因这种简单疗法不仅能减少呼吸困难，更重要的是，明显提高了婴儿免于脑瘫和发育迟缓的长期存活率。如果当初咖啡因疗法问世之初就对其疗效进行充分评估的话，那么就会大大减少有缺陷婴儿的数量。^{15,16}

抗生素预防早产

科学试验的结果可能会证明，人们对某些疗法有利的期望，或无害的假设，往往是不对的。医生们给患者开处方时，总是希望取得最好的效果，特别是在患者处于绝望时医生可能带来希望。例如，有理论称“隐性”感染（亚临床感染）可能诱发早产。这促使某些医生为孕妇开抗生素，因为希望这个方法可以延长妊娠期。然而从未有人认真想过这样用抗生素是否会带来更严重的问题。的确，证据显示孕妇们愿意使用抗生素，她们认为“就试一试吧，反正不会有坏处”。

后来人们开展了一项针对此疗法的科学试验，其结果具有明确的临床意义。首先，研究表明该治疗没有任何益处。而且更为严重的是，长期随访结果发现曾使用抗生素组的婴儿比对照组更容易出现脑瘫及语言、视觉和行动障碍。自从给孕妇开始使用抗生素后的几十年间人们都没有发现这些风险，但也没有通过科学试验获得评价其疗效的充分证据。一般情况下，对于同一种未经充分验证的疗法，患者通过“常规的”临床实践接受治疗与通过参与研究的方式接受治疗相比，前者风险更大。换句话说，如果患者不参加科学试验而服药，那么他们通常会遭遇更大的风险。^{17,18,19}

医生就处方中猜测成分的对话

下面是一段虚构的医生对话，其中一个全科医生说道：“很多我们的工作都是一种猜测，相信你我都不喜欢这样。证明治疗是否有效的唯一办法就是开展恰当的试验，但此工程浩大。那我们该怎么办呢？那就只能按照我们的想象来。我确信有些时候这样做没问题，因为这依据的是我们的临床经验等。然而其他时候我们很可能把错的做法当成对的。而且，由于我们所做的不能称为试验，没有任何规范，我们不能从中学到什么。”

Petit-Zeman S. Doctor, what's wrong? Making the NHS human again. London: Routledge, 2005, pp79-80.

乳腺癌

乳腺癌的治疗（见第 3 章）又是一个有关疗效不确定性的典型实例。医生对手术、放疗和化疗的使用差异相当大。非常早期阶段的乳腺癌患者及“伪癌症”患者的最佳治疗方案尚不明确，理想的腋窝淋巴结清扫范围也未明确。²⁰除此之外，还有诸多患者特别关心的问题仍未得到充分验证，如，如何能缓解治疗带来的乏力；什么是治疗上肢淋巴水肿（一种腋窝手术及放疗后引起的让患者痛苦和致残的并发症）的最佳方法。

如何解决疗效的不确定性

那么，我们该何去何从？首先，临床医生应获得有关该疗法的当前最佳证据，这些证据来自经验总结和对现有高质量研究的系统评价。经过此步骤后，如果治疗的不确定性仍未能消除，临床医生就需要和患者共同讨论这些不确定性并向他们解释原因。医生须考虑患者的意愿，和患者共同商讨确定最终的治疗选择。共同讨论可能使不确定性的问题增多，我们应该承认这些不确定性并着手解决。只有认识到不确定性的存在，才能慢慢使治疗越来越合理和安全。因此，不确定性是获得进展的先决条件，而非导致失败的因素。目前这种积极应对不确定性的态度已在一些专业指南中得到反映。英国医疗委员会出版的最新指南《医学实践规范》中建议，作为保持和提高医生绩效的一部分，他们“必须帮助解决疗效的不确定性”。²¹为此，临床医生必须同患者共同设计最佳的研究。

将治疗作为科学试验的一部分

当新旧疗法的效果存在重要的不确定性，并且还没有得到正确评估时，我们该怎么办？显而易见的答案是，应该和前文提到的治疗卒中的医生一样，当且仅当在临床试验中，才使用未经彻底检验的疗法，并通过研究设计，搞清楚疗法的利弊。

一位医学伦理学家这样说道：

“如果我们对某些治疗的相对优势不确定，那么我们在将其用于某一适应症时就不能肯定其疗效，也不能随便用于治疗患者。因此，在科学试验完成之前，坚持使用某种治疗方式既不合理也不伦理。由此看来，对于‘该患者的最佳治疗方案是什么？’这个问题，答案就是‘试验’。试验即最佳的治疗方案。这里指的是搞研究吗？是的。但我们这样做完全是因为存在疗效不确定性，并收集相关数据。在这种情况下“随机”选择治疗方案有影响吗？理论上是没有影响的。毕竟，当存在不确定性时，没有比这更好的决策机制了。”²²

将某项治疗作为其科学试验的一部分能起很大作用，因为这有助于改善患者的治疗效果。儿童白血病就是一个非常典型的实例。20 世纪 60 年代之前，几乎每个白血病儿童在确诊后都很快死亡。但是现在白血病的儿童约有 85% 的存活率。能取得如此显著的效果就是因为大多数白血病儿童参与了一项比较现有标准疗法和新疗法的随机试验。²³因而，对于大多数癌症患儿，最佳的治疗选择就是参与这样的医学研究。

患者能否解决不确定性的问题？

“所以，我们要怎么做才能帮助解决治疗中的不确定性呢？……虽然大家普遍认同患者

也是医学研究和卫生保健决策的参与者，但共同讨论治疗的不确定性非常复杂，医生们也会因此感到紧张不安。部分医生只是害怕会使患者更焦虑——这的确是一个值得担心的问题，但不是最重要的。另一部分医生则试图通过平衡以下两种选择以证明其行为的合理性：一是从伦理上讲，有责任告诉患者治疗不确定性的真相；二是从道德上讲，有义务减轻患者的精神压力。患者是否愿意生活在不确定性当中？我们需要找到答案。大概患者的忍受能力要比医生们想象的更强。”

Evans I. More nearly certain. *Journal of the Royal Society of Medicine* 2005;98:195-6.

如果尚未开展这样的试验，医生至少应该以标准的方式记录下使用新的未经验证的治疗的后果，例如，列一个清单，详细记录诊断该疾病时所用到的实验室检查和其他检查，以及评估疗效的方法。研究计划也可以像临床试验一样在数据库里注册（见第 8 章）。这样有助于了解该治疗的信息，有益于那些接受了未经验证的治疗的患者以及其他患者。医疗保健信息系统领域早已投入大量资金，相信通过该系统可以轻松采集有益于患者和其他公众的信息（见第 11 章）。²⁴

如果要更高效地解决疗效的不确定性问题，以上措施必须改进。我们将在本书后半部分谈到这一问题（见第 11 章和 12 章），特别是患者参与。但是还有一个特别的问题，上面已经提到过，这里我们还要强调。当疗效的信息不充分时，应确保临床医生仅在临床试验中使用该药物，进而增加我们对其价值和可能的危害的了解。然而，一些普遍的看法，包括研究监管体系的态度（见第 9 章），实际上阻碍了这种可降低风险的方法的实施。

30 年前，这一问题就困扰着一名英国儿科医生，他发现通过试验的途径，要给他一半的患者使用某种新疗法需要得到批准（给其中的另一半患者使用已有的治疗措施，并作为对照，来评价新措施的疗效），但如果他要把该新疗法作为常规治疗方案用于所有患者，则无需任何同意。²⁵这种不合逻辑的双重标准反复不断出现，并打击了那些希望减少疗效不确定性的临床医生的积极性。最终的结果是，医生们无法从治疗患者的经验中获得信息。就如美国社会学专家 Charles Bosk 所说：“只要我们不期望从经验中获取信息，任何治疗措施都可以很好地开展。”

要清楚地解释疗效的不确定性，从医生方面来讲需要一定的技巧和谦逊的态度。许多医生向即将参与试验的患者解释目前尚不知道哪种治疗最好都觉得很困难。但是，现在公众的态度发生了改变：那些以“上帝”自居的傲慢医生越来越不被患者所崇拜。我们应该着重训练那些敢于承认自己也是凡人，需要患者参与并协助研究以进一步确定疗效的医生。

对于许多临床医生和患者，其主要的障碍是不熟悉有关治疗的科学试验的特点，这个问题我们会在下文解决（见第 6 章）。

要点

- 效果显著的治疗很少见
- 治疗的不确定性普遍存在
- 通常情况下，不同治疗的效果都有微小的差别，合理地辨别这些差别很重要
- 当没有人知道该如何回答疗效中重要的不确定性时，需要采取行动来减少这种不确定性

- 在帮助患者参与到减少疗效不确定性方面，仍有很大的提升空间

第 6 章 验证疗效的科学试验

对于许多读者来说，验证疗效的科学试验可能很陌生，但原理其实并不复杂。日常生活中，很多时候我们对事物的直观把握就取决于它。但是学校对这部分内容的教学并不尽如人意，而且通常毫无必要地将其冠以复杂的文字描述。结果，许多人碰到这个主题便逃避，认为难以理解。我们希望本章以及后两章内容能让你相信，你实际上早已知道了其中关键的原理，从而能够轻松理解为什么它们如此重要。若想了解关于此问题的详细内容可通过浏览 www.testingtreatments.org 或詹姆斯·林德图书馆 (www.jameslindlibrary.org) 获取更多资料。

为什么需要验证疗效的科学试验？

疾病的自愈能力

许多疾病如果不治疗会更严重，而有些即使治疗了仍会变严重。但是，部分疾病不用治自己就会好转，即疾病有“自限性”。一位曾参与一项旨在验证普通感冒治疗效果试验的研究人员说：“感冒经积极治疗，七天内可痊愈，如果不治疗，也会在一周内好。”¹ 开个玩笑话，正所谓“老天治病，大夫赚钱。”当然，治疗也可能会使疾病变得更严重。

正因为人们通常可以不接受治疗而自愈，所以验证疗效时必须考虑疾病的自然病程和不接受治疗情况下的疾病结局。想想自己喉咙痛、胃痉挛或患皮疹的时候，即使不接受常规治疗，疾病也会自愈。但是，如果已经接受了治疗（即使是无效的治疗），你就会认为是治疗的作用让症状消失。简而言之，对疾病的自然病程（如疾病可能自愈）的认识可以避免不必要的治疗和盲目相信未经验证的治疗。

在疾病症状出现和消失的过程中，的确十分难确定是否是治疗发挥了作用。例如，关节炎患者急性发作时最有可能就医，而这种急性发作本身就不会持续很久。不管他们接受的治疗发挥主要作用还是辅助作用，有效还是无效，可能的结局都是疼痛在治疗后得到缓解，因为发作症状消退了。所以即使治疗对症状的缓解没有起到任何作用，医生和患者也会倾向把它归功于所采用的治疗。

误解治疗

“证据证明许多人发现服用烟草可以治愈各种疾病；另一方面，人们服用后也没有受到伤害。这种论断本身就是极大的错误，而且是莫大的荒谬。当某个病人患病到某种程度，他立即服用烟草，之后病自然而然好了，最后他恢复了健康。然后就得出结论，烟草的确创造了奇迹。”

James Stuart, King of Great Britaine, France and Ireland. A counterblaste to tobacco. In: *The workes of the most high and mightie prince, James*. Published by James, Bishop of Winton, and Deane of his Majesties Chappel Royall. London: printed by Robert Barker and John Bill, printers to the Kings most excellent Majestie, 1616: pp 214-222

乐观主义和心怀希望的益处

人们将自己病情好转归功于所接受的治疗，出现这种现象的心理原因现在是比较好理解的。如果一件事情发生在另一件事情之后，我们会倾向于认为第一件事情可能对第二件事情的发生产生了影响。我们也倾向于预测不存在的事物。这种现象体现在很多地方，如预测掷硬币、股票市场的价格和投篮情况等。同样，我们倾向于犯一种称为“证实性偏倚”的错误。我们很容易看到我们希望看到的事物，也就是说“所持的信念决定了所看到的事物”（境由心生）。任何支持自己信念的因素都会促使我们更加相信自己是正确的。相反，我们可能难以发现或难以接受有悖于自己观点的信息，于是会毫不经意地忽视它。

所信即所见

英国医生理查德·亚瑟（Richard Asher）在文章中提到：

“如果你坚信你的治疗是正确的，即使对照试验显示该治疗基本无效，你也会得到比较好的效果，你的患者恢复得比较好，你的收入也比较好。我认为，这说明为什么那些天资不足但轻信有余的大夫们能够获得巨大的成功，也解释了为什么那些功成名就的医生们总是对统计学和对照试验表现出极度反感。”

Asher R. Talking sense (Lettsomian lecture, 16 Feb, 1959). Transactions of the Medical Society of London, vol LXXXV, 1958-59. Reproduced in: Jones, FA, ed. *Richard Asher talking sense*. London: Pitman Medical, 1972.

当然，多数患者和医生希望治疗能起作用。他们会得出某种治疗可以起作用的结论，仅仅是因为他们相信这个治疗应该起作用，这种信念与得出的结论是一致的。他们不会主动寻找或者直接摒除与自己信念相反的信息。这些心理效应也解释了为什么即使治疗不含有效成分（一种通常被称为“安慰剂”的“假冒品”），那些相信治疗会缓解症状的患者，最后病情真的得到了好转。有报道称患者在服用糖片、注射蒸馏水、接受无效的理疗及假手术（只是切开了一个小伤口后又缝合了）后，他们的病情得到了缓解。

以一个对比不同减肥食谱的试验为例。研究人员招募了一个著名电视节目的若干观众，他们都想减肥，研究人员将他们分别分到六个不同的饮食小组中。其中一种饮食——白琳茶——已被公认为是成功的减肥方法。六个节食小组的所有成员的平均体重都减少了，但是一些小组的减肥效果比其他小组更好。然而，公布到电视上的结果透露出“胡萝卜节食法”根本不是什么减肥节食法。因为这个试验之所以采用这一方法，是为其他减肥方法提供一个“参照”——结果显示他们的体重之所以减轻，根本不是什么减肥疗法所致，而是由于其他因素所致的饮食方式的改变。

需要走出主观感觉

假如患者相信治疗确实有效，不是挺好的吗？为什么还要自找麻烦，花费很多来做一些研究来更加正式地评估疗效？为什么还要弄清楚这种对治疗的主观感觉是否能帮助以及如何帮助患者？至少有两个原因来解释。其一，无效的治疗会使我们错过有效的治疗。其二，许多治疗（即使不是绝大多数）有长期的、短期的甚至尚未发现的副作用。如果患者不接受这个无效治疗的话，就能够避免有害的副作用。因而，鉴别无效治疗或者会带来副作用的治疗方式是非常有必要的。研究也揭示了治疗为什么会起作用，同时表明了发现更好更安全治疗的可能性。

验证疗效的研究跟各行各业都有关系，但对于致力于使患者公平享用卫生保健资源的部门和机构，如英国国家卫生署和退伍军人健康管理局等而言，则显得尤为重要。在这种情况下，鉴于卫生保健可获得的资源有限，决策总要考虑采用什么样的治疗可获得更好的效果。如果一些患者接受了无效的治疗，那么可能就意味着剥夺了其他患者接受有效治疗的权利。

当然也不能说患者和医生对治疗的主观感觉和想法一文不值。相反，通常他们的感觉是对某种前景显著的治疗展开正式研究的起点，根据他们的主观感觉开展的研究有时能够同时鉴别该治疗的利和弊。例如，早在二十年前，由于一位孕妇接受了己烯雌酚疗法，她首次提出了这种药物造成她女儿患有罕见的阴道癌（见第2章，第15到16页）。再如，当一位高血压患者提出他接受的新治疗有副作用时，该患者及其医生都没有想到他的意见会引发一种空前畅销的药物问世——西地那非（万艾可）。

因此，个人对疗效的主观感觉也不容忽视。但要得出判断疗效的成熟结论，主观感觉并不是可靠的根据，更不用说要给出推荐意见了。

何为科学试验？

大多数人都知道，仅看媒体对某些医学新进展的表象报道是不对的。但可悲的事实是，就连那些声誉极好的杂志对疗效的报道我们也必须保持谨慎。对疗效的误导和过分宣传太常见了，因此评估疗效的可靠性就显得非常重要。

如果轻信了那些对疗效的表象报道，我们会面临两种风险。第一，我们会错误地认为某种有效的治疗实际上无效甚至有害；第二，我们会错误地认为某种无效甚至有害的治疗实际上有效。验证治疗的科学性旨在通过以下三个方面来获取疗效的可靠信息：1）对照的可比性，以尽可能减少对真实值的影响（偏倚）；2）考虑偶然性；3）评估所有相关且可靠的证据。本章以及后两章会阐述科学试验的这三个主要特征。

对照的可比性

对照是关键

验证疗效的科学性，对照很关键。临床医生和患者有时会在心里比较两种治疗的相对优点。例如，他们会在脑中形成这样一种印象，自己或他人目前接受的治疗与先前的治疗相比，会带来不同的反应。有时这种对照可以更规范些。早在十九世纪，波斯医生拉兹（al-Razi）为知道放血疗法是否对脑膜炎患者有效，便对比了放血与不放血两种治疗的结局。

比较接受不同疗法的各治疗组的情况，就可验证疗法是否有效。如果要保证对照的科学性，就需要使各组具有可比性，即比较组间的唯一不同在于所接受的治疗方法。其实这种观点不是现在才第一次提出。例如，1747年，詹姆斯·林德（James Lind）在索尔兹伯里号（HMS Salisbury）船上对6种坏血病的治疗方法进行对照之前，1）他特意选择了那些处于这种致命性疾病同一阶段的患者；2）保证患者有相同的基础饮食；3）保证患者的生活环境相同（见第1章，第1到3页）。这表明他清楚地意识到，这些非治疗因素而非治疗本身可能影响船员康复的几率。

以下这种情况会造成治疗的试验有偏倚。英国皇家医学院推荐发病初期的海员用硫酸治疗坏血病，而给生命垂危的海员推荐柑橘类的水果。这就造成了硫酸的治疗效果比柑橘类的水果好，实际上是完全错误的。除非对比组的其他各个相关方面都一样，否则将导致很大偏倚。

效果显著的治疗

有时，患者接受某种治疗后发现效果与以前接受过的治疗以及疾病自然病程导致的效果相比，差别十分显著，于是在没有充分验证该疗法的情况下，得出了确信治疗有效的结论（见第 5 章，第 50 到 53 页）。³ 例如，一位肺塌陷（气胸）的患者，给他胸部进针抽出气体，适当缓解了病情，效果非常显著。其他例子，如吗啡缓解疼痛、胰岛素治疗糖尿病昏迷、髋关节置换术治疗骨关节炎等，疗效也很显著。但是，治疗的副作用有时也相当大。一些药物会引发严重的，甚至致死性的过敏反应。如沙利度胺造成了罕见的肢体畸形（见第 1 章，第 4 到 5 页）。

但这些显著的效果，无论是有利还是有害，都不常见。大多数治疗效果平平，不过还是值得获取相关信息。例如，需要进一步确证吗啡的哪种用药方案更加安全有效；转基因胰岛素是否比动物胰岛素效果更好；新上市的人造髋关节比最便宜的一种贵了近 20 倍，考虑到患者的支付能力，要用的话是否值这个额外的费用等。通常这种情况下，我们需要避免没有偏倚的比较，否则会得出错误的结论。

效果一般但却有重要意义的治疗

比较不同时期患同种疾病且情况极其类似的患者接受不同治疗的效果

研究人员通常会比较现在接受某种治疗的患者，与以前患同种疾病的那些情况极其类似并接受其他治疗的患者。如果疗效显著的话，这类对照可以提供可靠的证据，比如某种新治疗可以让患者从普遍病死的疾病中存活下来。但如果不同治疗的效果差别不显著，却仍值得加以研究，那么这种“历史对照”就可能存在问题。尽管研究人员通过统计学调整和分析试图尽量保证比较组间情况相似，这些分析仍旧未考虑到，参与比较的各组患者所有的相关特征没有被详细记录。结果是，这种对照不完全可信。

当不同时期情况相似的患者接受相同治疗时也会出现这种情况。有研究纳入分析了 19 个这样的例子，它们报告了晚期肺癌患者在不同时间接受相同治疗，然后比较每年的病死率。尽管推测它们之间的病死率差别不大，但事实上，差别相当显著：病死率从 24% 到 46% 不等，甚至更高。⁴ 很明显，这种差别并非因治疗不同或是患者已知的情况不一样。病死率不同可能也反映了患者间存在未被发现的差别，或是其他因时间的推移而未被发现的改变（如随着时间的推移，护理或传染的控制会越来越越好），所以在做比较时并没有考虑到这些情况。

比较同一时期情况相似的患者接受不同治疗的效果

情况相似的患者在同一时期接受不同治疗，然后比较治疗的经验和结局，这一方法一直用来评估不同治疗的效果。然而，这也可能造成严重误导。与使用“历史对照”进行比较一样，这种评估方法十分具有挑战性，它需要明确患者在接受不同治疗之前，比较组的患者自

身情况是否足够相似。换言之，是否对照具有可比性。和“历史对照”一样，研究人员采用该种比较疗效的方法，可能会利用统计调整和分析来确保对照的可比性，但前提是，参与比较的患者的相关特征要记录和考虑进去。然而这些条件很难满足，因而我们应该谨慎对待此类分析结果。相信它们可能会导致巨大悲剧。

在一个关于激素替代疗法（HRT）的典型例子中，将绝经期间及绝经后接受了激素替代疗法的女性与情况相似且未接受该疗法的女性进行了比较。比较中发现，激素替代疗法能够减少心脏病和卒中发生的风险，如果该结果属实，那真是个好消息，很可惜它不是。随后研究人员通过设计具有可比性的对照组，以保证治疗前比较组间患者的情况足够相似，结果显示激素替代疗法恰恰带来相反的作用，它增加了患心脏病和卒中的风险（见第2章，第16到18页）。在这个案例中，前后两次心脏病和卒中发生风险之所以形成反差，是由于前一次接受了激素替代疗法的女性普遍比未接受的女性更健康，而不是因为激素替代疗法本身。由此可见，没有可比性的对照研究可能危害数以万计的人。

HRT的例子表明：保证组间可比的最好办法是在治疗之前进行对照分组。试验组和对照组的患者，不仅在已知与可测量的因素上相似，例如年龄和病情程度，而且要在无法测量却可能影响疾病恢复的因素上相似，例如饮食、职业、其他社会因素、对疾病的担忧或所推荐的疗法。在患者开始治疗后再分组，那么就很难保证，事实上根本无法保证这种相似度了。

但接下来又有个十分关键的问题：不同的结果反映的是不同疗效的差别还是比较组间患者本身的差别？

对不同治疗进行无偏倚地预先分组

1854年，军事孤儿院的军医托马斯·格雷厄姆·贝尔福（Thomas Graham Balfour）找到了设立比较组且保证可比性的方法。贝尔福想证实颠茄是否像一些人说的那样能治疗儿童猩红热。所以，“为了避免选择导致的误差”，他将孩子们交替分配到给药组和不给药组。这种交替分配或其他一些无偏倚的分组方式是验证疗效科学性的关键特征。这样的方式增加了比较组间的相似性，不仅保证已知且可测量的重要因素相似，而且保证了不可测但影响康复的因素相似，因此也避免了数据调整。

为了无偏倚地分组，科学试验的设计者要保证临床医生和患者都不知道或预测不到下一步分组的情况。如果他们知道就会有意或无意地被诱导，选择特定的治疗措施。例如，如果某位医生知道下一位要参加临床试验的患者分配到安慰剂（一种无效治疗）组，他或她可能不会让病情较重的患者参加，而会让病情较轻的患者参加。所以即使制定了无偏倚的分组方案，如果未对拥有决定患者是否参加科学试验权利的人隐瞒这种分组方案，那么就无法保证无偏倚地分配。而如果把分组的情况进行隐瞒的话，就没有人知道下一个患者将被分配到哪个组，也不会有偏离无偏倚分配原则的倾向。

产生分配序列号比简单的交替分配更不可预测，这有助于实施对分组方法的隐瞒，如以随机数字表为基础的分组，以及对该分组情况的隐瞒。隐瞒分组有多种方法，如可以用电话或电脑对确定符合纳入标准的患者进行远程随机分组；另一种方法是使用一系列有编号的信封，每个信封里包含一种分配方案，符合标准的患者选择并打开信封可以知道分组情况。为了保证这种方法有效，信封必须是不透明的，以防医生通过对光查看信封内的方案来“作弊”。

这个方法目前被认为是科学试验的一个重要特征。用随机数字法分配治疗方案的研究被称为“随机试验”。（见第3章，第26页的方框）

通过电话随机法隐瞒分配方案



通过电话随机法隐瞒分配方案

疗效比较中采用多种无偏倚（随机）分组方法

比较疗效可采用多种随机分组方法。例如，随机分组可用来比较在不同时间随机给同一个患者不同治疗，即所谓的“随机交叉试验”。基于此，如果要评估一种吸入性药物是否能帮助患者治疗顽固性干咳，可以设计一个持续几个月的研究。随机选取其中的几个星期，给患者使用包含药物的吸入器；而剩下的几个星期里，给患者使用相同的吸入器，只是里面不放药物。如果这种方式能够实现，那么得到的研究结果还是比较满意的。但是在很多现实情况下，这样的交叉研究是不可能实现的。例如，不同的外科手术就不能以这种方式比较，另外对某些“一过性”的疾病，急性疾病，如车祸后的大出血，也是不能采用的。

随机分组也能够用来对同一个患者不同部位给予不同治疗的比较。基于此，对于某种皮肤病，如湿疹或者银屑病，随机选取皮肤的患处涂含有药物成分的软膏，而其他地方涂不含活性成分的软膏。又如，在治疗双眼疾病时，随机选择某只眼睛进行治疗，而另一只不接受治疗的眼睛作为对照。

还有一种随机分组的方法可以用来比较不同人群或组群接受不同治疗的效果——也就是说，针对所有进入每一个初级保健诊所或者医院的人。这种对比被称为“群（或组）随机试验”。比如，为了评估墨西哥全民医疗保险计划，研究人员在卫生保健服务范围中匹配了74对地区——这些群组包含了美国7个州118,000个家庭。随机选取每组配对中的一组施行医疗保险计划。

	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY
May WEEK 1	A	B	A	B	B
May WEEK 2	A	A	B	B	A
May WEEK 3	B	A	B	A	A
May WEEK 4	B	A	B	B	A
May WEEK 5	A	B	A	B	B

将某人的时间随机化



将某人的眼睛随机化



将某区域的社区随机化



将某组的个体随机化

随机分组可能的不同分组方法

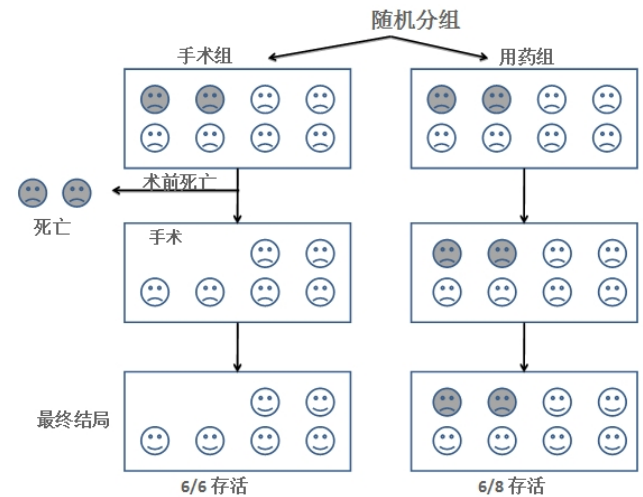
但现在随机分组最常见作用是决定如何将不同患者分配到不同治疗组。

疗效比较需要随访每位患者

在尽量确保比较组间具有可比性后，要注意避免因研究过程中忽视了某些患者而引入偏倚，这非常重要。不论他们接受的是哪种治疗（如果有的话），都要尽可能随访所有参与分配的人，同时要将其纳入所在组的主要结果分析中。这就是“意向性治疗”分析。否则，就不能保证科学对比了。

乍一看，将患者分组后，对照组的患者得不到治疗似乎不合逻辑。但如果不这样做，那么就会使试验不科学，其结果会有误导。比如，对于脑血管部分阻塞且出现阵发性眩晕的患者，他们卒中的可能性极大。研究人员实施一项试验来确认血管疏通手术是否能降低患者后续卒中的风险。他们恰当地比较了所有治疗组（无论其是否能在手术中存活）和非治疗组的患者。如果研究人员仅仅记录那些在手术后存活患者的卒中发生次数，在其他条件保证科学的情况下，会发现存活的患者卒中发生率更低，但他们已然忽略了一个重要的事实——手术本身也会引起卒中和死亡。这就会使对手术效果的试验不科学，因而手术的风险也应该被纳入评估。

图中所示的手术和药物治疗结果实际上相同。但如果两个被分配进行手术的患者在术前死亡并被排除到考虑范围外，这两组的比较就出现了偏倚，而这将会表明手术治疗效果更好，可实际并非如此。



所有参与随机分配的患者都需要纳入最终结局的原因（意向性分析）

处理按分配方案治疗的患者与实际接受治疗患者的差异

基于目前在本章中讨论的所有原因，你会意识到需要仔细设计治疗措施科学试验。呈现研究计划的文件称为研究方案。然而待实施的计划设计得再好，也可能并不会如预期所料——有时患者实际上所接受的治疗与计划当中的是不同的。例如，患者可能不会如预期的那样接受治疗，或者受限于物资和人员而无法开展某种治疗。如果发现了这种现实与计划的差异，我们就得仔细考虑并解决所带来的影响。

20世纪70年代和80年代，对儿童最常见的一类白血病——急性淋巴细胞性白血病的治疗取得了相当大的进展。然而，令人费解的是，从表面上看，美国儿童和英国儿童确实接受相同的药物治疗，但实际上美国儿童比英国儿童恢复得一些。在随访美国加利福尼亚的一个儿童癌症中心时，一位敏锐的英国统计学家注意到，美国白血病儿童接受的化疗远比英国儿童接受的更“激进”。这种治疗会出现令人难受的副作用（恶心、感染、贫血、脱发等），而这些副作用很难处理，所以英国医生和护士没有像美国同行那样“激进”，而是倾向于减少或暂停计划的治疗。这种“温和方法”会降低治疗的效果，这可能是英国和美国治疗效果不同的原因之一。

帮助人们按原计划进行治疗

不同治疗比较时，患者计划的和实际接受的治疗出现差别还会有其他方面的原因，这将会使对验证疗效的结果的评估变得复杂。我们不应该拒绝参与研究的人接受医学上必要的治疗。当一个未经证实但有希望带来有益效果的新治疗纳入科学试验中，人们应该确信参与研究的患者接受的都是确定有效的治疗。

如果人们知道患者接受的是什么治疗的话，就会出现几种可能的偏倚。一方面，患者和医生们可能会觉得接受“新”疗法的人非常幸运，这可能导致他们无意识地夸大该种治疗的益处。另一方面，患者和医生可能会觉得“旧”的疗法很难治好，这种失望可能导致他们低估任何积极的效果。知道谁会接受何种治疗，也会使医生给予接受“旧”疗法患者一些额外

的治疗和护理，以作为补偿，因为他们并没有被分配未经证实的新疗法组。这样当一组患者接受额外的治疗而另一组没有，则会使新疗法的评估变复杂，而且会使要比较的各组间不具有可比性，从而误导结果。一种减少计划治疗和实际治疗间误差的途径就是尽量保证新旧治疗组间各个方面可比，比如药物的颜色、味道和气味。

比较有希望带来有益效果的治疗与无效治疗（虚假的治疗或安慰剂）也需要采用上述方法，即虚假治疗在色、香、味等方面都要与“真正”治疗措施相似。这就是所谓的“盲法”或“伪装”。如果研究中能够实施“盲法”（实际中许多情况下不能实现），那么两个比较组的患者仅在某一方面表现不同，即新旧疗法的不同。同样，医生基本上不可能判断出他们关心的病人接受的是新疗法还是无效治疗。如果医生和患者都不知道接受的是哪种治疗，这种试验叫做“双盲”。于是乎，两个比较组的患者就会更加积极地依照所分配的治疗，临床医生也会更有可能对所有患者一视同仁。

验证治疗结局的科学性

虽然疗效比较中使用盲法的原因之一在于，能帮助患者和医生坚持被分配的治疗，然而一个更加公认的原因是，“盲法”能减少治疗结局评估出现的偏倚。

盲法成为上述原因之一还有一段有趣的历史。在十八世纪，法国国王路易斯十六号召对安东·梅斯梅尔（Anton Mesmer）的观点即“动物磁性”（又称催眠术）能带来有益效果进行研究调查。国王想知道这是由于“真正的力量”还是“思维幻想”形成的结果。在疗效验证的试验中，参与者蒙上眼睛并被告知是否接受催眠，事实上有时候他们被告知的和实际不符。结果显示，当被告知自己接受了催眠术，他们才会说体会到了“治疗”的效果。

对于某些治疗结局如存活率，因为某人是否死亡这个结局是确定的，评估结局不可能存在偏倚。但由于结局指标应该而且常常会涉及到某人对相关症状如病痛和焦虑的个人体验，因而大多数治疗结局带有一定的主观性。患者出于个人因素，可能会更加倾向于某种治疗。比如，当认为某种治疗对自己有益时，他们会对可能出现的有益症状更加敏感，反之亦然，人们会把有害的副作用归咎于他们担忧或不相信的治疗。

通常盲法是科学试验的必要特征。这意味比较组间的治疗必须以同样的方式出现。以多发性硬化症治疗的疗效验证为例，不知道患者是接受新药还是无效治疗（即采用了盲法）的医生与知道治疗组具体分配情况（即没采用盲法）的医生都检查所有的患者，结果显示，施盲的医生认为新疗法效果不大，而未施盲的医生认为新疗法有效。⁸这种差别的出现意味着新疗法是无效的，而且若知道治疗的具体分配会使未施盲的医生“看到他们认为的和希望看到的结局”。总而言之，评估治疗结局时主观因素越大，越有必要通过盲法来保证验证疗效试验的科学性。

有时可以对患者实施盲法，让他们不知道自己是否接受了真正的手术治疗。有一项研究针对膝关节骨关节炎患者，结果表明，清洗骨关节的手术方法与只是在麻醉情况下给膝关节皮肤做了一个小切口而假装清洗的方法相比，并没有明显的效果差异。⁹

但要对患者和医生都实施盲法，让他们都不知道治疗具体的分配情况，通常是不可能实

现的，如要比较手术和药物治疗的效果时，或者药物有典型的副作用时。某些结局虽然可能会有偏倚（如判断死亡原因或者X光结果），但可以让不知道治疗分配情况的人来独立评估治疗结局，从而避免偏倚产生。

发现并调查治疗预期外的副作用

发现治疗预期外的副作用

治疗预期外的结果无论好坏，通常首先由医生和患者察觉到。因为一项治疗要得到上市的许可，需要对治疗进行验证，这包括仅对几百人或几千人进行几个月的治疗，但这一阶段只能发现相对短期和频繁的副作用。罕见的副作用和那些需要经过一段时间发展的副作用很难在这一时期被察觉到，除非该治疗得到广泛且长时间的推广，而且覆盖的患者比上市前研究多得多。

越来越多的国家包括英国、荷兰、瑞典、丹麦和美国为临床医生和患者报告药物疑似不良反应提供了有利条件，之后会正式研究该不良反应。¹¹ 尽管这种报告的方式几乎没有成功发现过药物重要不良反应，但还是存在一些比较典型的例子。如 2003 年降胆固醇药物罗素伐他汀在英国上市后不久，有报告说发现了一种严重、罕见而且预期外的肌肉副作用，叫做横纹肌溶解。这种副作用使肌肉迅速分解，而且降解产物会造成严重的肾损害。进一步的调查发现大多数出现这种并发症的患者均服用了大剂量的罗素伐他汀。

调查对副作用的直觉

对副作用的直觉常常不准。那么应该如何对这些直觉进行调查，来证实其是否属实？这就需要对其进行试验，来证实或消除这些副作用，需遵守的原则与验证疗效的原则一样。也就是说要避免有偏倚的对照，确保研究对象可比性，并且要基于大量的研究数据。

黄卡计划

沙利度胺悲剧强调了上市后药物监测的重要性，之后，英国于 1964 年发起了黄卡计划。提交的报告由英国药品和保健品管理局（MHRA）分析。每年 MHRA 会收到 2000 多例关于可能的副作用的报告。起初，只有医生才有资格递交报告，后来鼓励护士、药师、法医、牙医、放射技师和验光师都递交报告卡。从 2005 年开始，鼓励患者和保健者报告疑似的不良反应。用户可以直接在线填写报告表或者通过电话在 www.yellowcard.gov.uk 网站上提交报告。

一位患者这样概括自己的经验：“能够通过黄卡计划报告副作用让自己有更大自主权。这意味着你可以直接报告而不需等待繁忙的专家来报告……这体现了以患者为中心的治疗理念，是患者参与的重大突破，标志着患者参与治疗之路的开端和人们态度的重大转变。”
Bowser A. A patient's view of the Yellow Card Scheme. In: *Medicines & Medical Devices Regulation: what you need to know*. London: MHRA, 2008. Available at www.mhra.gov.uk

与期望的治疗结果一样，预期外显著的结果比不显著的结果容易辨认和证实。如果预期外的疑似结果很不寻常，但在接受该治疗后经常出现，通常提醒医师和患者应该是什么地方

出错了。19 世纪末，瑞士外科医生柯赫尔（Theodor Kocher）从全科医生那儿了解到，早在几年前当自己切除了一位甲状腺肿女孩的甲状腺后，这个女孩变得迟钝而且嗜睡。他对这个案例和以前接受他治疗的甲状腺肿大患者进行深入调查后，发现对肿大甲状腺的完全切除术会并发呆小症和粘液性水肿，现在我们知道这是一种由于缺乏甲状腺激素导致的罕见且严重的疾病。¹²另外，服用沙利度胺（见第 1 章，第 4 到 5 页）产生的预期外的疑似效果得到证实，因为妊娠期服用该药与婴儿出生后四肢残缺的关系是显著的。这种婴儿畸形以前几乎未曾听说过。

用于评估各种治疗相对优势的随机对照试验也会发现不怎么明显的预期外副作用。两种预防新生儿感染的抗生素疗法的随机对照试验表明，其中一种药物干扰了肝脏代谢产物胆红素的代谢过程。这种代谢产物在血液中积累对接受抗生素治疗的婴儿的大脑有损害。¹³

有时对以前的随机试验进一步分析有助于辨别这种不显著的不良反应。对妊娠期妇女注射己烯雌酚后发现，这种药物可导致她们当中某些患者的女儿患癌症，此后就有人推测这种药物可能还有其他不良反应。通过联系参与随机对照试验的女性的子女，可以发现这些不良反应。随访研究显示，其子女会出现生殖器异常和不孕不育。最近一种治疗关节炎的新药罗非考昔（万络）被怀疑会引起心脏病发作，之后对相关随机试验结果的更详细研究表明，该药确实会引起这种不良反应（见第 1 章，第 5 到 7 页）。¹⁴

当研究预期外的不良反应时，要保证组间的可比性，随访参与随机试验的患者是一种非常令人满意的方法。然而，除非预先有相关规定，这并不是最佳的方法。如果能够按照规定收集随机试验参与者的详细联系方式，研究治疗措施可能的不良反应并不是很困难。因为我们可以再次联系他们并获得其更多关于患者健康的信息。

如果疑似的不良反应涉及的健康问题与采用的治疗完全不同，对其进行研究就容易多了。¹⁵例如，斯波克（Spock）博士建议婴儿应该趴着睡，他这里是针对所有的婴儿，而不是那些被认为风险高于平均水平的婴儿（见第 2 章，第 13 到 14 页）。推荐的措施（“让婴儿趴着睡”）与疑似的后果（婴儿猝死）之间缺乏联系，于是，当观察到这两者间存在相关性，就更能得出治疗 and 疑似后果间存在因果关系的结论。

相反，要研究治疗抑郁症的处方药会导致自杀心理这一预期外的副作用，则极具挑战性。如果没有做该药物与其他抑郁症疗法的随机对照试验，就很难断定服用与未服用该药的患者情况是否足够相似，也就很难保证对照的可靠性。

要点

- 验证疗效的科学性必不可少，否则我们会将无效的治疗认为有效，反之亦然
- 对照是所有治疗性科学试验的根本
- 比较不同治疗时（或比较治疗与不治疗时），组间可比性至关重要
- 评估治疗结局必须尽量减少偏倚

第 7 章 考虑偶然因素

偶然因素的影响和大数定律

关于疗效的可信证据有赖于避免偏倚的产生（并处理好那些目前尚不能避免的偏倚）。除非科学试验能达到上述要求，否则即便再大幅度地调整研究结果都将于事无补，得不到解决的问题依然会存在，其后果不仅危险，有时甚至还致命（见第 1、2 章）。但即使我们能成功地采取措施减小偏倚，人们仍可能会被偶然因素所蒙蔽。

众所周知，如果我们反复掷硬币，很少会见到连续五次或五次以上出现正面或反面的情况；且随着次数增多，最后正反面出现的次数极有可能十分相近。

两种治疗措施比较结果的差异可能仅仅是由偶然因素造成的。例如，患者接受 A 治疗后病死率为 40%，类似患者接受 B 治疗后病死率为 60%。假设分别有 10 位患者接受两种治疗，表 1 展示了你所预期的结果。接受两种治疗后死亡人数的差异被表示为风险率，这个样本的风险率是 0.67。

基于这样的小样本，我们能否给出方案 A 比方案 B 好的合理结论？很可能不行。因为其中一组的某些受试者可能因为偶然性而比另一组恢复得更好。如果在其他的小组间开展同样的试验，小组间死亡人数之比可能会颠倒（6:4），或相等（5:5），亦或出现其他的比例——这完全是偶然的。

	治疗方案 A	治疗方案 B	风险率 (A:B=)
死亡人数	4	6	(4:6=) 0.67
总人数	10	10	

表 1：通过这项小样本研究是否能可靠地估计治疗方案 A 与 B 的差异？

但如果分别有 100 名患者接受同样的两种治疗后依旧出现相同的比例（4:6），你会作何感想（表 2）？虽然反映差异的指标（风险率）与表 1 中的是相同的（0.67），但 40:60 比 4:6 的死亡人数更具有说服力，偶然因素作祟的可能性更低。所以，在比较两种治疗时，避免被偶然因素误导的方法就是：基于特定患者数（死亡、恶化、好转或是不变）足够多的样本得出结论。这有时也被称为大数定律。

	治疗方案 A	治疗方案 B	风险率 (A:B=)
死亡人数	40	60	(40:60=) 0.67
总人数	100	100	

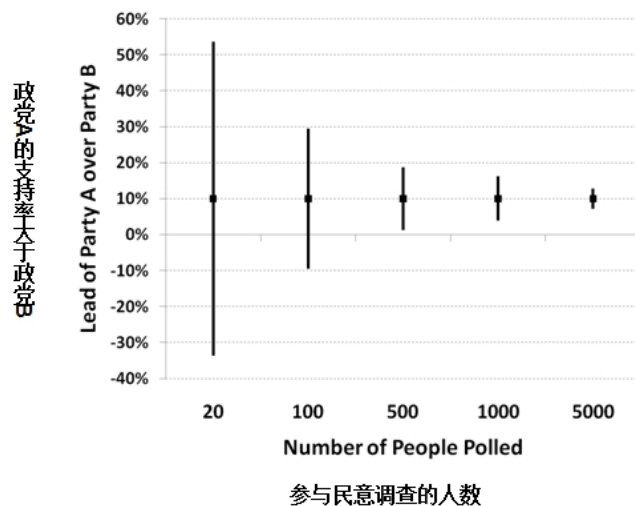
表 2：通过这项中等样本量的研究是否能可靠地估计治疗方案 A 与 B 的差异？

评价偶然因素在科学试验中可能发挥的作用

在解读科学试验中有关疗效比较的结果时，偶然因素会让我们犯两种错误：当实际疗效无差异时我们错误地认为有差异，或是实际疗效有差异而我们没有发现。被观察的特定治疗效果其样本量越大，上述错误发生的几率就越小。

考虑到疗效比较不可能将以往接受过或是未来将要接受治疗的所有个体包括在内，所以完全揭示两种治疗方案“真实的差异”是不可能的。但是，科学研究必须对这种真实的差异做出最佳的推测。

“可信区间”（Confidence Interval, CI）常用来表示估算出的差异的可信程度。它给出了真实差异可能存在的范围。即便不提可信区间这称谓，相信大多数人也知道它是什么。比如，在选举筹备阶段，一项民意调查显示 A 政党领先 B 政党 10% 的支持率；但如果我们留意媒体报道的话，我们会发现这两个政党支持率的差距最小可能 5%，而最大能达到 15%。上述的“可信区间”表明两政党的真实差异可能会落在 5%~15% 之间的某一点。随着越来越多的受调查者参加，结果的不确定性也越来越小，而两党支持率差异的估计值所对应的可信区间也就越窄。



表示政党 A 和 B 支持率差异的 95% 可信区间 (CI) 随着参加民意调查人数增加而变窄

就像我们能通过两个政党支持者的比例估算出两党支持率差异，进而得出该差异的不确定程度一样，我们同样也可以通过患者接受两种治疗后其病情缓解或恶化的人数之比估算出疗效差异，进而得出该差值的不确定程度。再举个例子，对比两种方案治疗心脏病的效果，观察的治疗结果数量越多，评价疗效差异的可信区间就会越窄。而可信区间“越窄越好”。通常会有一项指标来说明我们有多大把握认为真实值会落在可信区间的估计范围内。例如，95% 的可信区间意味着无论真实值是什么，我们都有 95% 的把握认为真实值会落在可信区间范围内。这也意味着，每 100 个“真实”值里可能会有 5 个 (5%) 落在范围外。

疗效比较的“显著性差异”意味着什么？

这是一个有趣的问题，因为“显著性差异”可以包含好几层意思。首先，它可以是对患者有重大意义的差异。然而，研究报告的作者们提到的“显著性差异”则通常是指“统计学差异”。但“统计学差异显著”在我们的实际日常生活中不一定会“显著”。即使“统计学差异”是由非偶然因素所致的疗效差异，它也可能几乎或完全没有实际意义。

以一个随机对照试验的系统评价为例：比较每天摄入一片阿司匹林与不摄入阿司匹林对数以万计的健康男性的影响。这项研究发现，摄入阿司匹林的人群心脏病发病率下降，这种差异是有“统计学意义”的，也就是说它不太可能由偶然因素引起。但这并不意味着它一定具有实际意义。如果一名健康男性心脏病的发病几率已经很低，此时利用药物降低发病率是不合理的，特别是阿司匹林会产生多种副作用，就拿出血来说，有时会致命。根据系统评价所提供的证据，如果 1000 名男性持续十年每天服用一片阿司匹林，其中有 5 个人在这期间可以避免心脏病发作，但 1000 人中有 3 人将会严重出血。

“统计学意义”意味着什么？

“老实说，这是个有趣的概念。它能告诉我们药物和安慰剂所造成的差异或不同的寿命预期所造成的差异是否有可能仅仅是偶然性导致。它也意味着，那些大到我们能观察出的差异是不太可能仅仅由偶然因素引起。

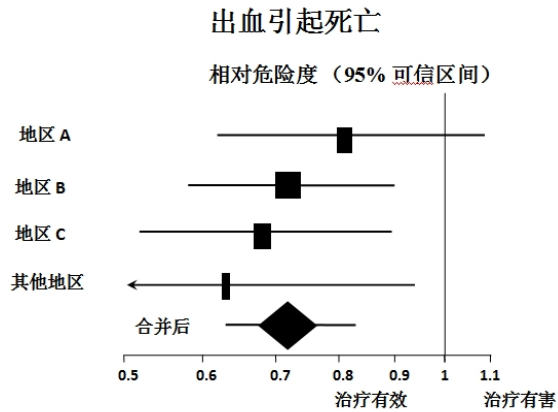
统计学家采用标准水平来描述‘不可能’。他们说的‘有意义’通常是指 5% 的水平（有时写作 $P=0.05$ ）。我们说差异有“显著性”，意即如果所有事件都是偶然因素造成的，那么这种情况 20 次里最多发生 1 次。”

Spiegelhalter D, 引自: *Making Sense of Statistics*. 2010.
www.senseaboutscience.org

在科学的疗效试验中获取足够大的样本量

有时候对疗效的评价可能仅从一到两个中心的研究中就能获得足够的样本量。但在研究治疗对类似死亡等罕见结局的效果时，就需要多中心、甚至多个国家的患者参与研究，以获得可靠的证据。例如，一个由 13 个国家 10000 名患者参与的试验显示，类固醇药物的使用会对严重脑损伤的患者起到致死性作用，而这种治疗已经被沿用了 30 多年。² 这个研究小组所做的另一项由 40 个国家 20000 名患者参与的试验显示，一种叫做氨甲环酸的廉价药物可以减少因损伤后出血而引起的死亡。³ 这些研究通过精心的设计来降低偶然因素所引起的偏倚和不确定性，它们是典型的科学试验，它们还为全球卫生保健提供了高度相关的高质量证据。事实上，在一项由英国医学杂志（BMJ）发起的民意调查中，上述氨甲环酸的随机对照试验被评选为 2010 年最重要的研究。

尽可能将疗效的评估建立在充分信息基础之上是非常重要的，下面的图表是通过上述获奖团队友情提供的数据，来说明如何降低由偶然因素导致的风险。图表底部的菱形方块代表了合并全部氨甲环酸试验后的结果。它表明这种药将出血引起的死亡数降低了近 30%（风险率刚刚超过 0.7）。合并后的结果提供了对这种药物效果最可靠的评估，尽管“A 地区”各中心的评估显示出这种药物的效果不太显著（无统计学意义，有可能低估药物的真实疗效），但“其他地区”里各中心的结果显示出相当显著的效果（有可能高估了疗效）。



各地区氨甲环酸对严重出血引起创伤患者死亡的情况及总体情况

某种程度上，通过合并多个国家、多个中心的试验数据，可以减小偶然误差。这些相似但又相互独立的试验其结果有时可通过统计学方法进行合并——这一过程被称为“Meta 分析”（见第 8 章）。尽管 Meta 分析的方法学已经被统计学家发展了多年，但直到 20 世纪 70 年代它才真正得以广泛应用，最开始是美国的社会学家，紧接着是医学研究者。截至 20 世纪末，Meta 分析已作为科学试验重要的元素被广泛接受。

例如，在五个不同国家中，有五项目被不同组织实施及不同资金支持的研究都在解决一个 60 年来未能解答的老问题：对于早产儿，“什么样的血氧水平最有可能保证他们幸免于重大残疾？”如果血氧水平太高，婴儿可能会失明；如果太低，他们可能会面临死亡或脑瘫。对于这些体弱的婴儿，不同的血氧含量所造成的差异很可能不显著，所以这就更需要足够的样本量来探究差异。因此这五个研究小组同意将自己负责的研究结果进行合并，这样所提供的结果将比任何一项单独研究所提供的都要可靠。

要点

- 在评价现有证据的可信度（基于证据质量和数量）时，必须考虑偶然因素的影响

第 8 章 评价所有相关的可靠证据

一项研究是否就足够了？

答案简单明了——“绝对不够”。单个对照研究几乎不可能得出充分可靠的证据，更不能用于支持治疗的决策。不过，这种情况有时也会发生。在这些罕见的单个研究中，有一项说明了心脏病发作期间服用阿司匹林能降低早期死亡的风险；¹另一项研究明确指出急性创伤性脑损伤患者服用类固醇会产生致死性后果（见下文和第 7 章，第 89-90 页）；还有一项研究证实咖啡因是现有唯一可预防早产儿脑瘫的药物（见第 5 章，第 57-58 页）。而通常情况下，单个研究只不过是解决一个相同或相似的问题时多项对照研究中的一个。所以单个研究得出的证据应和其他类似研究中的证据一起被评估。

20 世纪 60 年代，治疗的科学试验的先驱之一——英国统计学家 Austin Bradford Hill 提出了研究报告应该回答的四个问题：

- 为什么开展研究？
- 研究的内容是什么？
- 有何发现？
- 研究结果究竟意味着什么？

为什么开展研究？

“就临床研究的科学性和伦理性而言，解决需要被回答的问题是最根本的一条原则，而且在设计这些研究时还须确保其答案是有意义的。要满足以上任一要求的前提是确保最相关的研究已得到了恰当的开展……对已有证据不全面的了解将会违背与研究受试者的伦理契约，因为所提供的信息对受试者来说会既必要又有用。”

Robinson KA, Goodman SN. A systematic examination of the citation of prior research in reports of randomized, controlled trials. *Annals of Internal Medicine* 2011;154:50-55.

今天看来，这些关键问题依然很重要，却常常未能得到充分解决，甚至完全被忽视。而最后一个问题——研究结果究竟意味着什么？——其答案尤为重要，因为这可能影响到治疗决策和未来的研究。

以“对可能早产的妇女使用短程、经济的类固醇药物治疗”为例。1972 年该治疗的科学试验首次被报告：母亲接受类固醇治疗可降低早产儿死亡的可能性。十年后出现了更多试验，这些研究规模小且单个研究的结果让人眼花缭乱，因为这其中没有任何一个研究对以往类似的研究进行系统的总结。如果他们做了这一工作，就会发现已存在强有力的证据证明该药物的疗效。而事实上直到 1989 年该工作完成之后，大量的产科医生、助产士、儿科医生以及新生儿护士才意识到该治疗如此有效。可惜的是，之前已导致上万的早产儿白白夭折。

2

为解答“研究结果意味着什么？”这一问题，疗效比较的特定的科学试验证据必须同其他类似研究的证据一起解读。在报告新试验结果时，若没有系统地评价和参考其他相关证据，就会耽误对有益与有害的治疗的认知，并导致更多不必要的研究。

综合研究信息

一百多年前，英国科学促进协会的主席 Rayleigh 公爵就认为有必要在参考其他相关证据后对新的研究作出结论：

“众所周知，如果科学仅是由已有事实不断累积而成，那么它很快就会陷于停滞并被自身压垮……因此，获取新知识、消化吸收旧知识，这两个过程是同时进行的。因为两者都必不可少，所以我们需要抽出一些时间讨论它们的相对重要性……这些工作最需要的是发现问题与解释问题双管齐下，在呈现新事实的同时也应指出它和已有事实的关系，但恐怕人们通常都做不到这一点。”

Rayleigh, Lord. In: *Report of the fifty-fourth meeting of the British Association for the Advancement of Science; held at Montreal in August and September 1884.* London: John Murray, 1884: pp3-23.

对所有相关的可靠证据进行系统评价

我们应该将与某个特定研究的结果相关的所有可靠证据进行评价。这句话说起来很容易，但从许多方面来看都是一种挑战。评价非常重要，因为人们能够信赖它，这就要求必须系统地进行评价，否则就会误导大众。

系统评价的重要性

“系统评价和 Meta 分析对卫生保健日趋重要。临床医生借此跟上专业领域的步伐，同时它们也是制定临床实践指南的基石。资助[基金]机构可能会要求一篇系统评价来确保后期研究的合理性，同时一些医疗卫生期刊也在朝这个方向靠拢。和所有的研究一样，系统评价的价值取决于研究的内容、结果以及透明的报告。但系统评价报告的质量也像其他发表物一样良莠不齐，影响了读者对它们优劣的判断。”

Moher D, Liberati A, Tetzlaff, Altman DG. The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement (www.equator-network.org), 2009.

那些看似解决相同疗效问题的系统评价可能会得出不同的结论。有时这是因为所要解决的问题存在细微差异，或是研究者所采用的方法不同；有时也会因为研究者对结论做了“调整”。所以，找到解决我们关心的治疗问题的系统评价，判断哪种评价最有可能成功地降低偏倚并减少偶然因素的影响，以及哪种评价的结论完全忠于证据都是非常重要的。

减少系统评价的偏倚

偏倚不仅会歪曲单个试验的结果，导致错误的结论，它还会歪曲对证据的评价。例如，研究者可能只“挑选”那些他们认为和自己所期望的疗效结论相符的研究。

为避免这些问题，系统评价也应该像单个研究那样，将计划反映在研究方案里。研究方案需要明确呈现研究者将要采取何种措施来减少偏倚和评价过程中偶然因素的影响。这其中包括：阐明这项评价将要解决关于某一治疗方案的何种问题；评价过程中纳入研究的标准；

采取何种方法筛选出可能符合要求的研究；在纳入研究和分析数据的过程中采取何种步骤将偏倚最小化。

纳入系统评价的所有相关证据

纳入系统评价的所有相关证据——即便不考虑相关报告的语言和形式——还是面临着巨大的挑战，更不用说一些相关证据甚至从未被公开发表过。报告不全主要归咎于研究者因为对结果不满意，所以未能完成研究报告，或是将本应发表的报告没有投稿。而且药厂也会打压对自身产品不利的研究。期刊杂志在退稿时也存在偏倚，因为他们认为有些研究结果不足以“吸引眼球”。

不充分报告研究存在的偏倚是一种不科学和不伦理的行为，当前这一严重问题受到了广泛认同。特别是当人们试图选择治疗措施时会被误导，因为那些得出了“令人失望”或“阴性”结果的研究很少被报告，相反那些得出令人兴奋结果的研究很有可能被“过度报告”。

报告不充分的程度令人惊讶：至少有一半的临床试验从未得到完整报告。这种报告不充分的研究会造成偏倚，并且无论大规模还是小规模临床试验，都存在这种现象。处理这一问题的方法之一就是要求临床试验从一开始就公布研究计划，并鼓励研究者发表研究方案。

3

营销医学

“药厂的内部文件表明公众可得的证据可能并非其产品相应数据的真实体现。药厂及其相关医药传播公司的宗旨是医学文献的发表主要为市场利益服务。打压和篡改阴性数据及找人代笔[见第 10 章，124-125 页]已成为（药厂）控制医学出版为产品销售服务的工具，同时，散布疾病谣言和医生市场的细分也被用于使利润最大化。我们应认清尽管循证医学是理想境界，但营销医学才是当下真实的情况。”

Spielmans GI, Parry PI. *From Evidence-based Medicine to Marketing-based Medicine: Evidence from Internal Industry Documents. Journal of Bioethical Inquiry* 2010;7(1):13-29. Available online:

不充分地报告研究的偏倚甚至可能引起致命性后果。为了彰显诚信，英国一些研究者决定于 1993 年报告一项 13 年前进行的临床试验的结果。该试验研究了一种缓解心脏病患者心律失常的新药：九名患者用药后死亡，而对照组仅有一名患者死亡。他们写道“1980 年进行研究的时候，我们认为药物组病死率升高是偶然因素引起的……而该药物[劳卡胺]的研发由于商业因素被叫停，因此这一研究从未被发表；而现在它则成了‘发表偏倚’很好的例子。此处描述的结果……也许能对潜在的问题发出预警”。⁴而他们所提到的“潜在的问题”指的是一种与劳卡胺相似的药物，在其使用的高峰阶段，仅在美国就引发了成千上万的早亡（见第 2 章，第 14-15 页）。⁵

减少系统评价的偶然因素

在第 7 章（第 91 页）我们阐述了如何通过合并相似的单个研究数据（即 Meta 分析）来降低偶然因素的影响。所用的案例是在五个不同国家里分别开展并受到资助的五项研究，它们共同解决了一个困扰人们长达 60 年之久的难题：在不造成严重缺陷的情况下，将血氧含

量维持在什么程度能使早产儿的生存率最大化。该实例阐述了在所有可得研究结果出来之前就可以考虑如何设计 Meta 分析，而 Meta 分析也能被用于一组相似研究完成之后。

例如，1974 年一位瑞典医生开展了一项系统评价来比较乳腺癌单纯手术与手术联合放疗的结局。⁶他发现在所有研究中，接受放射治疗的女性病死率更高。用 Meta 分析对所有证据进行统计学合并之后，明确了高出的病死率并非由偶然因素引起。之后基于多个单病例资料的证据进行了更为详细的分析，证实了在当时应用放射疗法确实会增加病死率。⁷意识到这一点之后，更安全的实践规范出台了。

识别系统评价中的既得利益与过度报告

如果还有其他的利益冲突影响研究者开展和解读系统评价，那么后果会如何？评价员有可能收取了需要验证的医药公司推出的新疗法的酬劳。在评价用于治疗湿疹的月见草油疗效时，那些与药厂有关联的评价者得出的结论比那些没牵涉任何商业利益的评价者得出的结论更为乐观（见第 2 章，第 18-20 页）。然而，商业利益并不是导致评价产生偏倚的唯一因素，无论是研究者、卫生医疗专家还是患者，都可能因自身的偏见而导致偏倚。

遗憾的是，接受了既得利益的人们会放大偏倚，使治疗看上去比实际更有效（亦可见第 10 章）。⁸一些研究者通常（但非绝对）出于商业目的故意忽略一些已有的证据。他们在设计、分析以及报告研究的过程中粉饰某项治疗的结果，使其获得有力的支持。这是发生在 20 世纪 90 年代的案例，当时抗抑郁药赛乐特（帕罗西汀）的制造商未公开一项重要的证据，即该药实际上会使青少年患者的抑郁症状加重，促使他们中的一部分人试图以自杀的方式来摆脱该症状。⁹

此外，过度报告也是个问题。研究者把一个试验的结果（比作“香肠”）分篇报告，而不说明这些报告均来自一个研究，这一现象被称作“切香肠”。这样一来，单个结果“阳性”的试验会出现于多种期刊的不同文章中，造成偏倚。¹⁰此外，在开展试验之初就对每个研究进行注册并配以独一无二的标识，这样有助于减轻上述行为带来的后果。

如果没有对所有相关的可靠证据进行评价，后果会怎样？

治疗的科学试验包括系统地评价所有相关的可靠证据，从而明确：已掌握的证据有哪些，这些证据是来自动物实验或其他实验室研究，还是来自健康志愿者参与的新治疗试验，亦或是来自以往的有患者参与的研究。若这一步骤被忽略或做得不好，那么后果很严重：普通患者和研究的受试者会遭受折磨甚至出现不必要的牺牲，而医疗卫生和科学研究的宝贵资源也会被白白浪费。

患者可以避免的伤害

有研究比较了过去 30 多年间教科书上治疗心脏病的推荐意见和同期发表的治疗性科学试验的系统评价证据。¹¹结果显示，教科书中的推荐意见经常有误，而这是因为教科书的编者并未对相关证据进行系统评价。这样造成的影响是灾难性的，有的心脏病患者因此被剥夺了生存的机会（如溶栓药治疗）。另外仍有医生还一直推荐那些早已被科学试验证实具有致死作用的疗法（如抗心律失常药物治疗心脏病发作——见上文和第 2 章，第 14-15 页）。

科学知识日积月累，但科学家并未科学地积累证据

“25 年来，学术研究人员一直在谈论称为‘累积 Meta 分析’的东西：其本质是对特定的干预反复进行 Meta 分析，每当完成一项试验时，将数据添加进 Meta 分析来更新已有的结果，掌握结果的走向。最重要的是，通过这种方式可及时得到具有统计学意义的结果来说明问题，避免了拿更多人的生命冒险和后续不必要的研究。”

Goldacre B. Bad Science: How pools of blood trials could save lives.
The Guardian, 10 May 2008, p16.\

如果不能将研究结果整合到系统评价中，并产生新的证据，则会持续对患者造成损害。无需冷藏和交叉配血的血液替代物是输血治疗中较好的替代选择，但不幸的是，这些代用品增加了心脏病发作和死亡的风险。而对 20 世纪 90 年代后期报告的随机试验的系统评价才明确了血液代用品的危害，但我们本应该且有能力更早几年就认识到其危害。¹

受试者可避免的伤害

未能评估所有相关的可靠证据还会使研究受试者遭受本应避免的伤害。当有已知有效的疗法的情况下，研究人员被不断安排和允许不给受试者使用这些药物。例如，已有可靠证据证实肠道手术的患者使用抗生素可降低术后并发症引起的病死率，但其后很长一段时间里，研究者还继续在临床试验中不给对照组的受试者使用该治疗。研究者未能系统地评价已有的证据，会剥夺参与研究的一半受试者接受已知有效治疗的权利。审查过研究方案并有机会质疑研究者的伦理委员会和研究的资助机构显然忽视了这一重大缺陷。

如果研究者未能系统地评价有关治疗效果已有的证据，那么不仅是等待接受治疗的患者处境危险，健康的志愿者也会受到伤害。试验的第一阶段通常会纳入少量健康的志愿者。2006 年，六位青年男性志愿者在伦敦西部的一个私人研究机构中，被注射了一种从未被用于人体的药物。他们全部遭受了危及生命的并发症——其中一位志愿者失去了手指和脚趾——他们的健康状况也受到了长期的影响。如果有过类似药物严重不良反应的报告被公之于众，¹⁴ 或者研究者系统地评价了关于此类药物疗效已有的证据，¹⁵ 那么悲剧极有可能不会发生。¹³ 理想情况下，他们或许根本不会开展这项研究；亦或是即便他们决定试试，也会逐个而非同时向志愿者给药；而且他们本该提醒健康的青年志愿者可能发生的危险。¹⁶

卫生保健和研究中浪费的资源

没有对相关的可靠证据进行系统评价，这一做法即便不对患者和研究受试者造成伤害，也会导致卫生保健和医学研究资源遭到浪费。譬如，在 20 世纪 80 和 90 年代，共有 8000 多名患者参与了几项关于治疗卒中新药的试验。荷兰研究人员系统评价了这些药物研究的结果，但并未发现该药有任何益处（见第 10 章，第 121 页）。¹⁷ 之后他们决定评价此前关于该药的动物实验结果，还是没能发现任何益处。¹⁸ 如果这些开展临床试验和动物实验的研究人员事先系统地评价过以往动物研究的结果，那么上千名患者就很可能不必参加这些临床试验了。事实上，这样能让用于治疗卒中患者的资源得到更好的利用，并更有可能开展与该疾病

治疗更相关的研究。而这远不是一个个案。¹⁹

事先查找证据能否避免死亡？

“这本是一个不该发生的悲剧。Ellen Roche，一个 24 岁健康的志愿者在 2001 年 6 月参与约翰·霍普金斯大学的一项哮喘研究时死亡，其原因是该志愿者被要求吸入的化学药物引起了进行性肺衰竭和肾衰竭。这次事故显然是因为开展试验的研究者和批准该项研究的伦理专家组疏忽了该化学药物（即 Roche 吸入的六烃季铵）的危险性。讽刺的是，有关该化学物质危险性的证据可以轻而易举地从已经发表的文献中获得。《巴尔的摩太阳报》（Baltimore Sun）认为，指导医师 Alkis Togias 博士在研究该药物的副作用时做出了‘真诚的努力’，他显然只检索了有限的资源，包括 PubMed，只能检索到 1966 年以后发表的论文。而早在 20 世纪 50 年代发表的文献就提示了六烃季铵相关的肺损伤，还被多次引用。”

Perkins E. Johns Hopkins Tragedy. *Information Today* 2001; 18:51-4.

系统评价应该是新研究报告的起点和终点

有一项类固醇治疗急性创伤性脑损伤的疗效评价，其研究²⁰报告展示了如何回答 Bradford Hill 提出的四个问题。研究者称，他们着手进行研究是因为他们所做的现有证据的系统评价，以及临床上有关该治疗的其他所有证据均表明，该药物被广泛应用，但其疗效存在严重不确定性。报告称，在研究开始之前，他们已经注册并发表了研究方案。

医学期刊《柳叶刀》投稿须知：将研究结果和以往证据对比

系统评价

该部分应该描述作者是如何检索出所有证据的。作者应描述评价证据质量的过程，以及如何筛选并合并证据。

解读

作者应在这部分说明当他们的研究和以往的研究相比，该研究对于证据整体而言有何意义。

‘We ask that all research reports – randomised or not – submitted from Aug 1 . . . put the results into the context of the totality of evidence in the Discussion.’

“我们要求从 8 月 1 日起投稿的所有随机或非随机研究的报告……在讨论部分，把系统评价的结果，置于整个证据的大背景当中。”

Clark S, Horton R. Putting research in context-revisited.

Lancet 2010;376:10-11.

他们还介绍了用于将偏倚最小化并充分控制偶然因素影响的措施，如足够大的研究样本。他们报告了其研究结果，说明了严重颅脑损伤患者使用类固醇可能增加死亡风险。

最后还有很重要的一点，为阻止成千上万的人死于这种广泛使用的治疗措施，他们在报告中向读者提供了所需要的所有证据。因为他们通过自己的研究产生了新证据，并将其整合

进了他们对此前研究的系统评价中。

要点

- 单个研究很难提供充足的证据指导卫生保健中的治疗决策
- 评估不同疗法的优劣需基于针对所有相关可靠证据的系统评价
- 在单个验证疗效的研究中，必须采取方法减少偏倚和偶然因素的误导
- 不考虑系统评价的研究结果会使患者遭受不必要的伤害，还会浪费医疗卫生和研究的资源

第9章 对科学试验的监管：是推动还是阻碍？

现在你将会意识到，很多时候，对疗效的严格评价少之又少，而且疗效的不确定性依旧存在。更有甚者如第5章所提到的，一些主流的观点仍非常活跃，阻碍了医生与患者协作去探究治疗的效果。而且令人不解的是，大多数国家的医疗研究监管体系对这一问题的处理是将研究和治疗人为地分开。研究被看作是高风险行为，需要被严格监管，但常规治疗则被认为稳妥得多。但就像我们介绍的，即便不参加医学研究，普通患者也可能因为接受了未经评估或是草率评估的治疗而身处危险。

谁说医学研究有害健康？

“多数关于医学研究的伦理学探讨都致力于解决应如何监管医学研究这一问题。事实上，医学研究管理在很多方面都比医疗实践严格得多。但在研读了无数的医学研究指南之后，你可能会放弃‘医学研究像吸烟一样有害健康’的想法了。”

Hope T. *Medical ethics: a very short introduction.*

为什么研究被认为是有风险的且需要特殊监管，而常规治疗（会影响更多病人）却不是呢？人们没有遗忘研究者滥用权力的那段历史，在当时病人被视为一种为达目的而被利用的工具。另外研究确实会常常出问题，所以不可避免地产生了大量可怕的故事。也有人一直担心，一旦人们成为研究对象，那么他们个人的利益对于医学专业而言可能会远不如整个研究利益重要。

研究者各种各样的动机使得情况更加复杂：尽管一部分研究者开展研究的初衷就是造福大众，但其他研究人员明显是奔着钱去的，亦或是为了促进个人事业的发展，也有很多时候其动机让人琢磨不透。因而研究对于患者或公众而言可谓是一副令人恐怖的景象，这在某种程度上是因为对卫生研究的严格管制。

独立委员会，也叫研究伦理委员会（Research Ethics Committees, RECs, 欧洲名称）或机构审查委员会（Institutional Review Boards, IRBs, 美国名称），帮助人们免受以研究为名义的侵害。它们负责审核每项科研项目，并对它们实施与否给出意见；它们在研究的监管过程中也起到了重要作用，并向公众保证批准的研究是为公众利益着想而设计的。

委员会通常由无偿的志愿者组成，其中包括非专业人员。他们会评审多种不同的研究方案（即研究者对后续工作的计划），并向那些可能参与研究的人提供全部信息。委员会可以要求研究人员对计划书或受试者的信息进行调整。没有委员会的批准，研究将不能进行。由此委员会可以帮助并确保研究受试者避免承担不必要的风险，并向受试者和公众保证研究人员不会仅以个人意愿行事。

研究还要遵守许多其他形式的规范。大多数国家都出台了一些针对研究的法律。譬如，所有欧盟国都必须遵守《临床试验准则》，其中列出了和所谓的“医疗制品临床试验”（其本质是药物试验）有关的要求。部分国家还建立了用于所有或大部分类型的医疗卫生研究的监管系统。另外，有许多其他法律，尽管其出发点并非针对研究，但也能对研究产生潜在的作用。比如数据保护法原本是为了保护个人数据的机密性，但在许多国家也应用于医疗研究。

多数国家也有各种各样的机构涉足研究监管。

研究的开展也可以通过专业的行为规范和国际申明加以监管。例如，医生和护士需要遵守相关的专业体系行为规范，一旦违背，他们就可能会失去执业资格或者受到其他处罚。至于国际申明，如世界医学协会的《赫尔辛基宣言》，即使没有法律效力，但也对标准的制定有着深刻的影响。

监管系统是否让科学试验做得更好？

虽然目前的监管力度卓有成效，可当有人想要研究某种评价不充分的治疗，而不是将其应用于常规临床实践时，现有的监管系统却让他们寸步难行。在很多国家，高度复杂的体系（包括法律、机构、行为规范等等）给研究造成沉重的负担并浪费了大量的时间。研究者可能需要获取不同部门的多项批准，有时还必须面对相互矛盾的要求。

此外，总的来说，这一系统可能会严重拖延某些信息的收集，使其积极性备受打击，而正是这些信息能保证人们接受的医疗保健更加安全。以数据保护法和保密性行为规范为例，尽管两者初衷是美好的，但它们使得研究人员在收集常规的医疗记录数据用以明确治疗的副作用时困难重重。而对于计划开展临床试验的研究人员，则可能需要数年的时间才能招募到第一个患者，甚至试验招募受试者的过程都可能会因监管的需要而被放缓。但当研究人员遵循整套监管系统进行研究时，患者正在遭受不必要的病痛折磨甚至失去生命。

一个理想世界

“如果可能，在理想情况下，不考虑那些担心自身隐私胜过他人生命的人，我们可收集匿名的结果数据并将其与以往药物治疗比较……在理想世界，无论患者在哪接受什么治疗，当对何为最佳治疗真正存在不确定时，他们能够被直接有效的随机分到某个治疗组，而治疗进展随时得以监测。在理想情况下，这些观念将会自然而然地存在于我们对医疗卫生的认知中，以至于患者根本感觉不到被研究所打搅。”

Goldacre B. Pharmacology would be fascinating enough even if society didn't manage it really really badly. *TheGuardian*, 17 July 2010.

Available online: www.badscience.net/2010/07/pharmacology-would-be-fascinating-enough-even-if-society-didnt-manage-it-really-really-badly

在临床实践中，如果一种疗法在以往的“常规”实践中已被使用，那么只要患者同意，就意味着医生可以对患者实施未经验证的治疗。相比之下，想要通过开展研究来合理地评估同样的治疗就必须经历拖沓的审查程序。最终医生会失去科学验证治疗措施的积极性，反而继续在处方中使用未经评估的那些疗效不确定的治疗方案（见第5章）。

偏倚伦理学

“所谓的研究就是临床医生通过仔细研究、评价结局和发表结果来验证新疗法。此类研究的受试者应得到特殊保护。研究方案必须通过机构审查委员会（IRB）[相当于英国的伦理委员会]的审查。知情同意书必须经过严格的审查，部分研究可能会因这一步而被叫停。但还有种情况，临床医生可能根本不是出于探索的心态而验证新的疗法，而仅仅是因为他相

信新疗法对患者有效。在这种情况下，这样的尝试不是研究，也不需要IRB的批准，知情同意的获取也只是为了避免医疗纠纷而已。

不难看出，患者在第二种情况下（非研究）承担的风险远远大于第一种情况（正式医学研究的一部分）。此外，在第一种情况下，医生更具有伦理学观念。在第一种情况下医生是在对治疗进行评价，而在第二种情况中医生是在根据他并不可靠的直觉来选用治疗。然而，伦理学的规范旨在保护那些致力于创造最普遍的知识的患者，所以它只能用于约束那些有责任的研究者，而对不负责任的冒险者不起作用。”

Lantos J. Ethical issues – how can we distinguish clinical research from innovative therapy? *American Journal of Pediatric Hematology/Oncology* 1994;16:72-75

在规避风险和保护潜在研究受试者的过程中，研究监管系统显得过于谨慎，却忽略了一个事实：越来越多的患者和公众都已经作为合作伙伴在参与研究（见第 11 章）。然而，有一点值得肯定，研究监管人员开始意识到用于伦理审查的“通用型”方法可能过于繁琐。¹ 比如，英国正在评估“合理审查”的程序是否能在不引起任何重大伦理问题的前提下，更加简单灵活且安全地用于研究。

信息与知情同意

作为监管系统的手段之一，信息提供以及知情同意的相关要求会阻碍而非鼓励研究解决治疗方案的不确定性。考虑当前接受某种治疗的所有患者的利益，比仅关注少数参加临床试验的受试者的利益，要更为重要和符合伦理。² 因此，不论现在接受治疗的对象是否参与正式的治疗措施评估，他们的治疗知情同意的标准应该和受试者是一样的。为了作出与自己价值观和意愿一致的决策，患者应该及时掌握尽可能多的信息。

将治疗方案用于日常实践，就意味着人们会有不同的偏好和要求，而且可能因人而异。我们还发现人们需要的证据不仅可能会在数量和类型上有所不同，还会因他们对随时获取到的信息的理解能力差异而不同，也会因他们的焦虑和恐惧程度而不同。卫生专业人员应该根据不同个体在特定时间的需要，帮助他们快速准确地选择治疗措施。

反思知情同意

“一些人已经开始怀疑知情同意并非高质量生物医学实践的基础，而且……试图让其变成实践的基础既没必要也不现实。我们希望在更短的时间内改革并精简那些五十年来主宰知情同意的标准。”

Manson NC, O'Neill O. *Rethinking informed consent in bioethics*.

Cambridge: Cambridge University Press, 2007, p200.

然而，在研究中，监管部门会监督向受试者提供信息的过程，他们常常会执拗于在人们被邀请参与研究时就应“竭尽所能”地将所有可能的相关信息呈现给他们。这可能会使那些更愿意“把问题留给医生”的人感到不必要的灰心、失望或害怕，或是引起不必要的担心。

第5章（第57-58页）中提及的对早产儿使用咖啡因的临床试验，就很好地证明了坚持将所有信息毫无保留地提供给潜在受试者会造成怎样的危害。该研究纳入了世界各地的2000多名早产儿，但由于试验对象招募缓慢，其耗时超出所计划的一年时间。特别是在英国招募受试对象的进展得尤为缓慢，几个研究中心因为监管机构对审批流程的延误而退出研究。最为重要的是，研究伦理委员会坚持要求告知家长们使用咖啡因可能会诱发婴儿肌肉痉挛，但这一并发症只会在咖啡因的使用剂量超出十倍时出现。由此，父母们显然接受到的是令人恐惧却毫无必要的信息；而且如果咖啡因仅仅是作为常规治疗的一部分，或许他们就不会也不需要了解这些信息了。

少有证据能够证明这种被广泛应用的研究监管形式利大于弊。⁴事实上，现有的证据着实令人担忧。例如，在那些需要刻不容缓地给予治疗并评价其疗效的研究中，研究者往往必须“墨守陈规”，书写知情同意，这不仅会引起本应避免的死亡，还会低估治疗效果。⁵

获得知情同意可能是一项弊大于利的公共卫生干预。像其他初衷美好的干预措施一样，它的效果应该接受严格评估。如果研究伦理委员会愿意负起责任并有力证明自己的“处方”利大于弊，那么上述致命的后果可能在几十年前就被明确了。

在意识到医患之间的信任是一切满意咨询的基石后，用一种灵活的方式向潜在的研究受试者提供信息，将大大优于死板、标准化的方式。但限于监管系统干预研究的方式，目前临床医生无法自由选择如何向患者解释研究；而且他们常常发现很难谈论研究中固有的不确定性。正如第5章所述，医生在招募患者参与临床试验时，常会违心地说：“我不知道哪种治疗是最好的”，而患者通常不愿意听到这样的话。因此，无论是医生还是患者都需要对不确定性有更合理的认识，并更好地理解“研究为何必不可少”（见第11章）。

在高质量的医疗实践中进行知情同意的一般方法

“在围绕知情同意的讨论中，实质上缺少的是患者的理解，患者想要了解哪些信息，以及如何对待那些只希望了解最基础信息的患者。评估患者对所得信息的理解程度，这一工作几乎无人问津。临床医生发现很难确定患者或其家属对所得信息的正确理解程度。理解力会因信息提供者、信息的解读方式，以及提供信息的时间或环境而受影响。家长式的作风在医疗实践中并不为人接受；公认的方法是清晰地解释并给出满足患者需要的信息，这就是高质量的医疗实践。”

Gill R. How to seek consent and gain understanding. BMJ 2010;341:c4000.

学术争论，还是明智选择？

“12年前，我33岁，那一年我从医生变成了患者：被查出患有乳腺癌。当时我正在攻读博士学位，应用随机对照试验（RCTs）来评估口腔正畸学（我的专业）的治疗效果。讽刺的是，在开展自己的研究时，我意识到了参加临床试验的价值和治疗年轻的早期乳腺癌患者的不确定性。所以在接到诊断后我就向自己的主管医生咨询，是否有我可以参加的随机对照试验。他的回答让我很震惊。他告诫我说：“千万不要因为拘泥于学术争论而妨碍我最佳的治疗。”但何为最佳治疗？我全然不知，并且我也意识到了为什么医学界还在找寻‘什么是50岁以下早期乳腺癌患者的最佳治疗措施’的答案。而我接下来应该做什么？”

哪些是监管系统没有做到的？

尽管在研究开始前，监管系统就给研究人员施以严格要求，但仍有很多方面他们没有做到或是做得不好。许多系统并未充分确保申报的研究是切实需要的，例如，他们并未要求研究人员在开始新研究之前证明他们已经充分全面地评价了已有证据（见第8章，为何系统评价如此重要）。

此外研究监管的大部分精力都放在起始阶段，注重研究受试者的准入。而在以下方面投入的精力微乎其微：研究实施过程中的监管；确保研究者在研究后期（或结束后）及时发表研究报告；确保研究者阐明他们采取了何种手段来减少不确定性。

研究监管系统应该做什么

“如果伦理学家或其他人要给临床试验挑毛病的话，他们应该去关注那些不够充分科学的研究工作和重复的研究，最重要的是对受试者不合理的排除以及资源不公平、不合理的利用。但当前的讨论其实都忽略了一个问题：试验的根本目的是什么？是为了证明我们使用的治疗是安全的，并且比其他治疗更有效。伦理学上没有捷径，临床试验也不例外。”

Ashcroft R. Giving medicine a fair trial. *BMJ* 2000;320:1686.

人们在受邀参加疗效研究时要确保该研究有价值，这样他们所做出的贡献才是有意义的。监管系统需要更多行动来确保以上两点，并清除不必要的障碍，为那些探究患者真正关心的问题研究铺平道路。越来越多的人意识到了科学试验是每个人的责任。随着患者与公众开始参与研究的策划与实施（见第11章），他们在解决监管障碍方面也越来越多有发言权。

要点

- 研究监管过于复杂
- 当前研究监管系统阻碍了那些能改善医疗卫生水平的科学试验
- 尽管有繁琐的监管准则约束研究者，但监管系统几乎未能保证申报的研究是真正必要的
- 研究监管并未有效地监控和随访已批准的研究

第 10 章 高质量、低质量和不必要的研究

我们在前文中已强调了合理设计临床治疗试验和着手解决那些对患者及公众重要的问题的必要性。如果能做到这些，即便研究结果和期望不相符，我们仍应感到欢欣鼓舞，因为它降低了不确定性，使我们距重大的发现又迈进了一步。

尽管有很多的高质量研究正在严格按照设计和报告标准¹稳步自我完善，但出于各种各样的原因，低质量和不必要的研究仍在继续。鉴于“需要更多的研究”这一永久的需求，明智的选择是不贪多，而是关注那些针对患者需要的研究，并明确这些研究的必要性。我们将在本章节就以上问题展开讨论。

高质量的研究

卒中

卒中是引起死亡和长期瘫痪的主要原因。首次卒中的病死率介于 1/6~2/6 之间，随后发生的卒中病死率约为 4/6。引起卒中的主要病因之一是向大脑供血的颈动脉狭窄。附着在颈动脉内壁的脂肪物质有时会脱落并阻塞细小的动脉分支，从而引起卒中。在 20 世纪 50 年代，外科医生开始使用颈动脉内膜切除术来移除这些脂肪沉淀物。手术目的旨在降低卒中的风险，但就手术本身而言，存在引起并发症的风险。

尽管颈动脉内膜切除术不断得到推广和普及，但评价其利弊的随机试验直到 20 世纪 80 年代才得以展开。显然研究结果对患者和他们的医生来说都意义重大。两项经过精心设计的试验（分别在欧洲和北美地区开展）纳入了那些已出现颈动脉狭窄症状的患者（如，轻微卒中或短暂的卒中样症状），目的是比较手术治疗和当时可得的最佳非手术治疗的效果。有几千名患者参与了这两项长期研究。在 20 世纪 90 年代发表的研究结果指出：手术治疗可以降低卒中或死亡的风险，但这都取决于颈动脉狭窄的程度。总的来说，手术本身会使轻度狭窄的患者引发卒中，对其产生危害。这些重要的研究结果对临床实践产生了直接的影响。^{2,3}

孕妇的先兆子痫

另一项高质量研究的典型例子和孕妇有关。全世界每年约有 60 万妇女死于妊娠相关的并发症。此类死亡多发生在发展中国家，其中很多和妊娠相关抽搐（阵发）有关，这种情况称为子痫。子痫的发作往往是灾难性的，它会引起孕妇和胎儿的死亡。妇女子痫的先兆症状发作（先兆子痫，也称毒血症）时，会出现高血压和蛋白尿。

1995 年的一项研究发现，作为一种廉价的常用药，硫酸镁能预防子痫发作妇女的复发。该研究还指出硫酸镁的效果优于其他抗惊厥药，尤其是一种极为昂贵的抗抽搐药物。因此，研究者认为探索硫酸镁是否能预防有先兆子痫的孕妇抽搐发作意义重大。

为此，研究者设计了 **Maggie** 试验来解答这一问题，并取得了重要成就。该试验在全球 33 个国家中选出 1 万多名先兆子痫发作的孕妇，在给予常规治疗的同时，向 1/2 的妇女注射硫酸镁，另 1/2 使用安慰剂（虚拟治疗）。试验结果清晰有力地证明硫酸镁不仅可以降低 50% 抽搐发生的几率，而且，尽管试验中硫酸镁注射治疗未能显著降低胎儿死亡风险，但有证据

证明它能降低孕妇死亡风险。同时，硫酸镁副作用轻微，而且对母婴无害。^{4,5}

我参加 Magpie 试验的经历

“我很高兴自己能成为如此重要试验中的一员。我在怀孕第 32 周时出现肿胀，随后逐渐加重，在第 38 周时被确诊为先兆子痫，并接受了剖腹产手术。幸运的是我和宝宝术后都恢复得很好。先兆子痫让我心有余悸，我也衷心希望这项试验的结果能对像我一样的妇女有所帮助。”

——Clair Giles, Magpie 试验受试者

MRC News Release. Magnesium sulphate halves risk of eclampsia and can save lives of pregnant women.

London: MRC, 31 May 2002.

儿童感染 HIV 病毒

高质量研究结果也对感染 HIV（人类免疫缺陷病毒，艾滋病的原因）的儿童产生了巨大影响。2009 年末，来自联合国艾滋病规划署（the joint United Nations Programme on HIV/AIDS, UNAIDS）的数据显示，据估计全世界有 250 万感染艾滋病病毒的儿童，其中 230 万生活在撒哈拉以南的非洲地区。每小时约有 30 名儿童死于艾滋病。⁶与儿童免疫系统功能低下相关的一类细菌感染（如肺炎）是引起死亡的主要诱因。复方新诺明作为一种普及较广、价格低廉的抗生素，被长期用于治疗非艾滋病引起的儿童和成人的呼吸系统感染。有研究指出复方新诺明可缓解 HIV 感染的成人细菌感染后出现的相关并发症。⁷

当有初步证据显示复方新诺明能降低艾滋病患儿的细菌感染时，一组英国研究人员联合他们在赞比亚的同事，开展了一项大规模研究来评价复方新诺明作为一种潜在预防药物的效果。这项试验开始于 2001 年，历时两年，有超过 500 名儿童参与，比较了复方新诺明和安慰剂的效果。结果比预想要来得快，复方新诺明不仅将艾滋病相关的病死率降低了 43%（复方新诺明组 74 例死亡而安慰剂组 112 例死亡），还降低了入院率。因此，独立委员会审查了研究结果并要求终止试验。

这项试验一个立竿见影的效果就是赞比亚政府采取措施，给所有参与试验的儿童发放了复方新诺明。更广泛的效果是世界卫生组织（World Health Organization, WHO）和联合国儿童基金会（UNICEF）及时地修改了对艾滋病患儿的用药建议。^{8,9}

以上组织此后也持续将复方新诺明列为一种用于治疗儿童 HIV 感染者的廉价、安全和挽救生命的药物。¹⁰

低质量的研究

精神障碍

遗憾的是，研究并非总是高质量完成，并非总有意义。以一种颇为痛苦的疾病为例——迟发性运动障碍。在治疗精神障碍（尤其是精神分裂症）时，有一类指定用药称为精神安定剂（抗精神病药），迟发性运动障碍是由于长期使用此类药物而引起的一系列副作用。迟发

性运动障碍最突出的特征是嘴唇和脸部反复不自主地运动，具体表现为鬼脸状、咂嘴、频繁伸舌和鼓腮等，有时还会伴发四肢抽搐。服用抗精神病药超过三个月的患者有 1/5 会出现以上副作用。

20 世纪 90 年代，一组研究人员系统地回顾了前 30 年中对迟发性运动障碍采取过的治疗措施。在 1996 年的报告中他们惊奇地发现，经确认共开展过 500 项随机试验，其中包括 90 种不同的药物治疗方案。然而没有一项试验能提供有用的数据。有些试验因纳入患者人数过少使得结果不可靠；其他试验由于治疗干预时间太短而被认定没有意义。¹¹

这一组研究人员后来发表了一篇全面的调查报告，总体上评价了和精神分裂症治疗相关的随机试验的内容和质量。他们共纳入了 2000 项试验，但结果令人失望。要肯定的是，多年来，药物治疗在某些方面改善了精神分裂患者的预后和转归。例如，大多数患者可以在家里或是社区中接受治疗。然而，即便是在 20 世纪 90 年代（甚至今天），大多数药物仍处于临床试验阶段，因而它们用于门诊治疗的效果还不确定。更重要的是，不同研究对疗效的评价天差地别，令人震惊。研究者发现，曾针对 600 多种治疗方案（主要为药物治疗，也包含心理疗法）开展过试验，却有 640 种不同的标准来评价研究结果，而其中 369 种标准只用过一次。这使得在比较不同试验的结果时障碍重重，而且医生或患者也无法理解这些结果的实际意义。研究者还发现了一些别的问题，很多研究规模太小或是持续时间太短导致结果没有意义；研究中经常使用明显超出常规剂量且副作用突出的药物和新药进行疗效对比，并忽略其他目前可得的耐受性更好的治疗，这明显是不科学的试验。这篇评价的作者们最后总结到：半个世纪以来，相关研究的质量低、持续时间短、临床应用有限，因而需要精心设计、恰当实施和规范报告的试验来填补空缺。¹²

对分娩孕妇进行硬膜外麻醉

硬膜外麻醉缓解孕妇分娩时疼痛的早期试验作为一个反面案例，强调了评价患者重要结局的重要性。20 世纪 90 年代，研究人员评价了以往对比硬膜外和非硬膜外麻醉的对照试验。据他们估计，尽管过去的 20 年内有几百万妇女接受了硬膜外阻滞麻醉，但似乎只有不到 600 人参与过设计合理、无偏倚的试验来对比该方法和其他镇痛方法。他们筛选出 9 项符合要求的试验，这些试验基本上都将激素水平和其他一些被认为能反映生产时压力的物质作为评价指标。一些试验还关注了婴儿的结局。然而除了两项试验外，其他试验都没有考虑到孕妇自身对疼痛的先后比较。换言之，这些试验中的研究人员竟忽略了一项最重要的研究结果——孕妇的疼痛是否得到了有效的缓解。¹³

不必要的研究

早产儿呼吸窘迫

一些研究既不属于高质量研究也不是低质量研究，而是显然没有必要的研究。以一项和早产儿有关的研究为例。早产儿的肺大多尚未发育完全，有引起危及生命的并发症的风险，如呼吸窘迫综合症。早在 20 世纪 80 年代初，已有证据很好地说明了对有早产风险的孕妇使用类固醇类药物可降低早产儿的呼吸窘迫综合症发生率和病死率。然而在此后的 10 年中仍有大量试验将类固醇类药物和安慰剂对比或是进行空白对照。如果当时已经对之前那些试验的结果进行了系统评价并做了 Meta 分析（见第 7、8 章），那么后续的这些试验很可能就不

会开展了，因为系统整合后的证据已经说明更多的试验是没有必要的。然而，不必要的试验使参与进去的半数患者错过了有效的治疗。

卒中

另一个不必要的研究案例是关于用一种叫尼莫地平的药物(钙拮抗剂的一种)治疗卒中，究其原因也是由于没有汇总和分析以往的研究结果。如果有可能在患者卒中时控制大脑受损的程度，那么他们瘫痪的几率就会大大下降。基于这一考虑，20世纪80年代初在动物实验取得喜人成果后，尼莫地平就开始用于患者的临床试验。尽管1988年发表的一项的临床试验认为尼莫地平对卒中患者有益，但另外几项关于尼莫地平和其他钙拮抗剂的试验结果却与之相悖。直到1999年，在对共纳入近8000名患者的多项临床试验累积的证据进行系统评价之后才发现此类药物是无效的(见第8章，第102页)¹⁴。既然尼莫地平的使用是基于确凿的科学证据，为什么还会出现这样的结果？

人体研究的结果被曝光后，动物实验的研究结果才开始被仔细地审视。直至动物实验经系统地评价之后人们才清楚地认识到，这些实验的设计质量普遍很低，而且其结果被大量偏倚所干扰，最终使结果不可靠。换言之，从一开始就没有令人信服的理由支持对卒中患者开展试验。¹⁵

抑肽酶对术中和术后出血的影响

研究资助者、学术机构、研究者、学术伦理委员会和学术期刊都对不必要的研究起到了推波助澜的作用(见第9章)。正如我们在第8章中所提到的以及前两个不必要研究的例子所揭示的，在没有对已知研究进行系统评价前，不应该设计或是开展新的研究。

2005年，一篇令人震惊的研究报告揭露了有关抑肽酶(一种用于减少术中和术后出血的药物)对照试验的情况。抑肽酶确有其效，但令人震惊的一点是虽然已有大量可靠证据证明该药能大大减少输血，仍有对照试验在不断开展。¹⁶在该报告发表时，有64项试验的结果已被发表。在1987年到2002年之间，后续的抑肽酶试验结果中引用之前相关研究结果的比例从最高33%跌落至新近发表的10%。在44篇后续论文中，只有7篇引用了规模最大的一项试验(其规模为所有试验规模的28倍)的结果；而44篇论文中没有一篇引用了1994年和1997年间发表的相关系统评价。

这篇文章的作者强调，科学理应是一个积累的过程，但是许多科学研究者并没有科学地去积累证据。不仅大多数新的研究设计未基于已有证据的系统评价，而且新的研究证据也很少被纳入到系统评价更新的结果中(见第8章)。

歪曲研究的优先性

大多数组织机构都在支持生物医学的研究，而且大量研究人员也将精力投入其中，他们的目标非常明确：为提高大众健康水平而提供信息。然而在每年数以百万计的生物医学研究报告中又有多少能真正对这一崇高的目标作出贡献呢？

患者关注的问题

布里斯托大学的研究人员提出了一个根本性的问题：“对膝关节骨关节炎患者和他们的医生而言，研究中什么程度的问题才算是他们关注的问题。”¹⁷ 他们一开始组建了分别由患者、风湿病学家、理疗师和全科医师构成的四个观察小组。这些小组都一致表明不建议开展更多由制药厂资助的对比不同种类非甾体抗炎药（此类药物包含布洛芬）和安慰剂的试验。此外，患者还希望严格评估外科手术和物理疗法，评估患者教育方案的应对策略，从而能帮助患者成功应对这一慢性、致残和引起疼痛的疾病。当然，以上这些形式的治疗和应对方案无法像药物那样提供可观的经济收益，所以常常被忽视。

用以上的标准评判，其他方面的治疗性研究中有多少会出现类似的问题——患者和医生所关注的问题与研究者正解决的问题不一致呢？遗憾的是，这种不一致更像是一种常态而非例外。^{18,19,20,21}

药品配方的局部改动使药物焕然一新、效果更佳的情况少之又少，但在对关节炎及一些其他慢性病的治疗性研究中这种模式却成为主流。这真是对资源极大的浪费！

谁来决定研究什么？

上述情况显然令人失望，那么这是怎么回事呢？一个重要的原因是研究者所关注的内容被外界因素扭曲了。例如，制药公司开展研究是为了自身的基本需求，即完成自己对股东们应尽的义务——盈利。而制药商对患者和医生的责任则排到了第二位。商业是被巨大的市场所驱动的，而市场可以是考虑是否使用激素替代治疗的妇女，也可以是陷入压抑、焦虑、失落甚至忍受疼痛的人群。而在近几十年中，这种以商业目的为主导的研究模式很少能产生重要的新疗法，甚至扰乱了“大众市场”。相反地，对不同种类的药物，商家还经常会制造很多和它们相似的新药，也就是所谓的“me-too”药。这让人不禁想起当超市只能供应面包的日子里，那些白色面包片上千变万化的图案。至此，我们也就不会惊讶于制药商对市场的投入超过对研究的支出了。

但商家又是如何说服开处方的医生选择这些新产品，而非现有的更廉价的药物呢？一个惯用的伎俩是委托他人开展多项小规模的研究来证明使用新药比不用药的效果好，而不去设计试验来比较新药和已有药物的效果。可悲的是，制药商总是能顺利找到那么一些医生，愿意劝说自己的患者参加此类毫无意义的研究。而最终这些医生常常会在处方中使用通过上述途径研发的药物。此外，药物审批权威机构还常常推波助澜，坚持新药只需要和安慰剂进行对照，而不是那些已有的有效治疗措施对照。

“Me-too 药”在加拿大的影响

“1996 年至 2003 年，不列颠哥伦比亚省大多数（80%）的药品消费增幅都来自于具有专利新药的使用，但这些新药和那些 1990 年之前已经存在廉价药物相比并没有实质性的进步。‘me-too 药’比那些经过时间考验的‘老药’高出的巨额费用值得我们认真审视。借鉴新西兰所使用的药物定价手段或许能节省开支，并用于其他医疗保健需求。举例来说，如果不列颠哥伦比亚省 2003 年消耗的“me-too 药”中有一半的定价和以往使用的“老药”相当，那么将能节省 3.5 亿美元（占处方药总费用的 26%）。省出的这部分资金能够使得超过 1000 名的新医生加入医疗行列。”

目前全球销量前 20 名的药物中就包含了已有药品目录中出现的一些新的专利药…… ‘me-too 药’ 很可能在多数发达国家内左右着消费趋势。”

Morgan SG, Bassett KL, Wright JM, et al. ‘Breakthrough’ drugs and growth in expenditure on prescription drugs in Canada. *BMJ* 2005;331:815-6.

另一种手段就是找人代笔。代笔就是一名专业写手写出文章后却被冠以别人的名字。大多数人可能马上会联想到“名人自传”就是这样被“代笔”而成的。其实学术出版物中也会混有代笔之作，而且其潜在后果令人担忧。有时制药商会雇佣传媒公司来准备一些美化和宣传公司产品的文章。文章一旦完成，就会有学者签约，并因为“代言”了该文章而获得“报酬”。随后，文章会被投稿出版，评论类文章这时将倍受青睐。制药商还比较青睐杂志的增刊，即额外的独立出版物，但能使用母刊的名字，而且常常由制药企业资助，不会像母刊一样需要通过严格的同行评审。通过以上方式制造和推广的市场信息导致产品被过分吹嘘，而且其带来的负面作用也被弱化了。（见 97 页，第 8 章）

医生和制药公司

“没有人知道制药公司总共给了医生多少钱，但是根据美国排名前九的制药公司的年报，我估计这个数字约为每年几百亿美元。通过这种方式，制药商就很好地控制了医生对自己产品的评价和使用。这在很大程度上也有赖于医生，尤其是知名医学院中的资深教授，因为他们能左右研究结果、药物试验方法，甚至还能定义是什么导致了疾病。”

Angell M. Drug companies & doctors: a story of corruption.
New York Review of Books, January 15, 2009.

制药公司还会在医学期刊上给自己的产品做广告。典型的广告包含了有证据来源的参考文献来支持他们的产品介绍。乍一看这很有说服力，但是当所引用的证据被逐个分析时情况就不同了。即便这些证据来源于随机对照试验，让读者感到其评估可靠，但其真实面目并非如此。研究人员分析了主流医学杂志中的此类广告，了解这些随机对照试验证据的真实情况，他们发现被引用试验中只有 17% 是高质量的，这一小部分试验回答了有关药物效果的问题，且非药物公司赞助。而众所周知，那些由药物公司资助的研究更容易得出对其产品有益的结果。^{25, 26}

骗子、傻瓜和坏蛋？

两名研究者在英国医学杂志（*British Medical Journal*）圣诞特刊上发表了一篇诙谐幽默的文章，他们恶搞编造了一个名为 HARLOT plc 的皮包公司，为试验的投资商们提供整套的服务。内容如下：

“那些制造假药和伪劣医疗器械，但又希望扩大自身市场份额的制造商们，那些希望提供更多不必要诊断和治疗服务的专业医疗机构们，还有那些希望实施不合理而又出于私利政策的地区和国家医疗保健部门们……哄人‘我也一样药’ [我们的 E-Zee-Me-Too 试验设计小组]能够向你们保证阳性的研究结果。”

令人吃惊的是，真有人向该文的作者们咨询有关“神奇”的 HARLOT plc 皮包公司的情况。

Sackett DL, Oxman AD. HARLOT plc: an amalgamation of the world's two oldest professions. BMJ 2003;327:1442-5.

一些知名期刊（如《柳叶刀》）对此类事情的评论，促使我们去关注那些部分知名期刊发表临床研究的不良动机，以及大学和制药公司越来越可疑关系。《新英格兰医学杂志》的前任编辑直截了当地提出一个问题：“学术医学难道是用来推销的吗？”²⁸

忽视了患者利益的生物医学研究模式并不仅仅只受商业利益优先的坏影响。许多高等学府和研究机构的专家都相信医疗卫生的进步有赖于对疾病发病机理的揭示。因此他们在实验室中开展大量动物研究。尽管这些基础研究的必要性无可非议，但几乎没有证据支持它们应该比人体试验获得更大的资助力度。^{29,30}而最终的结果就是实验室研究过剩，而与患者的相关程度却无从考证。

竭尽全力只为找基因

“基因革命一直被认为是解决人类所有问题的希望。当我们能够定位和复制特定基因时，我们能建更好的房子，减少污染，更勇敢地面对癌症，成立更多儿童保育设施的专项基金，对国家体育馆的选址和设计达成共识。很快，所有新生儿都能在基因水平场所被分娩。基因神通广大，比如能让女生在 GCSEs 考试（中学考试）中的表现超过男生。基因拥有着无限的可能性……因此，我们也就进入了一个不确定的世界，但同时也是一个充满希望的世界。因为无论基因问题带来什么样的道德困扰，终有一天人们能隔离相应的基因，解决这个问题。”

Iannucci A. The Audacity of Hype. London: Little, Brown, 2009, pp270-1

造成这种不平衡的原因之一是学术环境中夸大了基础研究（尤其是基因研究）所能带来的临床进展（见 43~44 页，第 4 章基因研究部分）。而作为一名卓越的医生和遗传学研究者，大卫·韦瑟罗尔（David Weatherall）先生在 2011 年观察发现：“很多主要致死性疾病反映的都是大量表达效能较低的基因联合外界和社会环境影响的作用。研究工作为一些疾病的发病过程提供了宝贵的信息，但是也强调了在发病机制后隐匿的个体性和变异性。显然，我们距离以基因修饰为基础的个体化医疗时代还有很长的一段路。”³¹

如今，DNA 双螺旋结构的发现已经过去 50 多年了，早期关于“基因革命”为卫生保健所能带来种种好处的吹嘘也渐渐偃旗息鼓。人们开始面对现实。一名科学家对基因技术促进新药研发表达了看法：“我们已经进入了面对现实的阶段……基因的各个方面都必须和环境、临床用药等因素关联起来。因为一种药物对一位患者无效并非意味着就是基因变异引起的。”³²科学杂志《自然》的一位主编在一期杂志中祝贺了人类基因组序列发现十周年：“……药物靶向治疗部分癌症中的特定基因缺陷，比如一些罕见的遗传疾病，已经取得了一些进步。但是后基因组生物学的复杂程度使得之前让治疗从涓涓细流壮大如洪的希望破灭了。”³³

科学研究未能使银屑病患者受益

“极少有研究会比较治疗银屑病的不同方法，或长期疗效。对一些慢性终身疾病，其研究所持续的时间之短令人难以置信。我们似乎只满足于现有的治疗比空白对照有效。很明显，

研究者已经完全忽视了患者的体会、视角、意愿或满意度。”

——R Jobling，银屑病学会主席

Jobling R. Therapeutic research into psoriasis: patients' perspectives, priorities and interests. In: Rawlins M, Littlejohns P, eds. Delivering quality in the NHS 2005. Abingdon: Radcliffe Publishing Ltd, pp53-56.

通过精心设计的针对患者的临床试验来验证源于基础研究的治疗理论，这是一条必由之路。但对这些理论的研究往往不够彻底，未能验证它与患者的真正相关性。20多年前，研究人员明确定位了引起囊性纤维化的基因缺陷，而现在出现这种症状的患者依旧有一个最根本的疑问——他们何时才能看到这一发现为他们的健康带来福音？

即便研究似乎看起来与患者相关，研究人员也常常在设计研究时忽略患者的意愿。有一个很好的实例是：让一些患了肺癌的医生，站在患者的角度逐一考虑自己是否愿意参与所提供的六项肺癌试验（均符合纳入标准），有 36~89%的医生表示他们不愿意。³⁴

同样，在有关银屑病（一种困扰全球约 1.25 亿人的慢性致残性皮肤病）的临床试验中，患者的利益未能被体现。^{35,36} 例如，英国银屑病学会发现，研究者们在许多研究中持续使用一种极其不正规的评分系统来评价不同的治疗方案。这一评分系统有很多缺陷，其中一项是以患者全身皮肤受侵范围和损伤程度为评估标准；但问题是，对患者而言，最大的困扰是面部、手掌、足底和外阴部的皮肤损害。

要点

- 不必要的研究是对时间、精力、金钱及其他资源的浪费，也会对患者造成潜在的伤害，有悖于伦理
- 只有在对过去所有相关研究做最新的系统评价，且结果显示有必要时，才应开展一项新的研究，而且还要在注册之后方可进行
- 最新研究产出证据后应对以往的系统评价及其他相关证据进行更新
- 很多研究是低质量的，并且其出发点让人怀疑
- 商业和学术界会扭曲研究
- 和患者息息相关的问题常常得不到解决

第 11 章 做好的研究，人人有责

我们在前文中已展示了针对疗效的低质量或不必要的研究会浪费多少时间、金钱和精力，即那些不论是现在还是将来都无法解决对患者来说是重要的问题的研究。我们也希望已让您意识到今后更好的治疗试验应该来自于患者、临床工作者、公众和研究者卓有成效的合作。

患者和公众如何能有助于提高研究质量？

以往壁垒重重的医疗领域已经在逐渐敞开大门，接纳新鲜的观点和之前那些“门外汉”，而且家长式的作风也在逐渐消失。因此，患者和公众也能越来越多地参与卫生保健研究的过程，这其中包括研究什么和怎样开展研究。¹ 在研究的过程中与患者协作、将其作为合作者，这一举动得到国际社会越来越多的支持。同时，那些希望在研究中有患者和公众参与的专业人员也得到有用的指导意见。^{2,3,4}

患者的经验既能加深思考，又能提供不同的视角。他们第一手的信息能很好地提供人在面对疾病时的反应，以及作出的反应对治疗决策的影响。通过大量研究（问卷调查、研究报告的系统评价、个体试验报告和影响评估）^{5,1,6,7} 积累的证据证实了患者和公众参与能够对疗效研究起到促进作用。

患者的选择：大卫和歌利亚

“谁有权力判断所研究的课题是否最大程度上满足了那些承受着种种痛苦的患者的需求？为什么最相关的问题没有被提出？目前是谁在提出问题？谁应该来提出问题？谁应该来决定优先顺序？患者才是提出与自身最相关的卫生主题的最佳人选。只有他们能讲述自己的切身感受、健康状况、生活质量及其相应程度。患者就像是大卫，他们必须准备好自己的投石器来和歌利亚抗争，而歌利亚则是那些需要凭借证据在市场上出售产品并盈利的制药公司，以及那些受好奇心驱使而同时又要保证自己的科研经费、学术地位和事业发展的研究人员。只有当核心出发点是围绕患者利益时，商业利益、学术探讨、科研经费和论文才能被纳入考虑范畴。那些致力于提高研究质量并有主见的患者和组织应该准备好自己的投石器，选择适宜的石块，瞄准，取得胜利。”

Refractor. Patients' choice: David and Goliath. Lancet 2001;358:768.

在诸多新兴机构中，Cochrane 协作网 (www.cochrane.org)，一个汇聚了对当前可得最佳治疗性证据进行系统评价的工作者的国际协作网，在 1993 年其创建之初就已经开始鼓励患者参与。2004 年成立的詹姆斯·林德联盟 (www.lindalliance.org) 就将患者、医护人员、临床医师召集在一起，来确定那些他们一致认为最重要的且尚未解决的疗效方面的问题，并根据紧要程度对这些问题排序。这类有关疗效不确定性的信息能让那些医疗卫生研究的资助者们了解哪些才是患者和医生真正关心的问题。⁸ 自 2008 年起，欧盟委员会资助了一项促进患者群体在临床试验中作用的项目，旨在将相关经验通过研讨会、报告书及其他交流方式在整个欧洲推广。⁹ 在其他国家，也对此类研究活动展开了积极广泛的讨论。

至关重要的合作关系

“英国国民保健署（NHS）所开展的关注人群的研究必须要有患者和公众的参与。

不论研究有多么复杂，抑或是研究者有多么出色，患者和公众总是能提供独特和弥足珍贵的见解。他们对研究的设计、实施和评估的建议常常使研究本身更加有效、可靠和划算。”

Professor Dame Sally Davies. Foreword to Staley K. Exploring impact:

public involvement in NHS, public health and social care research.

Eastleigh: INVOLVE, 2009. Available from: www.invo.org.uk.

不同的角色通过多种方式不断地被融入到研究中¹⁰，这使得患者及公众能和医疗卫生专家共同努力，而且新的合作方式也在不断地被完善（见下文 *消除患者和研究者之间的隔阂* 及第 13 章的第 2 点 *合理地设计和开展研究*）。¹¹ 这些在研究活动的方方面面都有所体现：

- 提出需要解决的问题
- 项目的设计，包括选择重要的结局指标
- 项目的管理
- 患者知情手册的制定
- 研究结果的分析 and 诠释
- 传播和实施指导治疗决策的结果

研究中患者的参与

研究中患者的参与是如何产生的呢？正如我们在第 3 章提到过的，之前强加到女性乳腺癌患者的过度治疗引起挑战和变革的过程，这不仅归功于新一代的临床研究者，当然还有患者的功劳。临床工作者和患者的协作保证了研究证据既符合严格科学的标准，又满足妇女的需求。当女性患者对根治性乳房切除术质疑时，她们所传递的信息是她们所关注的不仅仅是癌症的根除：她们要求参与制订有效的疾病治疗方案。

对于那些想作为合作研究者全程参与的患者和公众，存在几条可行的途径。例如，可以以个人身份或是作为某健康（或疾病）支持小组的一员参与研究，也可以参加像小组座谈会之类的促进会活动。先不论他们参与的种种途径，只要他们熟悉了研究方法学的特点要素，那么他们就能在和卫生专家合作时变得自信并有所作为。因此，他们将会要求获得和他们所扮演的角色相关的高质量信息和培训。我们将会在第 12 章继续解释为什么此类信息公布的方式（尤其是以数据的形式）对充分理解研究至关重要。当然也存在许多不那么突出的方式能使患者和公众投身于科学研究，只要我们能营造出一种协作的氛围来接纳患者的见解和建议。

今天那些踊跃参与研究的患者应该在回顾过去时对早期“患者先驱”的先驱性运动报以感激，这些先驱意识到自己应直言不讳并向现状发出挑战，而这一切是因为他们需要获得准确的信息。例如，在 20 世纪 70 年代早期，美国一小群由 Rose Kushner 领导的乳腺癌患者开始进行自我教育，这样自己才能有有用武之地，接着她们开始教育别人。Kushner 既是一名乳腺癌患者，又是一位自由撰稿人，她在 20 世纪 70 年代早期就开始挑战传统独裁的医患关系和根治性手术的必要性。¹² 在彻底研读了有关根治性乳房切除术疗效的证据之后她写了一本书。到 70 年代末，她的影响力和受关注程度使得她开始参与美国国家癌症研究院对新研究标书的评审。¹³ 与之相似的是，在英国，由于信息的匮乏，妇女们开始有所行动了。举

例来说，20 世纪 70 年代，Betty Westgate 成立了乳房切除术协会；20 世纪 80 年代，Vicky Clement-Jones 成立了 CancerBACUP 基金会（现在是 Macmillan 癌症基金会的一部分）。

外行帮助反思艾滋病（AIDS）

“在艾滋病领域里公信力斗争是多元化的：这其中包括了极其广泛的参与者。而‘外行’参与科研结论的公布和评价有助于形成对艾滋病的认识，正如他们使我们怀疑以往认知里的‘门外汉’是谁，‘专家’又是谁。现在到了一个考量专业知识结论或其代言人是否可信的紧要关头。但从更深层次上考虑，上述问题还涉及到公信力评价真正的机制：科研结论是如何评判的？由谁来决定？（正如这项研究显示，）科学领域内的争论同时也是有关科学以及科研应怎样开展或是谁应该开展的争论。”

Epstein S. *Impure science: AIDS, activism and the politics of knowledge*.

London: University of California Press, 1996.

20 世纪 80 年代后期，美国的人类免疫缺陷综合征或艾滋病患者对自身疾病有很好的认识。他们通过相关政策保护自己的权益不受研究机构的侵害，这为患者参与研究设计铺平了道路。最终当患者在研究中面对治疗选择时，这种参与为他们提供了选择的余地，多种灵活的设计也鼓励了更多人参与。这一模式在 20 世纪 90 年代早期被英国人所效仿。伦敦的 Chelsea 和 Westminster 医院所开展的研究就有艾滋病患者群体的参与，他们协助设计了研究。

14

这些艾滋病患者中的积极分子使得研究者端正了自身：一些研究者将有组织的患者群体视为灾难，但事实上这是向研究者对不确定性的解释所发出的正当挑战。而此前研究者的研究方法都忽视了患者所关注的结局。另一方面，患者能够渐渐体会到对新药效果草率作出评价的隐患以及在“有前景的”抗艾滋新药未经严格评估前就被要求投放市场的危害。研究者可能仅就新药“善意的投放”会在今后延长不确定性对患者所带来的痛苦为由，对这一方式提出异议。而患者的反对则源于他们认为这最终会加快患者自身和研究者认识到从容开展疗效对比评价的必要性，而且患者和研究者两大阵营应联合设计研究，同时顾及到双方的需求。

15

20 世纪 90 年代开展的一项艾滋病试验就很好地证明了研究中患者参与的重要性。当时齐多夫定这个药刚刚开始被用于艾滋病的治疗。其在疾病进展期患者身上的应用是有益效应的良好证据。那么，接下来讨论的问题显然就是在感染早期就使用齐多夫定是否能延缓病情进展并提高远期生存率。于是，为验证这个可能性，美国和欧洲都开展了试验。当发现一个可能的但仍不确定的有益效应时，美国的试验很早就停止了。尽管美国的试验已经得出结果，但在患者代表的参与和支持下，欧洲的试验继续进行直到得到明确的结果。试验的结论与之前认为的不同：在感染早期使用齐多夫定没有任何好处。这时候，药物仅有的作用就成了它令人讨厌的副作用。¹⁶

患者怎样破坏验证疗效的科学试验

患者参与研究并不总是对验证疗效的科学试验起到促进作用。2001 年，一项对研究人员的调研显示患者参与临床试验有时对研究确实起到了非常积极的作用，但它也暴露了一些非常现实的问题。这些问题的产生主要是因为人们对这种类型的协作缺乏经验。首先，患者

的参与会极大地推迟研究的启动。同时，人们也会对利益冲突和“代表性”产生顾虑，因为一些患者仍没有意识到不应该仅仅将他们自己的利益带到试验的管理会议上去。⁵

许多这样的问题似乎都归因于患者对研究如何开展及受到资助缺乏认识和理解。绝望的境地有时会促使患者孤注一掷，从而接受那些未经充分评估而且有可能弊大于利的治疗，对那些奄奄一息的患者也是如此。我们已经提到，通过游说患者和他们的律师，从而使得那些“有前景的”抗艾滋病新药被“善意地”投放到市场，这将会造成负面影响：它延迟了对患者非常重要的治疗结局的确定。而最近，个人和患者群体通过反面和误导性宣传，对多发性硬化症和乳腺癌药物的处方药产生了影响。

在 20 世纪 90 年代中期，干扰素在其疗效证据不足的基础上被用于治疗复发-缓解型多发性硬化症患者。很快，各型多发性硬化症的患者都要求使用这些昂贵的药物，而卫生保健服务机构也同意资助他们使用。干扰素也成为了这种衰弱性疾病公认的标准用药。因此，我们永远不会知道如何正确使用干扰素治疗多发性硬化症——此类研究从未开展而且为时已晚。然而，随着时间的推移有一项事实已经渐渐浮出水面——干扰素有着诸如“流感样”症状的恼人的副作用。

赫赛汀（曲妥珠单抗），正如我们在第 1 章，第 9 到 12 页所说明的，它并不是一个对所有患乳腺癌的妇女都适用的特效药。首先，其有效性取决于一个被特定基因修饰过的肿瘤，而只有 1/5 的乳腺癌患者才会携带该基因。另外，该药物对心脏有潜在的严重副作用。然而，患者的拥护助长了媒体的狂热，并导致政治家们也跟随舆论的大潮：带着缺乏对已有证据的考虑，以及对该药利弊平衡的进一步研究，赫赛汀的使用得到正式认可。

儿童消费力和新药

‘新药究其本质而言是不完整的产品，因为包括其安全性、有效性和包括成本在内的全部信息尚不能完全获取到。

值得注意的是，对“新药”的大力推崇并不是报纸所独有的，还常常可以在其他的媒体渠道、医学界和科学界中被轻易发现。

“儿童消费力”（儿童持续不断地缠着家长直到家长同意为他们购买）通常是和儿童广告有关的一个概念。而此处需要提出的问题是：我们是否也在见证着“患者消费力”或是此类直接面向患者的广告？这些广告使产品的认知度得到了提升，接着患者、慈善机构甚至临床医生开始要求此类产品投入使用。如果是这样，我们就需要进一步去了解是谁在推动这一类型的营销，他对医生和患者行为究竟会产生怎样的影响，以及这种行为是否为当前行为监管规范所允许。’

Wilson PM, Booth AM, Eastwood A et al. Deconstructing media coverage of trastuzumab (Herceptin): an analysis of national newspaper coverage.

Journal of the Royal Society of Medicine 2008;101:125-32

患者组织：他们的声音是否独立？

在患者组织和药厂的关系中还有另一个鲜为人知的利益冲突。大多数患者组织只有很少的资金，依赖于自愿捐助，而且只能获得很少的独立资金。来源于药厂的资金以及与其开展

联合项目可以帮助患者组织发展，使其变得更有影响力，但也会扭曲和不尊重患者的日常安排，包括他们的研究日程。这个问题的严重性难以衡量，但一项评估同欧洲药品管理局（EMA）合作的患者或消费者组织受公司赞助级别的调查为我们提供了全新的视角。该局协调着整个欧洲新药的评价和监管，值得称道的是，该机构积极组织患者和用户团体参与其监管活动。然而，通过在2006年至2008年对23个此类团体的调查发现，其中有15个组织从制药商或药厂机构接收过部分或大量的资金。此外，只有不到一半上述组织的向药品管理局如实汇报了所接受资金的来源或数量。¹⁷

市民参与改善卫生保健

“当不同阵营的利益汇聚时，出售医疗服务的一方和开具处方的那一方就会混合成一股巨大的影响力，几乎每次都会将决策者推向同一个方向：更多的试验、更多的程序、更多的床位、更多的药……”

由于十多年来在这个领域里只有个别人发出一些声音，我感觉到在讨论中常常缺少一种真正代表公众利益的声音。受赞助的宣传团队会急切地庆祝一种新的治疗或技术，而不愿公开指出其疗效有限、成本过高或是其本身就是一种危险因素。而且，像许多记者一样，政治家们常常会受到资深医学专家和激进的支持者不必要的惊吓——他们利用公信力来进行营销：拓宽疾病的定义，提倡最昂贵的治疗方案。

卫生保健领域中新的公民游说团体的出现深谙正确和错误使用科学证据的方式，因此这可能会引发关于开支优先次序更透明的辩论。这样的公民团体往往能揭露媒体的误导性营销，并为公众和政策制定者更全面地提供有关卫生政策的风险、利益、成本等方面真实而精准的评估。”

Moynihan R. Power to the people. BMJ 2011;342:d2002.

在某些情况下，患者组织是由药厂扶持的，以此来宣传他们的产品。例如，有一家制造干扰素的公司，为了使英国国民保健署批准干扰素用于治疗多发性硬化症（见上文），成立了一个名为“争取行动”的新的患者组织。^{18,19}该患者组织公开传达的所有信息都是干扰素有效但有些昂贵，而实际的问题在于该药物是否真正有效果。

消除患者和研究人员之间的隔阂

我们要警惕以上那些由于患者参与科学试验而引起的问题，以及他们可能会无意间破坏试验的方式。像大多数事情一样，良好的初衷并不能保证利大于弊的结果。不过，也有实例能很好地说明研究人员和患者合作能够改善研究相关性和设计。因此，许多研究人员会积极寻找能够协作的患者。

有一个例子能很好地反映协作筹备工作的价值，研究人员与患者及潜在患者探讨了一些在紧急情况下开展科学试验所遇到的棘手问题。如果想要对急性脑卒中的救治成功，那么这些治疗需要在脑卒中发作后尽可能早地实施。因为研究人员不确定实施的最佳方式，所以他们向患者和医护人员寻求帮助。他们召开了一次探讨会，邀请了一组患者和医疗卫生专家参与，并确定小组讨论中有老年人参与。最后，研究人员阐明了这项试验的计划，而且患者还帮助研究人员起草并修订了试验信息的宣传册。²⁰

这一全面深入的前期研究使得研究伦理委员会及时地支持一项随机(对照)试验的开展。核心小组的参与者已经意识到试图获得急性病患者的知情同意所面临的伦理方面的两难处境,这些急症可能会使他们思维不清或是无法沟通,即使这不是指患者没有意识。患者可以提出建议和解决方案,使研究的设计为各方都能接受,同时还大大完善了宣传册中的内容。

社会学家正以研究小组成员的身份,越来越多地参与到患者疾病敏感方面的正式探讨,这也完善了试验的实施过程。在对局限性前列腺癌的患者开展临床试验时,研究人员想比较以下三种完全不同的治疗方案的效果——手术、放疗或“观察等待”,而这无论是对于开展试验的临床医生还是对于犹豫是否参加试验的患者来说都极为困难。临床医生极不情愿去讲述“观察等待”这种方案,因而它被留在最后介绍,而且他们在介绍这种方案时缺乏信心,因为他们误以为被要求参加这项试验的男性患者会觉得这个选择无法接受。社会学家被要求去研究可接受性这个问题,进而协助确定试验是否真的可行。

社会学家研究的结果给了我们很好的启迪。²¹研究表明如果试验中的“观察等待”被描述为“主动监测”,那么它将会成为第三种能够接受的选择,如果医生在邀请患者参与试验时不把它留到最后描述,并且明确地以患者可以理解的术语介绍主动监测,就更容易接受了。

这项消除了医生和患者之间隔阂的研究,指出了那些给两大阵营带来困扰的特定难题,而这都可以通过对治疗方案更好的描述来解决。其中一项研究结果表明,随着时间的推移,被邀请参与试验的男性患者接受率由原来的40%增加到70%。这种快速的生长意味着如果准备工作尚未进行,对局限性前列腺癌患者的所有治疗方案的效果就不会像现在这么早的出现。而且,因为前列腺癌是一种常见病,未来将有许多人会从中受益,这比他们原来想象的要早。

协作前景一片大好

患者和公众能通过多种方式参与治疗性试验。正如我们前文所述及的,这些都可能会成为原动力——那些指出在理解方面存在差距的人,以及对新的实施方法的需要。研究人员可能会为他们的加入提供方便;他们可能会参与一项工作的某几个阶段而并非全部;他们可能会发现一种特定的不确定性,从那一刻起开始参与其中,而解决这一不确定性需要通过相应的宣传和实施,并将项目的研究结果整合到一篇会更新的系统评价中去;他们可能以不同的方式参与同一个项目。有时他们还会自己开始一项工作。这其中不存在硬性的规定:在某个特定的研究中,那些适当的策略和方法将被采用。正如上述局限性前列腺癌试验所证明的,方法时刻在被完善——即使是在一个项目的进行过程中。

当患者和研究人员一起工作时,他们为减少治疗的不确定性提供了一个强大的组合并使所有人得益于此。使得这种协作机制开展的各个方法均适合于相应的单个研究,在国家级研究机构的支持下,未来的前景将一片大好。

要点

- 患者和研究人员合作有助于发现并减少治疗的不确定性
- 患者的参与能够提高研究质量
- 有时患者会无意中破坏验证疗效的科学试验
- 患者组织和药厂的关系可以歪曲治疗效果的信息

- 想要做出有效贡献，患者需更好地了解研究的基础知识，才能可容易获得无偏倚的信息
- 没有一个“正确的方法”来实现研究中的协同参与
- 患者参与应基于特定的研究目的
- 患者参与研究的方法学在不断更新

第 12 章 那么，什么让卫生保健变得更好？

在前几章，我们举了很多例子来说明为什么治疗应该建立在可靠的研究之上，同时这些研究的设计目的须是为解决对患者意义重大的问题。无论我们是普通公众、患者还是医疗卫生专家，治疗效果都会或多或少地影响到我们的生活。验证疗效的科学试验产生的强有力的证据确实很重要。

在这一章里，我们一起来看看这些证据会怎样通过影响卫生保健实践来促成医患双方共同做出对个体的治疗决策。明智的决策需要基于高质量的证据，还应能够告诉我们不同治疗方案可能产生的后果。不过，对于不同个体这些后果的意义与价值也不尽相同。因此，即便是使用相同的证据，不同患者所作出的决策也可能不同。举例来说，较之于常人，功能健全的手指对于音乐家，灵敏的嗅觉对于厨师，敏锐的视觉对于摄影师，往往会更为重要。因此，他们也许会愿意为这些他们视若珍宝的东西做出更大的努力或是承担更高的风险。证据和决策的关系相当复杂。因此，你可能会发现，本章的大部分内容都在回答如何处理这种关系的一些常见问题。

但在此之前，我们还需要更深入地思考“共同决策”，并指出在实践中的体现形式。共同决策的方式决定于以下两者各自的决定，即家长作风的专业人员和被遗弃的患者。患者常常抱怨的是自己获得的信息太少，其实，他们对自己想要承担的责任有着不同的预期。^{1,2}

共同决策

“共同决策的定义是：‘患者参与临床决策的过程。’它的理念在于，专业人员的职责在于（应该）足够清晰、坦诚地向患者说明问题，患者从而能够明白存在于大多数医疗决策中的不确定性，并最终接受必须在矛盾的选项中作出抉择这一事实。医生的专业知识在于根据疾病的轻重缓急，作出合理的诊断并提出治疗方案。而患者扮演的角色则是提出自己的价值观和个人意愿（这往往和他们所处的社会环境有关）。”

Adapted from Thornton H. Evidence-based healthcare. What roles for patients? In: Edwards A, Elwyn G, eds. *Shared decision-making in health care. Achieving evidence-based patient choice*. Second edition. Oxford: Oxford University Press, 2009, p39.

有些患者并不乐于对自己的疾病和治疗方案了解得太多，希望完全听从专业医生的建议，但也有很多患者热衷于了解更多的信息。对于这些想了解更多信息的患者来说，我们应确保他们能获得好的阅读材料，并且有经验丰富的医生能够指导他们去哪里和以什么样的方式去获取此类有用的信息。

不同人对什么叫做“理想的咨询”有着相当不同的理解。有些人希望独立探索，而有些人则希望得到指导。患者在医生鼓励下更好地参与决策，或许会成为最有效的方式之一；而且一旦患者尝到甜头，这种方式便会成为患者的首选。就像我们接下来将要提到的，患者一个简单的问题就有可能打开两者的话匣子。更加重要的是，无论何种程度的参与，患者得到公平对待之后，往往会觉得自己真正参与了治疗。

本书的一些观点在你眼中是什么样的？

尽管没有哪两项咨询完全相同，但对如何尽可能做出最佳决策的指导原则都是一样的，我们在本书开头也提到过。我们最终要达到的目标是：在当前最佳相关证据的指导下，医患双方都对他们共同参与的咨询过程感到满意。患者咨询医生各种各样的健康问题：有些是短期的，有些是长期的，有些性命攸关，有些只是“小毛病”。他们的个人情况千差万别，但他们的問題都需要得到解决，因为这最终关系到他们的决策。

为更好地说明这一点，我们首先以一种常见疾病膝关节骨关节炎（又叫“磨损性”关节炎）的医患讨论为例。接下来我们会解答一些有关使用研究证据指导临床实践的基础问题，这些问题是那些情况各异的患者在咨询医疗专业人士时希望得到回答的，也是阅读了本书前几章的读者可能提出的。



医患沟通以及一些可能的问题

共同决策：常见问题的咨询

医生：是这样的，现在您患有中度膝关节骨关节炎，这是中老年人的常见病。我们通常称它为“磨损性”关节炎。这种疾病的症状通常很不稳定，时好时坏，但总的来说，随着发病时间的延长，疾病会缓慢进展。最近您感觉怎么样？

患者：额，最近如果我干活稍稍用力，我的膝关节就很疼，这种状况可能持续几个小时，有时候我还会因为这个难以入睡。最近，情况好像越来越糟了，我担心可能需要做膝关节置换术了。

医生：膝关节置换术当然是治疗此病的选择之一，但是我们通常只是在其他相对简单的治疗方法都无法奏效，万不得已的情况下才用它。

患者：那您的建议是……

医生：一般的镇痛剂或者是抗炎药可能会缓解病情。除药物治疗外，一些特殊的锻炼也可以加强膝关节周围的肌肉，并有助于保持膝关节的功能和减轻您的疼痛。您想了解更多这方面的信息吗？

患者：吃药会让我的胃不舒服，我还是想听听关于锻炼方面的知识。

医生：好，我给您一些关于如何进行锻炼的手册，也可以介绍您去找我们的理疗师。在这段时间里，如果疼痛持续明显的话，您可以放心规律地服用一些对乙酰氨基酚。

患者：这对我很有帮助，您还有其他的治疗方法吗？

医生：对于严重的骨关节炎我们确实还有更进一步的方法。但对于您这种情况来说，如果您能够做到我提出的几点的话，相信您很快就会发现随着锻炼的加强，您的肌肉会更有力量，症状会慢慢改善。随着您疼痛减轻，睡眠也会改善，然后慢慢地您就可以做更多事情了。您也可以考虑去游泳或者溜狗，这样不仅起到锻炼的作用，还能让您心情愉悦，而且还可以帮助您减肥。我想我们可以暂不考虑更激进的治疗选择，先看看您的锻炼效果和疼痛的变化。但如果您觉得病情加重，随时再来找我。

将研究证据应用于实践时遇到的问题

问题一：对于一个生命垂危的患者，难道不应该尝试所有值得一试的方法吗？

在这种情况下，新出的“特效药”或者是经过一些知名人士在大众媒体上高调宣传的治疗方案都会显得十分诱人，这其中可能就包含在市场上热销但却未被验证过的“替代”药物。主流的治疗方案在这个时候可能显得并没有那么吸引人，前景也不被看好，但是，大多性命攸关的情况下使用的药物都是在经过千锤百炼后其有效性和安全性才得以证实。因此，一开始就寻找最佳证据不仅可以节省大量时间和金钱，还可避免悲剧的发生。

一般情况下，主流医学已经明确了其所提供药物的有效性和安全性存在不同程度的不确定性，并通过试验的方式将这种不确定性降到可以接受的水平。此外，还通过不断系统地评价相关证据来改善现有治疗方案。然而，只有患者意识到这是唯一能让治疗改善的途径后，才能保证其稳步发展。

我们可以理解，患者在性命攸关时，会极度渴望尝试任何治疗方案，包括那些未经验证的‘治疗方案’。但让患者考虑参加一些旨在比较一种新的治疗和现有最佳治疗的临床试验，对他们而言会更好。这种比较不仅会反映新治疗带来的益处，还会发现其可能造成的危害。在性命攸关的情况下需要的治疗力度非常大，而且任何治疗都有一些副作用。因此，对一项新的治疗进行彻底而充分的试验显得非常重要，只有这样才能系统地记录下结果，来判断其是否能为患者带来帮助。

问题二：虽然患者想知道治疗方案是否‘有效’，但假如他们不想知道所有细节呢？

为了让患者在充分知情的情况下作出选择，我们需要在过度提供信息和剥夺患者知情权之间找到平衡点。我们同样应牢记，患者也需要在患病之初和随后权衡利弊进行抉择时获得信息。一个理想的咨询过程中，医患双方都应该感到相当满意，因为患者掌握了足够让他们马上作出决定的相关信息，而医生也将当前最好的处理方案告诉了患者。但问题并非到此为止，如果患者经过更多的思考后想了解更多细节，在这个时候，医生应该能够协助患者提出问题，并进一步理清思路。

有些选择会涉及到艰难的取舍；我们最终不得不“两害取其轻”。举个例子，在第4章，我们讲到了主动脉瘤，它会和心脏相连的主动脉扩张，而主动脉瘤破裂则是致命的。目前我们主要采取手术来纠正这个问题，但患者会有1%~2%的几率死于手术。所以我们就需要在手术致死率和主动脉瘤破裂致死率之间进行权衡。从长远的角度来看，采取手术无疑是最好的方法，但一些患者可能并不愿意接受手术，或者有些患者想将手术延期，比如在她们女儿的婚礼结束之后。所以说，我们并不需要盲从一些“唯一有希望的”方法，我们更要学会权衡风险并把握时机。

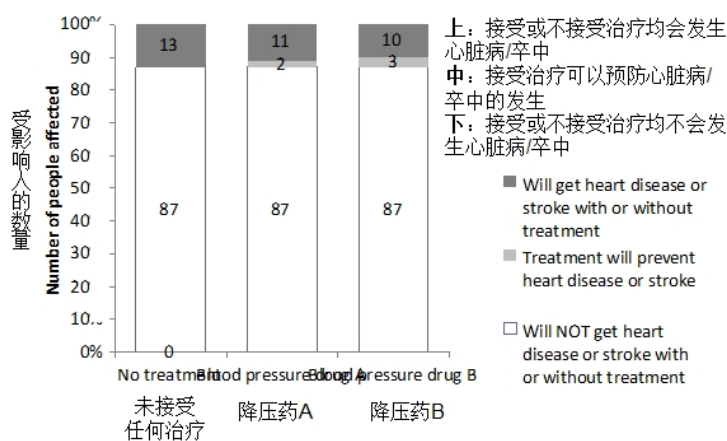
问题三：统计数字让人一头雾水——我们的患者是否必须要了解这些数字？

数字所呈现出的信息可能会让人望而却步——甚至会彻底误导人们。但如果你真的想要对比两种不同的治疗方案，或是你想了解和自己情况一样的人是如何受到疾病影响的，数字总能从某些层面上提供一些帮助。但一些数字的呈现方式确实比另外一些的要好。

让外行（甚至是医生）能够了解数字的含义的最佳方法就是采用百分比来表示。这就意味着我们要采用一些整数。例如，100个人当中有15个人患病我们就说是15%。除了用文字来表述数值，我们还可以采用某些类型的图表来表示。比如彩色直方图、饼图，或方框里的圆脸、笑脸和哭脸等，当然还有图表。通过这种“决策辅助”的形式来呈现数字可以帮助更多的人了解数字的含义。

下面我们用直方图来解释降压药在接下来的十年里对有高血压病史的患者发生心脏病和卒中风险的影响。³

在接下来的十年里，100个和你类似的患者的病情会如何？



在十年里，每100个有高血压但不接受治疗的患者，会有13个人得心脏病或卒中。如果100个人都吃了降压药A，那么只有11个人会得心脏疾病或卒中，另外2人可避免患病。如果所有的100个人都吃了降压药B，只有10个人会得心脏病或卒中，另外3个可避免患病。这样我们就可以明白了。但现在情况是，即使是这样简单的数据，也只有我们的统计学家才能看明白。

接下来我们再看看通过表格而非直方图来呈现的数据。在这里，我们将重点放在相对较好的药物上——药物 B：

我们逐个将数值转换成自然频率（简单计数），可以得到如下的表格：

	不吃药	吃 B 药
发生心脏病或卒中（10 年内）	每 100 个人中有 13 个人	每 100 个人中有 10 个人
没有发生心脏病或卒中	每 100 个人中有 87 个人	每 100 个人中有 90 个人
总计	100	100

不吃药，得心脏病/卒中的概率是 13%（每 100 个人中有 13 个人），但吃 B 药得病的概率为 10%（每 100 个人中有 10 个人）——相差 3%（每 100 个人相差 3 个人）。因此吃 B 药能让每 13 例心脏病患者中的 3 例不会发生心脏病或卒中。也就是说，相对危险度降低率为 3/13，或者 23%。换句话说，治疗的**绝对**危险度降低率为 3%，但**相对**危险度降低率为 23%。这是相同数据的两种不同表达方式。

相对危险度降低率通常数值更大，有时候可能还会高出很多，这样我们就可以吸引更多的注意力。因此如果你看到标题说“卒中风险降低率为 23%”，这样的数据毫无意义。因为并没有说明它的试验人群是谁，也没有说明试验的时间范围。最重要的是，它没有指出不接受治疗患上卒中的风险。这种数据很有可能就是相对危险度降低率（但只有验证后才知道）。

有时候数值会有很大差异。对一项前列腺癌筛查，看看报纸表达数据的方式，“降低高达 20% 的病死率”看起来很不错。而这项结果也可以说“每筛查 1410 个人，能够防止 1 个人死亡”，当然我们还可以说成相当小的 0.07%。这意味着每一万个接受筛查的人当中，只有 7 个人能防止过早死亡。在这里，20%是相对危险度降低率，0.07%是绝对危险度降低率。但由于前列腺癌病死率极低，后者比前者要低得多——这样的数值不太可能作为报纸的标题来吸引读者的眼球。值得我们注意的是，如果一个标题的数值看起来相当的乐观，那么它很有可能就是这种情况。⁴

别被吸引眼球的数值愚弄！

“如果你胆固醇过高，那么你在 50 岁时得心脏病的概率就比一般人高 50%，这听起来很糟糕。我们换一种说法，如果你胆固醇过高，那么你得心脏病的风险就增高了 2%。这个听起来还行。但事实上两者是一样的（以上为假设的数据）。让我们再换一种说法，每 100 个胆固醇水平正常的患者，在他们 50 岁的时候，会有 4 个人得心脏病；但每 100 个胆固醇水平过高的患者，在 50 岁时，就有 6 个人得心脏病。这样一看，它们确实只相差两个百分点。”

Goldacre B. *Bad Science*. London: Fourth Estate 2008, pp239-40.

所以说数字真的很重要，正确地表达它们能够帮助人们作出决策。病人应该毫不犹豫地要求医生以更清晰易懂的方式解释数字的含义（如有需要，也可以采用一些更加形象的表述方式）。如果治疗的决策由医患双方共同作出，两者都应该很清楚每一个数字代表的真正含义。

问题四：患者如何知道研究证据对他们适用？

所有的决策都参考之前类似的经验，无论是一条经验还是多条经验。类似于随机对照试验的科学试验就是一种通过设计将偏倚最小化的好经验。无论好坏，总是难以避免这样的不确定性：之前的经验能够在多大程度被用于治疗下一名患者？所以，如果科学试验中所研究的患者情况相似，所处阶段或严重程度也相似的话，那么对个体最有可能的推测就是他会出现和其他人一样的结局，除非有确切的原因认为他和别人的情况不一样。

当然，即使数据能够用到患者身上，他们或许还会担心：“由于个体差异，自己对同样的治疗，反应是否会不一样？”治疗措施的“科学试验”只会告诉我们大体的情况，但不能很好地保证它对所有人的效果都一样，当然也不可能预测出谁最终会出现不良反应。研究证据会指导我们得到可能的最佳治疗，之后在患者身上使用。就像治疗皮疹一样，循证的治疗可能需要将药物用于体表某一部位皮肤，再拿另外一部位作参照（见第6章，第74页）。通过对比两个部位的不同反应，医生和患者都会看到治疗是否有效，或者是否出现了不良反应。事实上，在第一次接受某种皮肤治疗的时候，我们经常会使用这种“局部测试”的方法进行试验，就像我们治疗面部粉刺一样。

但大多数情况下，我们并不可能直接做比较。对于一些慢性病和没有威胁到人身安全的问题，比如说疼痛或瘙痒，我们可以在同一个病人身上反复试验。这种方式也叫做“单病例随机对照试验”，意思是病人的个数（N）实际上只有一个。在这样的例子里，我们在第6章曾提到的对照可比性的原则并没有改变，其中包括无偏倚的盲法结局评价等。在理想情况下，我们会用安慰剂对照来试验皮肤治疗方案或药物的疗效，但这往往很难操作。

然而在很多情况下，我们不能“先试一试，再看看效果怎么样”，因为此类结局往往所需周期过长或不确定性过高。比如，我们不可能在一名患者出现卒中后再验证阿司匹林是否能预防卒中。这就给大多数预防型药物和许多急症的治疗提出了问题，如脑膜炎、肺炎或者蛇咬伤。在这些情况下，我们根本就不可能先看效果如何才做出决定。所以我们必须考虑的是能否和如何借鉴其他研究的经验。

在实际情况中，如果我们对证据所提供的信息感到满意，我们就需要进一步提出这样的问题：患者的病情（或者健康人群的预期风险水平）跟研究纳入人群的相似度有多高？这很重要。因为在一般情况下，病情较为严重的一方在治疗上获益更大。因此，如果两者病情一样，或实际比研究中的病情更严重，我们通常会对证据的实用性更有把握。反之，如果实际病情较轻（或健康人群的预期风险较低），那我们需要考虑的关键问题则是，是否值得去获取比研究显示的要小的结果。

问题五：“基因检测”和“个性化医疗”中医生可根据需要针对每个人制定具体的治疗方案，这意味着我们就不需要其他方案了？

尽管个性化治疗听起来相当诱人，且有可能应用于某些情况，但如果说我们要把这种方式作为主要的治疗方式，几乎不可能。正如我们在第4章（见第43-44页）讨论基因检测的时候提到的，大多数疾病并不仅仅是几个基因简单相互作用的结果，而是基因与环境因素更为复杂的相互作用的结果。

基因分析能够帮助一些患有遗传性疾病的家庭或个人做出权衡，如亨廷顿病、地中海

贫血（遗传性血液病）和一些其他疾病（大多数都是一些罕见病）。基因信息确实给上述疾病的家庭带来了福音。但对于我们所关心的常见疾病，基因分析在家族史和临床检查的基础上能提供的更多信息微乎其微。尽管在未来，这种情况可能会有所改善，但是就目前而言，我们不必太相信所谓的对通过基因检测就能预测常见疾病的危险。

需要声明的是本书的作者并没有哪一个人曾经建立基因档案，我们也没有这方面的打算。因此，除非发生以下两种情况，否则，当我们不建议基因检测的时候，大家无需感到惊讶：(1)家族史提示可能患有已知的遗传疾病；(2)或在少数已知疾病中，通过一种或者几种基因的检测，我们可以知道应该如何治疗它。

问题六：如果病人所患的疾病正在进行一些临床试验，而医生刚好又不知道，他们如何才能找到这些临床试验？（见附加资源）

正在就诊的患者被纳入临床研究的概率小于百分之一，这一比例会因病种和环境大幅波动。甚至在癌症研究中心（那里频繁进行各类试验），这一幅度依旧很大。大多数患癌症的儿童会参与试验，但在成年人中，这个概率不到十分之一。大多数临床试验中，患者都是在临床中心进行登记。如果一个中心没有参与试验，那么就不能招募患者。因此，患者要找那些能够参与临床试验的中心。仅有少数在社区进行的试验患者可以直接参与。这些试验通常研究的是如何解决患者的精神疾患，诸如抑郁症或焦虑症。最近，有些试验通过互联网的方式让人们直接参与。最近有一项研究评价了在运动之前先做拉伸训练的效果，就是以这种方式让患者登记：这些人可能甚至都没有进诊所，却已经通过互联网获得了所有的指导并接受了随访。

如果医生在登记患者进入试验时显得不那么情愿，患者就需要寻找原因了。有可能患者真的不太合适，也有可能仅仅因为医生想要推卸由繁琐的监管规定强加给他们的额外工作（见第9章）。那些确信他们可以参加临床试验的患者应该坚持他们的想法。如果确实有合适的临床试验，并且患者也想参与进来，那么医生应该给予支持。

问题七：辨别证据是否可信的最佳方法（通过互联网或是其他途径）是什么？应该注意什么问题？

遗憾的是，不存在完全可靠而又简单的标识来辨认可靠的信息。如果你不准备查找原始研究，你就只能参考他人的意见。因此，对他人（或其他机构）的能力做出正确评价很重要，并且还要注意是否牵涉到利益冲突（是否别有用心）。如果没有，那么你就要问问自己是否相信他们已经发现并且评价了最好的研究：是否进行了详细说明并参考了已发表的研究？

举个例子，有一个人他想知道β-胡萝卜素（类似于维生素A）是否有抗癌作用。以“β-胡萝卜素和癌症”作为关键词通过Google检索，有超过八十万条的结果。其中，前十条里，有四条原始研究结果，六条综述或专家意见，而在这六条里，又有三条在同一页面中为维生素或其替代药物做宣传。看来情况真的不容乐观！

甚至还有一个网站提到：

“问：β-胡萝卜素能否预防癌症？”

答：研究表明，β-胡萝卜素能降低患癌症的风险。黄色、红色和深绿色蔬菜中含有β-胡萝卜素，水果中也含有。毫无疑问，服用β-胡萝卜素替代物和食用含有它的蔬菜水果效果

一样。但这也有风险。研究表明，服用β-胡萝卜素的人患肺癌风险增加。”

除了宣传外，他们还经常提到“研究”，却未见其更详细的说明或列出已发表研究的参考文献——这是关键所在。对于作者是否进行检索，是否评价“研究结果”，还是仅仅参考了他偶然看到的结果顺眼的文章，我们都无从考究。

让我们来对比一下“维基百科”前十个检索结果：

“Cochrane 协作网于2007年发表于《美国医学会杂志》(JAMA)的一篇文章对所有科学期刊的随机对照试验进行了评价，发现β-胡萝卜素增加了1~8%的死亡率(相对危险度:1.05; 95%可信区间: 1.01-1.08)^[15]。但这篇Meta分析纳入了两项关于吸烟者的大型研究，所以结果是否适用于一般人群还有待验证。”^[16]

以上文字说明了证据类型(随机对照试验)，并且给出了参考文章(方括号里边的数字)。所以，没有广告而有证据的详细信息，让人更信服。

问题八：有哪些可信的信息来源值得推荐？（见附加资源）

当前还没有一个包含了所有疾病和所有治疗方案的信息库。患者可以根据本书的几个原则，形成自己的检索技巧。比如，除了本书第6至8章提到的方法，《聪明的健康选择》(Smart Health Choices)⁵这本书也列举了几个小技巧，包括如何发现好的证据，应该注意哪些问题。

目前在网，很少有证据是基于系统评价得出的。可供选择的有包含了“大白话摘要”的Cochrane系统评价数据库(www.Cochrane.org/Cochrane-reviews)和德国药物评审机构(IQWiG)网站(德语，但在www.informedhealthonline.org有英语译本)。还有很多网站并不是基于当前最佳证据的系统评价，但提供了高质量的信息，比如NHS Choices(www.nhs.uk)以及PubMed(www.pubmed.gov/health)都提供了高质量证据。

当然，我们也需要时刻提高警惕。特别是需要提防利益冲突，如那些通过诱导人们相信某些信息从而赚钱，或想推销某种商品的网站。这些情况相当隐蔽，正如我们在第11章提到的，有些患者团体悄悄地接受药厂的资助，并提供错误的信息。

问题九：如何避免“被鉴定”为“患者”并接受不必要的治疗？

医学在过去取得了长足的进步：用于预防和治疗传染病的疫苗以及抗生素、关节置换术、白内障手术、儿童癌症治疗等，不胜枚举。但这种成就也使得医学滥用在其并不适用的地方。对于一个举着锤子的人，他的眼里只有钉子；而对于一个掌握新疗法的医生（甚至药厂），则会认为哪里都有毛病。比如，当治疗糖尿病和高血压的新疗法问世时，医生立马会建议他的患者采用这种新方法，尽管这些患者只是稍感不适。这样一来，被贴上“糖尿病”或“高血压”的标签的人就会剧增，那些原本正常的人群被“医疗化”对待。

除了治疗的不良反应（有时候由不必要的治疗引起），这种“贴标签”的做法也会造成心理和社会层面的问题，它破坏了许多人对健康状态的感觉，给工作和医疗保险业造成麻烦。因此，认识到这一连串后果，对患者和公众都相当重要。在匆匆忙忙同意接受治疗之前，必须停下脚步进行思考，权衡利弊得失。正如我们曾在第4章谈到的，筛查这种方式很有可能

造成过度诊断的问题，并把一个正常人贴上标签，最终还有可能导致过度治疗。

对此，我们的第一道防线就是警惕“标签”的存在并做进一步调查。类似于“一个健康的人就是那些没有被充分检查的人”，这种看似轻率的论点其实蕴含深意。因此，探究疾病风险高低的做法总是明智的。就像我们在之前谈到的，患者应该向医生询问如果对疾病采取观察治疗会有什么后果，病情如何监测，何时采取措施。当了解患者不想马上接受治疗或试验时，有些医生可能舒了一口气。但有些医生可能就会陷入“贴标签”的陷阱里：标签=疾病=强制治疗。他们并没有意识到其实患者有可能更希望先观望情况，再作出决定。

谁才有糖尿病？

我们怎样才能判断谁有糖尿病呢？当我还是一名医学生的时候，我们的数值是这样规定的：如果你的空腹血糖值超过 140，你就会被诊断为糖尿病。但在 1997 年，糖尿病诊断与分级专家委员会重新定义了这个疾病，他们把数值改为 126。过去，空腹血糖值在 126~140 的人是正常的，但在现在就是糖尿病患者。这个微调直接导致糖尿病患者剧增了一百六十万。

这会造成问题吗？谁也说不上。因为我们修改了标准，现在需要治疗更多的患者。这意味着我们可能会降低这些新患者的糖尿病并发症风险。但也因为这些患者的病情更轻（血糖值相对更低，为 126~140），他们本来就具有较低的并发症风险。

Welch HG, Schwartz LM, Woloshin S. *Overdiagnosed: making people sick in the pursuit of health*. Boston: Beacon Press, 2011: p17-18.

我们该何去何从？

我们以上谈到的所有问题，包括不同患者的意愿和价值观，如何理解统计数字，如何运用数字，以及如何采用有效的治疗来应对日益严重的病情，全都建立在医患双方以及卫生部门以及它所服务的民众之间的良好沟通之上。因此，在这里，我们想用共同决策的“萨尔茨堡宣言”来结束本章，它列举了如何让不同人群共同协作的几点建议。^{6,7}

共同决策的萨尔茨堡宣言

我们希望临床医生能够做到：

- 意识到他们有义务与患者共享重要的决策
- 促进互相交流，鼓励患者提问，帮助患者认识自身情况，让患者表达自己意愿
- 在告知风险时，提供与最佳临床实践一致的不同治疗方案的准确信息，并说明不确定性和治疗的利弊
- 根据患者的实际需要，提供相关信息，并给患者足够的时间思考以便理解信息
- 提醒患者，大多数情况下都不需要马上做决定，为患者及其家属提供资料，以帮助他们做决定

我们希望医生、研究人员、编辑、记者和其他人能够做到：

- 确保提供清晰的、循证的、最新的以及声明了利益冲突的信息

我们希望患者能够做到：

- 提出他们的观点、问题以及他们所关心的事情
- 认识到他们在对自己的治疗上有平等的参与权
- 寻找并运用高质量卫生保健证据

我们希望决策者能够做到：

- 采纳鼓励共同决策的政策，包括考核指标，以促进其提高和改善
- 修正知情同意法以促进共同决策技巧和工具的发展

我们提出以上倡议，是基于以下几点现状：

- 如今，大多数患者接受的治疗都受限于临床医生的能力和知识面，而不是受到广泛认可的最佳临床实践标准或是患者对治疗的意愿
- 医生总是迟迟不能意识到以下问题：患者很希望能够和他们一起了解自己的健康状况和当前可用的治疗方案，以及在制定决策时考虑到他们个人的意愿
- 许多患者及其家属发现自己很难参与到卫生保健的决策中来。有些患者缺乏质疑医疗专家的勇气。许多人对健康问题及其关键因素了解有限，也不知道该去什么地方寻找清晰、可靠且易于理解的证据

第 13 章 合理的研究：未来更美好的蓝图

毫无疑问，医学研究在提高生活质量和延长人类寿命方面功不可没。然而，正如本书所述，当前商界和学术界作为研究的推动者，在确定和处理患者的优先顺序上远没做好。

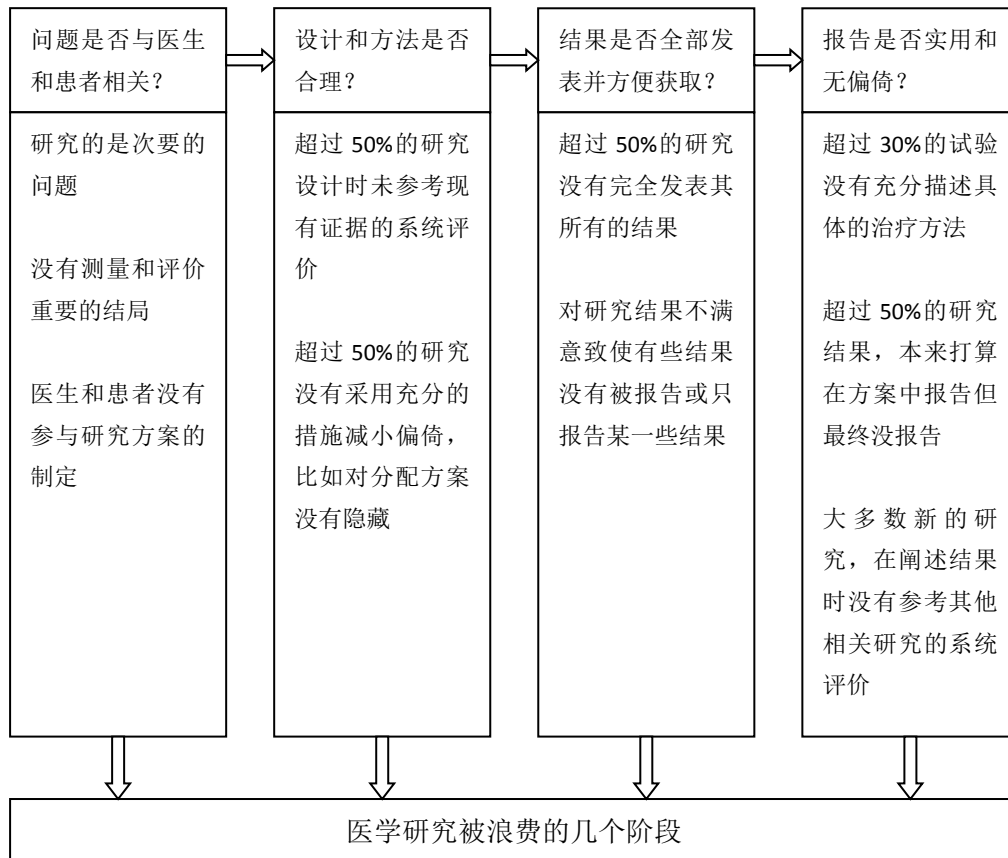
全球每年有超过一千亿美元的巨额资金用于医学研究¹，但其中很大一部分都被投入到了实验室研究和动物研究，而不是那些可能得出与患者更直接相关证据的临床研究。

当要决定研究那些有关疗效的问题时，患者认为的优先顺序常被忽略。药厂雄厚的财力对研究问题的决策影响很大。因为药厂会慷慨地对临床试验招募的每位患者和经常参与试验的学术研究者及其工作单位支付报酬（成千上万英镑或美元），但这些试验关注的问题仅关乎药厂利益，而非关乎患者利益。

令人遗憾的是，投入医学研究的大量资金被浪费在以下连续的阶段：提出错误的研究问题；开展不必要的或低质量（设计拙劣）的研究；研究结果未能全部发表和获取；产出了有偏倚和没用的报告。这应该与每个人都息息相关——研究人员、研究资助方、临床医生、纳税人，尤其是患者。

在展望未来更美好的蓝图之前，我们简要列出了更好的研究所应具备的关键要素：

1. 提出正确的研究问题
2. 合理设计并实施研究
3. 发表所有结果并方便获取
4. 产生实用无偏倚的研究报告



1. 提出正确的研究问题

有时候医生并不知道哪种治疗措施可能对患者效果最好，因为现有的治疗方案未被正确研究。那些可能会对患者的治疗产生关键作用的研究，由于对药厂或学术界可能没有或只有很小的利益，所以重要的问题仍未得到解决，正因如此才造成巨大的危害。举这样一个例子，用类固醇治疗脑损伤导致的残疾患者是否会提高或降低生存率。一项设计良好的试验表明类固醇治疗可能导致了数以千计的脑损伤患者死亡，但在此之前类固醇已经被用了几十年。²但为何这项研究的提案一开始就遭到了药厂和一些大学研究人员的反对呢？他们之前参与评估了一种对患者而言，测量指标并不是那么重要的昂贵新药（所谓的神经保护剂）疗效的商业试验，所以他们不希望看到有其他研究来争夺受试者。

我们着力解决这些问题的另一原因是想有助于确保现有的宝贵卫生保健资源不被浪费。在 20 世纪 40 年代，人血白蛋白溶液静脉滴注被用于烧伤患者和其他危重病患者的复苏，理论上看来它可以降低患者病死率。令人讶异的是直到 20 世纪 90 年代，这个理论都未得到科学试验的验证。正值此时，一项对相关随机（对照）试验的系统评价指出，没有证据表明与普通生理盐水相比人血白蛋白溶液可以降低病死率。而系统评价显示，实际上非要说人血白蛋白对病死率的影响，那便是增加病死率。³这篇系统评价的结果促使澳大利亚和新西兰的医生们联合起来，就人血白蛋白溶液和生理盐水（一种可替代的复苏溶液）的疗效进行首次大规模的随机对照试验。⁴而这项本该在半个世纪前就应该做的试验并没有发现人血白蛋白溶液优于生理盐水的证据。由于白蛋白比生理盐水贵 20 倍，可想而知在过去的 50 年里甚至更久，全球卫生医疗支出中为此浪费了巨额资金。

2. 恰当地设计和实施研究

调查发现许多临床试验报告质量差，这促进了临床试验报告标准的制定与应用。该标准明确了受邀参与研究的患者数和拒绝参与的人数。从试验一开始，各治疗组的所有结果都要报告。但仍有许多问题需要完善：(a)研究要解决问题的选择；(b)如何构建问题以确保要评价的治疗结局是患者非常关心的；(c)如何使患者方便获取信息。（见第 11、12 章）

为了解试验方案的可行性和可接受性，有患者团体参与的一些探索性工作将会很有用。这样做会发现方案设计的缺点，有助于确定相关性更强的结局指标，甚至有可能发现试验根本无法开展。^{5,6}

这样做不仅可节省大量的时间和金钱，还能避免失败。正如我们在第 11 章描述的（140-141 页）局限性前列腺癌患者的临床试验，它就阐述了医生是如何通过斟字酌句来说明试验目的和治疗方案，从而使得研究设计得到改善。了解患者观点会使研究易被接受，因为受试者的担忧和信息需求得到了认同，并且向潜在参与者提供的信息也考虑了这些发现。

7

3. 发表所有结果并使其方便获取

研究结果的选择性报告会引入严重的偏倚。一些“阴性”研究当其结果不符合调查者或资助方的预期时，它们就永远不会被发表。没有哪一篇已发表的报告会提及这些问题，这些试验会消失无踪。⁸不仅如此，已发表的某些试验中报告的结果也有可能是被选择过的，也就是说部分结果因为在试验过程中发现不是“阳性”而被排除。⁹患者已遭这种有偏倚报告试验的残害甚至致死。这种行为既不伦理，也不科学。

4. 产生实用无偏倚的报告

即使是已发表的研究，都常忽略那些可供读者评价和应用结果的重要部分。2000 年 12 月，一篇对发表在著名期刊上的 519 个随机（对照）试验的评价发现，其中 82% 没有描述隐蔽分组方案，52% 没有提供减少观察者偏倚方法的细节，如我们在第 6 章曾提及的，这两项特点对好的研究¹⁰至关重要。更有甚者，就连治疗措施的描述也未能逃脱这种粗陋的报告。一项试验显示，通过分发专门的手册（对比没有手册）可以帮助肠易激综合症患者，但该试验并未提及册子的内容或通过何种方式获取。因此，这种“治疗措施”实际上并不能为其他患者或医生所用。对主要期刊的试验分析发现约 1/3 缺少这类关键信息¹¹，而这仅仅是例子之一。

最后，大多数已发表的试验并没有将它们的结果建立在先前类似的试验上。正如我们在第 8 章提到的，没有这关键的一步就不可能知道结果真正意味着什么。在 1997-2009 年的 12 年里，我们对五种主要医学期刊发表的随机（对照）试验进行每四年一次的核查，揭示了问题的所在。总的来说，94 篇研究中只有 25 篇（27%）参考了相似试验的系统评价；只有 3 篇真正包含了更新的评价，它们整合了最新的研究结果并指出其对总体证据的影响。但令人沮丧的是，随着时间的流逝，报告质量的改善与否依旧无据可循。¹²这会使得医生使用哪种疗法将取决于他们碰巧阅读到的期刊。

未来更美好的蓝图

医学研究可以因正当理由而开展，可以很好地完成，可以被很好地报告。就个别而言，下面提供的建议没有哪一个是新的；但如果将其组合，并由患者和医生联合实施，我们的八个行动指南会构建一幅科学验证和应用治疗的未来更美好的蓝图。

1.增加关于如何判断疗效可靠性的常识

情形发生了改变，即越来越多的公众意识到偏倚和偶然因素会通过不同方式严重扭曲有关疗效的证据。但科学研究最重要的特点之一——辨别偏倚并将其降到最低，目前很难被当作“常识”。我们需要下更大的决心，减少理解上的重要隔阂，并让这些知识从学龄期起成为常规教育中的一部分。

2.提高制作、保存和传播关于疗效研究证据的系统评价的能力

很多关于疗效的问题都能很容易地得以解决，可以通过系统地评价已有证据、持续更新此类评价，以及将研究结果有效地传播给专家和患者。但要轻松获取已有证据用于制定系统评价，仍有很长的一段路要走。卫生系统的主要目的之一是解决证据不足的问题，这样才能使得有关疗效的可靠信息得到整合，并容易获取。

3. 疗效不确定时，应鼓励大家诚实面对

承认不确定性对医疗卫生专家而言通常是困难的，患者有时也不愿接受。因此患者会被给予一种安全的错觉，对证据中的不确定性浑然不知。医患双方要想成功合作并有效地完成疗效评估，那么二者都必须明白治疗措施未经充分评价可能会造成严重的后果；他们也必须进一步熟悉获取可靠证据所需要的方法。因此我们必须找到促成医患合作的最佳方法。

4.确定并优先研究患者和医生认为需要解决的重要问题

那些基础研究以及致力于医药公司利益最大化的研究，在可预见的未来，不太可能给患者带来益处，但却占据了研究资助单位和学术机构研究活动的绝大部分。对于那些不能赚钱但对患者至关重要的问题，即使已有公共资金扶持，我们也必须竭尽所能争取资源，才能保证研究能够进行。我们应该留意到：现在的研究更多的是为解决患者和医生对治疗效果的疑问，而资助方也会优先考虑研究这些问题以降低治疗的不确定性。

5.正视治疗共识的双重标准

那些承认疗效的不确定性并准备通过正规的治疗比较来解决这些问题的医生，在与患者沟通的过程中，相对其同事而言需要遵守更为严格的标准。这种违反常理的双重标准既不合逻辑，又不堪一击。当不确定某治疗的效果时，标准的解决办法就是开展随机（对照）试验或采用其他无偏倚的评价方法来评估疗效。我们应该确保参与疗效的研究并不是一种冒险行为，即“标准”做法总是安全有效的。

6.解决研究领域的低效问题

很多人惊讶地发现，研究者在为新研究寻求资助和伦理审批时，并不需要对已有证据进行系统评价。这就不可避免地造成了以下结果：低质量和根本没有必要的研究在无科学依据下以及伦理允许范围之外持续进行。我们应该敦促资助方和研究伦理委员会，确保研究者不会在未参考现有相关证据的系统评价的情况下就开展新的研究。一项新研究的报告，应始于参考那些表明开展更多研究必要性的系统评价，终于阐述新的研究结果对证据总体的影响。

7. 杜绝发表有偏倚的研究

我们需要从头到尾采取各种措施，来彻底杜绝发表有偏倚的研究。试验在一开始就应该被注册，并公开其计划书，以接受详细审查；而在试验完成时，所有试验的结果都应该被发表并提供原始数据，以便进行审查和进一步分析。

8. 要求商业信息和其他利益冲突透明化

大量现有证据表明，有时候既得利益和其他利益在研究设计、实施、分析、解读以及应用等多个方面会凌驾于患者的利益之上。这就使得双方的信任关系摇摇欲坠，也就不能更有效地保证研究服务于患者利益。所有试验的参与者，从商业公司到患者群体都应该将所得利益透明化，而非仅仅公开患者的利益。

现在开始行动

科学试验的改革迟迟未动。如果专家和患者共同行动，我们提出的方案将切实可行。读者们，你们应该呼唤变革——就在此刻！

行动计划：你能做的事情

确定对你重要的疗效问题。

学会辨别不确定性，说出来，提出疑问，寻求真实的答案。

不要担心询问医生目前有哪些治疗方案，或是接受某种治疗方案会怎样，不接受又会怎样。

在考虑可能的治疗方案时，你可在 www.ohri.ca/DecisionAid 或决策辅助网上找到有用的决策辅助信息。在本书的补充资源中（想了解更多关于共同决策的内容吗？）也可见到。

浏览信息可靠的网站，如英国国家卫生署决策网（www.nhs.uk）。或者参考本书第 12 章内容以及补充资源。

正确质疑那些毫无根据的宣称治疗取得“突破性进展”言论和媒体报道；注意媒体报道“数字”的方式，特别是标题中醒目的数字！

敢于挑战那些提供给你和你的家人的基于信仰和教义却无可靠证据可循的治疗方案。

小心被“贴上”多余的疾病标签，同时要避免过度检查（见第 2、4 章），即查明某疾病对于自己的风险高低。问问如果不马上采取措施会怎样。

只有符合以下情况，方可参与临床试验：1) 研究方案已注册并公开可得；2) 方案参考了现有证据的系统评价，并能说明试验是合理的；3) 收到书面保证，其所有研究结果都会被发表，并发给所有希望了解情况的参与者。

鼓励并与卫生专家、研究人员、研究资助者以及其他人士合作，他们试图通过研究来解决关于你所看重的疗效方面未能解决的问题。

鼓励更为广泛地开展针对偏倚和偶然因素影响的教育，劝说当选的政治代表及其他想要在学校课程中强调此类教育的人：行动起来，并从小学做起。

参考文献

序言—Ben Goldacre

- 1 L exchin J, Bero LA, Djulbegovic B, et al. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ* 2003; 326:1167-70.
- 2 S chwitzer G, Ganapati M, Henry D, et al. What are the roles and responsibilities of the media in disseminating health information? *PLoS Med* 2(7):e215.
- 3 Wilson PM, Booth AM, Eastwood A, et al. Deconstructing media coverage of trastuzumab (Herceptin): an analysis of national newspaper coverage. *Journal of the Royal Society of Medicine* 2008;101:125-32.
- 4 Shang A, Huwiler-Müntener K, Nartey L, et al. Are the clinical effects of homoeopathy placebo effects? Comparative study of placebo-controlled trials of homoeopathy and allopathy. *Lancet* 2005;366:726-32.
- 5 Bjelakovic G, Nikolova D, Gluud LL, et al. Antioxidant supplements for prevention of mortality in healthy participants and patients with various diseases. *Cochrane Database of Systematic Reviews* 2008, Issue 2. Art. No.: CD007176.

引言

- 1 Nolte E, McKee CM. Does health care save lives? Avoidable mortality revisited. London: Nuffield Trust, 2004.
- 2 Nolte E, McKee CM. Measuring the health of nations: updating an earlier analysis. *Health Affairs* 2008;27(1): 58-77.
- 3 Gigerenzer G. Reckoning with risk. London: Penguin Books, 2003. Ci ting Franklin B. Letter to Jean Baptiste Le Roy, 13 November 1789. Writings, vol x.
- 4 Goldacre B. Bad Science. London: Fourth Estate, 2008, px.
- 5 Matthews A, Dowswell T, Haas DM, et al. Interventions for nausea and vomiting in early pregnancy. *Cochrane Database of Systematic Reviews* 2010, Issue 9. Art. No.: CD007575.
- 6 Irwig L, Irwig J, Trevena L, et al. Smart health choices: making sense of health advice. London: Hammersmith Press, 2008. pdf freely available at: www.health.usyd.edu.au/shdg/resources/smart_health_choices.php and from www.jameslindlibrary.org.
- 7 Woloshin S, Schwartz LM, Welch HG. Know your chances: understanding health statistics. Berkeley: University of California Press, 2008. pdf freely available at www.jameslindlibrary.org.

第1章 新的不一定更好

- 1 Vandenbroucke JP. Thalidomide: an unanticipated adverse event. 2003. Available from: www.jameslindlibrary.org.
- 2 Stephens T, Brynner R. Dark medicine: the impact of thalidomide and its revival as a vital medicine. Cambridge, Mass: Perseus Publishing, 2001.
- 3 Thomson D, Capstick T. How a risk management programme can ensure safety in thalidomide use. *Pharmaceutical Journal* 2004 Feb 14:194-5.
- 4 Krumholz HM, Ross JR, Presler AH, et al . What have we learnt from Vioxx? *BMJ* 2007;334:120-3.
- 5 Merck statements dated 7 December 2009, on Vioxx proceedings in Ontario, Canada, and 4 March 2010 on Vioxx judgment in Australia, available at www.merck.com.
- 6 Cohen D. Rosiglitazone what went wrong? *BMJ* 2010;341:c4848.

- 7 Lehman R, Yudkin JS, Krumholz HM. Licensing drugs for diabetes: surrogate end points are not enough, robust evidence of benefits and harms is needed. *BMJ* 2010;341:c4805.
- 8 Blackstone EH. Could it happen again? The Björk-Shiley convexo-concave heart valve story. *Circulation* 2005;111:2717-19.
- 9 Wilson PM, Booth AM, Eastwood A, et al . Deconstructing media coverage of trastuzumab (Herceptin): an analysis of national newspaper coverage. *Journal of the Royal Society of Medicine* 2008;101:125-32.
- 10 Timmins N. Drugs watchdog gets harsh treatment. *Financial Times*, 8 October 2005, p6.
- 11 Hawkes N. Wonder drug is “cure” for cancer, say doctors. *The Times*, 20 October 2005.
- 12 Press MF, Sauter G, Bernstein L, et al . Diagnostic evaluation of HER-2 as a molecular target: an assessment of accuracy and reproducibility of laboratory testing in large, prospective, randomized clinical trials. *Clinical Cancer Research* 2005;11(18):6598-607.
- 13 NICE draft guidance on trastuzumab (Herceptin) for early breast cancer (press release), 9 June 2006. www.nice.org.uk/page.aspx?o=328789.
- 14 Cumming J, Mays N, Daubé J. How New Zealand has contained expenditure on drugs. *BMJ* 2010;340:1224-6.
- 15 NHS NICE Technology Appraisal TA34. Guidance on the use of trastuzumab for the treatment of advanced breast cancer. Issue date March 2002; review date April 2005. www.nice.org.uk/TA34.

第2章 预期的效果没有变为现实

- 1 Gilbert R, Salanti G, Harden M, et al . Infant sleeping position and the sudden infant death syndrome: systematic review of observational studies and historical review of clinicians' recommendations from 1940-2000. *International Journal of Epidemiology* 2005;34:74-87.
- 2 Furberg CD. Effect of antiarrhythmic drugs on mortality after myocardial infarction. *American Journal of Cardiology* 1983;52:32C-36C.
- 3 Chalmers I. In the dark. Drug companies should be forced to publish all the results of clinical trials. How else can we know the truth about their products? *New Scientist* 2004, 6 March, p19. Citing Moore T, *Deadly Medicine*. New York: Simon and Schuster, 1995.
- 4 Cowley AJ, Skene A, Stainer K, et al . The effect of lorainide on arrhythmias and survival in patients with acute myocardial infarction: an example of publication bias. *International Journal of Cardiology* 1993;40:161-6.
- 5 Chalmers I. Evaluating the effects of care during pregnancy and childbirth. In: Chalmers I, Enkin M, Keirse MJNC, eds. *Effective care in pregnancy and childbirth* . Oxford: Oxford University Press, 1989:3-38.
- 6 Ulfelder H. The stilbestrol disorders in historical perspective. *Cancer* 1980;45:3008-11.
- 7 Office of Technology Assessment. *Identifying health technologies that work: searching for evidence*. Washington, DC: US Government Printing Office, 1994.
- 8 Hemminki E, McPherson K. Impact of postmenopausal hormone therapy on cardiovascular events and cancer: pooled data from clinical trials. *BMJ* 1997;315:149-53.
- 9 Anonymous. HRT: update on the risk of breast cancer and long-term safety. *Current Problems in Pharmacovigilance* 2003;29:1-3. Citing results of Women's Health Initiative randomized controlled trial (*JAMA* 2003;289:3243-53) and Million Women Study (*Lancet* 2003;362:419-27).
- 10 Roberts H. Hormone replacement therapy comes full circle. *BMJ* 2007;335:219-20.

11 Williams HC. Evening primrose oil for atopic dermatitis: time to say goodnight (editorial). *BMJ* 2003;327:1358-9.

12 Hoare C, Li Wan Po A, Williams H. Systematic review of treatment for atopic eczema. *Health Technology Assessment* 2000;4(37):1-191.

13 Takwale A, Tan E, Agarwal S, et al . Efficacy and tolerability of borage oil in adults and children with atopic eczema: randomised, double blind, placebo controlled, parallel group trial. *BMJ* 2003;327:1385-7.

第3章 越多未必越好

1 Crile G. A plea against blind fear of cancer. *Life*, 31 October 1955, pp128-32.

2 Baum M, Houghton J. Contribution of randomised controlled trials to understanding and management of early breast cancer. *BMJ* 1999;319:568-71.

3 Veronesi U, Cascinelli N, Mariani L, et al . Twenty-year follow up of a randomized study comparing breast-conserving surgery with radical mastectomy for early breast cancer. *New England Journal of Medicine* 2002;347:1227-32.

4 Baum M. Breast beating: a personal odyssey in the quest for an understanding of breast cancer, the meaning of life and other easy questions. London: Anshan, 2010.

5 Japanese Breast Cancer Society. Results of questionnaires concerning breast cancer surgery in Japan 1980-2003. *Breast Cancer* 2005;12(1).

6 Early Breast Cancer Trialists' Collaborative Group. Effects of adjuvant tamoxifen and of cytotoxic therapy on mortality in early breast cancer. An overview of 61 randomized trials among 28,896 women. *New England Journal of Medicine* 1988;319:1681-92.

7 Clinical Trial Service Unit website: www.ctsu.ox.ac.uk.

8 The Cochrane Collaboration website: www.cochrane.org.

9 Kolata G, Eichenwald K. Health business thrives on unproven treatment, leaving science behind. *New York Times Special Report*, 2 October 1999.

10 Farquhar C, Marjoribanks J, Bassler R, et al . High dose chemotherapy and autologous bone marrow or stem cell transplantation versus conventional chemotherapy for women with early poor prognosis breast cancer. *Cochrane Database of Systematic Reviews* 2005, Issue 3. Art. No.:CD003139.

11 Farquhar C, Marjoribanks J, Bassler R, et al . High dose chemotherapy and autologous bone marrow or stem cell transplantation versus conventional chemotherapy for women with metastatic breast cancer. *Cochrane Database of Systematic Reviews* 2005, Issue 3. Art. No.: CD003142.

12 Piccart-Gebhart MJ, Procter M, Leyland-Jones B, et al . Trastuzumab after adjuvant chemotherapy in HER-2-positive breast cancer. *New England Journal of Medicine* 2005;353:1659-72.

13 Romond EH, Perez EA, Bryant J, et al . Trastuzumab plus adjuvant chemotherapy for operable HER-2-positive breast cancer. *New England Journal of Medicine* 2005;353:1673-84.

14 Carlson GW, Woods WC. Management of axillary lymph node metastasis in breast cancer: making progress. *JAMA* 2011;305:606-7.

第4章 越早未必越好

1 Raffle A, Gray M. Screening: evidence and practice. Oxford: Oxford University Press, rev. repr., 2009.

2 Sense About Science. Making sense of screening. London: Sense About Science, 2009.

Available from www.senseaboutscience.org.

- 3 Goodman MT, Gurney JG, Smith MA, et al. Sympathetic nervous system tumors. In: Ries LAG, Smith MA, Gurney JG, et al (eds). Cancer incidence and survival among children and adolescents: United States SEER Program 1975-1995. National Cancer Institute, SEER Program. NIH Pub. No.99-4649. Bethesda, MD, 1999. SEER Pediatric Monograph available at <http://seer.cancer.gov/publications/childhood>.
- 4 Mullassery D, Dominici C, Jesudason EC, et al. Neuroblastoma: contemporary management. Archives of Disease in Childhood – Education and Practice 2009;94:177-85.
- 5 Morris JK. Screening for neuroblastoma in children. Journal of Medical Screening 2002;9:56.
- 6 Raffle A, Gray M. op. cit., pp89-92.
- 7 Welch HG. Should I be tested for cancer? Maybe not and here's why. Berkeley and Los Angeles: University of California Press, 2004, p77.
- 8 Cosford PA, Leng GC, Thomas J. Screening for abdominal aortic aneurysm. Cochrane Database of Systematic Reviews, 2007, Issue 2, Art. No.: CD002945.
- 9 Welch HG. Screening mammography – a long run for a short slide? New England Journal of Medicine 2010; 363:1276-8.
- 10 Heath I. It is not wrong to say no. Why are women told only the benefits of breast screening and none of the possible harms? BMJ 2009; 338:1534.
- 11 Gøtzsche PC, Nielsen M. Screening for breast cancer with mammography. Cochrane Database of Systematic Reviews 2011, Issue 1. Art. No.: CD001877.
- 12 Kösters JP, Gøtzsche PC. Regular self-examination or clinical examination for early detection of breast cancer. Cochrane Database of Systematic Reviews 2003, Issue 2. Art. No.: CD003373. (No change, Update, Issue 3, July 2008.)
- 13 McPherson K. Should we screen for breast cancer? BMJ 2010;340:c3106.
- 14 Cancer Research UK. Prostate cancer – UK incidence statistics. Updated 23 December 2010. <http://info.cancerresearchuk.org/cancerstats/types/prostate/incidence>.
- 15 Chapman S, Barratt A, Stockler M. Let sleeping dogs lie? What men should know before getting tested for prostate cancer. Sydney: Sydney University Press, 2010. pdf available from: <http://ses.library.usyd.edu.au/bitstream/2123/6835/3/Let-sleeping-dogs-lie.pdf>.
- 16 Holmström B, Johansson M, Bergh A, et al. Prostate specific antigen for early detection of prostate cancer: longitudinal study. BMJ 2009;339:b3537.
- 17 Djulbegovic M, Beyth RJ, Neuberger MM, et al. Screening for prostate cancer: systematic review and meta-analysis of randomised controlled trials. BMJ 2010;341:c4543.
- 18 Stark JR, Mucci L, Rothman KJ, et al. Prostate cancer screening: the controversy continues. BMJ 2009;339:b3601.
- 19 National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. New England Journal of Medicine 2011;365:395-409.
- 20 Moynihan R. Beware the fortune tellers peddling genetic tests. BMJ 2010;341:c7233.
- 21 Thornton H. The screening debates: time for a broader approach? European Journal of Cancer 2003;39:1807-9.
- 22 Adapted from Wilson JMG, Jungner G. Principles and practice of screening for disease. Public health paper no 34. Geneva: World Health Organization, 1968.

23 C OMARE 12th Report: The impact of personally initiated X-ray computed tomography scanning for the health assessment of asymptomatic individuals. Press release, 19 December 2007. www.comare.org.uk/12thReportPressRelease.htm.

24 Department of Health. Better protection for patients having ‘MOT’ scans: http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/MediaCentre/Pressreleasesarchive/DH_115243.

25 Food and Drug Administration. Radiation-emitting products: Computed tomography and full-body CT scans – what you need to know: www.fda.gov/radiation-emittingProducts.

第 5 章 应对疗效的不确定性

1 Cabello JB, Burls A, Empanaza JI, et al . Oxygen therapy for acute myocardial infarction. Cochrane Database of Systematic Reviews 2010, Issue 6. Art No.: CD007160.

2 Glasziou P, Chalmers I, Rawlins M, et al . When are randomised trials unnecessary? Picking signal from noise. *BMJ* 2007;334:349-51.

3 Goh CL. Flashlamp-pumped pulsed dye laser (585nm) for the treatment of portwine stains: a study of treatment outcome in 94 Asian patients in Singapore. *Singapore Medical Journal* 2000;41:24-28.

4 Druker BJ, Talpaz M, Resta DJ, et al . Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *New England Journal of Medicine* 2001;344:1031-7.

5 Goldman J for the British Committee for Standards in Haematology. Recommendations for the management of BCR-ABL-positive chronic myeloid leukaemia. London: BSH, 2007.

6 Purohit N, Ray S, Wilson T, et al . The parent’s kiss: an effective way to remove paediatric nasal foreign bodies. *Annals of the Royal College of Surgeons of England* 2008;90:420-2.

7 Sa nghavi DM. How should we tell the stories of our medical miracles? *Lancet* 2010;375:2068-9.

8 Léauté-Labrèze C, Dumas la Roque E, Hubische T, et al . Propranolol for severe hemangiomas of infancy. *New England Journal of Medicine* 2008;358:2649-51.

9 Huikeshoven M, Koster PHL, de Borgie CAJM, et al . Re-darkening of port-wine stains 10 years after pulsed-dye-laser treatment. *New England Journal of Medicine* 2007;356:1235-40.

10 Waner M. Recent developments in lasers and the treatment of birthmarks. *Archives of Disease in Childhood* 2003;88:372-4.

11 Anti-Thrombotic Trialists’ (ATT) Collaboration. Clinical Trial Service Unit website: www.ctsu.ox.ac.uk/projects/att.

12 Lin CWC, Moseley AM, Refshauge KM. Rehabilitation for ankle fractures in adults. Cochrane Database of Systematic Reviews 2008, Issue 3. Art. No.: CD005595.

13 Lindley RI. Personal communication, 2005.

14 Wardlaw JM, Murray V, Berge E, et al . Thrombolysis for acute ischaemic stroke. Cochrane Database of Systematic Reviews 2009, Issue 4. Art. No.: CD000213.

15 Schmidt B, Roberts RS, Davis P, et al ; for the Caffeine for Apnea of Prematurity Trial Group. Long-term effects of caffeine therapy for apnea of prematurity. *New England Journal of Medicine* 2007;357:1893-902.

16 Caffeine citrate (Comment) in Neonatal Formulary 5. Available from: www.blackwellpublishing.com/medicine/bmj/nnf5/pdfs/comment/caffeine1.pdf.

17 Kenyon S, Pike, K, Jones DR, et al . Childhood outcomes after prescription of antibiotics to

pregnant women with preterm rupture of the membranes: 7-year follow-up of the ORACLE I trial. *Lancet* 2008;372:1310-18.

18 Kenyon S, Pike K, Jones DR, et al . Childhood outcomes after prescription of antibiotics to pregnant women with spontaneous preterm labour: 7-year follow-up of the ORACLE II trial. *Lancet* 2008;372:1319-27.

19 Erythromycin (Comment) in: *Neonatal Formulary* 5. Available from: www.blackwellpublishing.com/medicine/bmj/nnf5/pdfs/commentary/erythromycin.pdf.

20 Giuliano AE, Hunt KK, Ballman KV, et al . Axillary dissection vs no axillary dissection in women with invasive breast cancer and sentinel node metastasis: a randomized clinical trial. *JAMA* 2011;305:569-75.

21 General Medical Council. *Good Medical Practice* . London: GMC, 2006, p13.

22 Ashcroft R. Giving medicine a fair trial. *BMJ* 2000;320:1686.

23 Pritchard-Jones K, Dixon-Woods M, Naafs-Wilstra M, et al . Improving recruitment to clinical trials for cancer in childhood. *Lancet Oncology* 2008;9:392-9.

24 Equator network resource centre for good reporting of health research studies: www.equator-network.org.

25 Smithells RW. Iatrogenic hazards and their effects. *Postgraduate Medical Journal* 1975;15:39-52.

第 6 章 正确的验证疗效的试验

1 Hopkins WA. Patulin in the common cold. IV. Biological properties: extended trial in the common cold. *Lancet* 1943;ii:631-5.

2 Sanders TAB, Woolfe R, Rantzen E. Controlled evaluation of slimming diets: use of television for recruitment. *Lancet* 1990;336:918-20.

3 Glasziou P, Chalmers I, Rawlins M, et al . When are randomised trials unnecessary? Picking signal from noise. *BMJ* 2007;334: 349-51.

4 Pocock SJ. Randomised clinical trials. *BMJ* 1977;1:1661.

5 Balfour TG. Quoted in West C (1854). *Lectures on the Diseases of Infancy and Childhood*. London: Longman, Brown, Green and Longmans, p600.

6 King G, Gakidou E, Imai K, et al . Public policy for the poor? A randomised assessment of the Mexican universal health insurance programme. *Lancet* 2009;373:1447-54.

7 Peto J, Eden OB, Lilleyman J, et al . Improvement in treatments for children with acute lymphoblastic leukaemia: The Medical Research Council UKALL Trials, 1972-84. *Lancet* 1986;i:408-11.

8 Noseworthy JH, Ebers GC, Vandervoort MK, et al . The impact of blinding on the results of a randomized, placebo-controlled multiple sclerosis clinical trial. *Neurology* 1994;44:16-20.

9 Moseley JB, O'Malley K, Petersen NJ, et al . A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *New England Journal of Medicine* 2002;347:81-8.

10 Venning GR. Validity of anecdotal reports of suspected adverse drug reactions: the problem of false alarms. *BMJ* 1982;284:249-54.

11 McLernon DJ, Bond CM, Hannaford PC, et al on behalf of the Yellow Card Collaborative. Adverse drug reaction reporting in the UK: a retrospective observational comparison of Yellow Card reports submitted by patients and healthcare professionals. *Drug Safety* 2010;33:775-88.

12 Kocher T. Ueber Kropfexstirpation und ihre Folgen. *Archiv für Klinische Chirurgie*

1883;29:254-337.

13 Silverman WA, Andersen DH, Blanc WA, et al. A difference in mortality rate and incidence of kernicterus among premature infants allotted to two prophylactic regimens. *Pediatrics* 1956;18:614-25.

14 Zhang J, Ding E, Song Y. Adverse effects of cyclooxygenase 2 inhibitors on renal and arrhythmia events: meta-analysis of randomized trials. *JAMA* 2006;296:1619-21.

15 Vandenbroucke JP, Psaty BM. Benefits and risks of drug treatments: how to combine the best evidence on benefits with the best data about adverse effects. *JAMA* 2008;300:2417-9.

16 Whittington CJ, Kendall T, Fonagy P, et al. Selective serotonin reuptake inhibitors in childhood depression: systematic review of published versus unpublished data. *Lancet* 2004;363:1341-5.

第7章 考虑偶然因素

1 Antithrombotic Trialists' (ATT) Collaboration. Aspirin in the primary and secondary prevention of vascular disease: collaborative meta-analysis of individual participant data from randomised trials. *Lancet* 2009;373:1849-60.

2 CRASH trial collaborators. Final results of MRC CRASH, a randomised placebo-controlled trial of intravenous corticosteroid in adults with head injury – outcomes at 6 months. *Lancet* 2005;365:1957-9.

3 CRASH-2 trial collaborators. Effects of tranexamic acid on death, vascular occlusive events, and blood transfusion in trauma patients with significant haemorrhage (CRASH-2): a randomised, placebo-controlled trial. *Lancet* 2010;376:23-32.

4 Askie LM, Brocklehurst P, Darlow BA, et al and the NeOProm Collaborative Group. NeOProm: Neonatal Oxygenation Prospective Meta-analysis Collaboration study protocol. *BMC Pediatrics* 2011; 11:6.

第8章 评价所有相关的可靠证据

1 IS IS-2 (Second International Study of Infarct Survival) Collaborative Group. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17,187 cases of suspected acute myocardial infarction: ISIS-2. *Lancet* 1988;332:349-60.

2 Reynolds LA, Tansey EM, eds. Prenatal corticosteroids for reducing morbidity and mortality after preterm birth. London: Wellcome Trust Centre for the History of Medicine, 2005.

3 Dickersin K, Chalmers I. Recognising, investigating and dealing with incomplete and biased reporting of clinical research: from Francis Bacon to the World Health Organisation. James Lind Library 2010 (www.jameslindlibrary.org).

4 Cowley AJ, Skene A, Stainer K, et al. The effect of lorainide on arrhythmias and survival in patients with acute myocardial infarction: an example of publication bias. *International Journal of Cardiology* 1993;40:161-6.

5 Moore T. *Deadly Medicine*. New York: Simon and Schuster, 1995.

6 Stjernswärd J. Decreased survival related to irradiation postoperatively in early operable breast cancer. *Lancet* 1974;ii:1285-6.

7 Stjernswärd J. Meta-analysis as a manifestation of 'bondföruft' ('peasant sense'). *JLL Bulletin: Commentaries on the history of treatment evaluation* 2009 (www.jameslindlibrary.org).

8 Fugh-Berman AJ. The haunting of medical journals: how ghostwriting sold "HRT". *PLoS Medicine* 2010;7(9):e1000335.

9 Whittington CJ, Kendall T, Fonagy P, et al. Selective serotonin-reuptake inhibitors in

childhood depression: systematic review of published versus unpublished data. *Lancet* 2004;363:1341-5.

10 Spielmans GI, Biehn TL, Sawrey DL. A case study of salami slicing: pooled analysis of duloxetine for depression. *Psychotherapy and Psychosomatics* 2010;79:97-106.

11 Antman EM, Lau J, Kupelnick B, et al . A comparison of results of meta-analysis of randomized control trials and recommendations of clinical experts. *JAMA* 1992;268:240-8.

12 Natanson C, Kern SJ, Lurie P, et al . Cell-free hemoglobin-based blood substitutes and risk of myocardial infarction and death: a meta-analysis. *JAMA* 2008;299(19):2304-12.

13 Chalmers I. TGN1412 and The Lancet's solicitation of reports of phase 1 trials. *Lancet* 2006;368:2206-7.

14 Jack A. Call to release human drug trial data. *Financial Times*, 8 August 2006.

15 Kenter MJH, Cohen AF. Establishing risk of human experimentation with drugs: lessons from TGN1412. *Lancet* 2006;368:1387-91.

16 McLellan F. 1966 and all that – when is a literature search done? *Lancet* 2001;358:646.

17 Horn J, Limburg M. Calcium antagonists for acute ischemic stroke. *Cochrane Database of Systematic Reviews* 2000, Issue 1. Art. No.: CD001928.

18 Horn J, de Haan RJ, Vermeulen M, et al . Nimodipine in animal model experiments of focal cerebral ischemia: a systematic review. *Stroke* 2001;32:2433-8.

19 O'Collins VE, Macleod MR, Donnan GA, et al . 1,026 experimental treatments in acute stroke. *Annals of Neurology* 2006;59:467-77.

20 CRASH trial collaborators. Final results of MRC CRASH, a randomised placebo-controlled trial of intravenous corticosteroid in adults with head injury – outcomes at 6 months. *Lancet* 2005;365:1957-9.

第9章 对正确疗法的试验监管：是推动还是阻碍？

1 Emanuel EJ, Menikoff J. Reforming the regulations governing research with human subjects. *New England Journal of Medicine* 2011;10.1056/NEJMs1106942.NEJM.org.

2 Chalmers I, Lindley R. Double standards on informed consent to treatment. In: Doyal L, Tobias JS, eds. *Informed consent in medical research* . London: BMJ Books 2001, pp266-75.

3 Fallowfield L, Jenkins V, Farewell V, et al . Efficacy of a Cancer Research UK communicating skills training model for oncologists: a randomised controlled trial. *Lancet* 2002;359:650-6.

4 Chalmers I. Regulation of therapeutic research is compromising the interests of patients. *International Journal of Pharmaceutical Medicine* 2007;21:395-404.

5 Roberts I, Prieto-Marino D, Shakur H, et al . Effect of consent rituals on mortality in emergency care research. *Lancet* 2011;377:1071-2.

第10章 高质量、低质量和不必要的研究

1 Equator network resource centre for good reporting of health research studies: www.equator-network.org.

2 European Carotid Surgery Trialists' Collaborative Group. Randomised trial of endarterectomy for recently symptomatic carotid stenosis: final results of the MRC European Carotid Surgery Trial (ECST). *Lancet* 1998;351:1379-87.

3 Cina CS, Clase CM, Haynes RB. Carotid endarterectomy for symptomatic carotid stenosis. *The Cochrane Database of Systematic Reviews* 1999, Issue 3. Art. No.: CD001081.

4 The Magpie Trial Collaborative Group. Do women with pre-eclampsia, and their babies,

- benefit from magnesium sulphate? The Magpie Trial: a randomised, placebo-controlled trial. *Lancet* 2002;359:1877-90.
- 5 D uley L, Gülmezoglu AM, Henderson-Smart DJ. Magnesium sulphate and other anticonvulsants for women with pre-eclampsia. *Cochrane Database of Systematic Reviews* 2003, Issue 2. Art. No.: CD000025.
- 6 Global Report. UNAIDS report on the Global AIDS epidemic 2010: www.unaids.org/globalreport/Global_report.htm.
- 7 Grimwade K, Swingler G. Cotrimoxazole prophylaxis for opportunistic infections in adults with HIV. *Cochrane Database of Systematic Reviews* 2003, Issue 3. Art. No.: CD003108.
- 8 C hintu C, Bhat GJ, Walker AS, et al . Co-trimoxazole as prophylaxis against opportunistic infections in HIV-infected Zambian children (CHAP): a double blind randomised placebo-controlled trial. *Lancet* 2004;364:1865-71.
- 9 MRC News Release. Antibiotic drug almost halves AIDS-related death in children. London: MRC, 19 November 2004.
- 10 World Health Organization and UNICEF. Co-trimoxazole prophylaxis for HIV-exposed and HIV-infected infants and children: practical approaches to implementation and scale up. WHO and UNICEF, 2009.
- 11 S oares K, McGrath J, Adams C. Evidence and tardive dyskinesia. *Lancet* 1996;347:1696-7.
- 12 Thornley B, Adams C. Content and quality of 2000 controlled trials in schizophrenia over 50 years. *BMJ* 1998;317:1181-4.
- 13 Howell CJ, Chalmers I. A review of prospectively controlled comparisons of epidural with non-epidural forms of pain relief during labour. *International Journal of Obstetric Anesthesia* 1992;1:93-110.
- 14 Horn J, Limburg M. Calcium antagonists for acute ischemic stroke. *Cochrane Database of Systematic Reviews* 2000, Issue 1. Art No.: CD001928.
- 15 Horn J, de Haan RJ, Vermeulen RD, Luiten PGM, et al . Nimodipine in animal model experiments of focal cerebral ischemia: a systematic review. *Stroke* 2001;32:2433-8.
- 16 Fergusson D, Glass KC, Hutton B, et al . Randomized controlled trials of aprotinin in cardiac surgery: using clinical equipoise to stop the bleeding. *Clinical Trials* 2005;2:218-32.
- 17 Tallon D, Chard J, Dieppe P. Relation between agendas of the research community and the research consumer. *Lancet* 2000;355:2037-40.
- 18 Cream J, Cayton H. New drugs for Alzheimer's disease – a consumer perspective. *CPD Bulletin Old Age Psychiatry* 2001;2:80-2.
- 19 C ohen CI, D'Onofrio A, Larkin L, et al . A comparison of consumer and provider preferences for research on homeless veterans. *Community Mental Health Journal* 1999;35:273-9.
- 20 Griffiths KM, Jorm AF, Christensen H, et al . Research priorities in mental health, Part 2: an evaluation of the current research effort against stakeholders' priorities. *Australian and New Zealand Journal of Psychiatry* 2002;36:327-39.
- 21 Oliver S, Gray J. A bibliography of research reports about patients', clinicians' and researchers' priorities for new research. London: James Lind Alliance, December, 2006.
- 22 Chalmers I. Current controlled trials: an opportunity to help improve the quality of clinical research. *Current Controlled Trials in Cardiovascular Medicine* 2000;1:3-8. Available from:

<http://cvm.controlled-trials.com/content/1/1/3>.

- 23 Editorial. Safeguarding participants in controlled trials. *Lancet* 2000;355:1455-63.
 - 24 Fugh-Berman A. The haunting of medical journals: how ghostwriting sold "HRT". *PLoS Medicine* 2010;7(9):e10000335.
 - 25 Heimans L, van Hylckama V, Dekker FW. Are claims of advertisements in medical journals supported by RCTs? *Netherlands Journal of Medicine* 2010;68:46-9.
 - 26 Lexchin J, Bero LA, Djulbeovic B, et al. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ* 2003;326:1167-76.
 - 27 Weatherall D. Academia and industry: increasingly uneasy bedfellows. *Lancet* 2000;355:1574.
 - 28 Angell M. Is academic medicine for sale? *New England Journal of Medicine* 2000;342:1516-8.
 - 29 Grant J, Green L, Mason B. From bench to bedside: Comroe and Dripps revisited. HERG Research Report No. 30. Uxbridge, Middlesex: Brunel University Health Economics Research Group, 2003.
 - 30 Pound P, Ebrahim S, Sandercock P, et al. Reviewing Animal Trials Systematically (RATS) Group. Where is the evidence that animal research benefits humans? *BMJ* 2004;328:514-7.
 - 31 Weatherall D. The quiet art revisited. *Lancet* 2011;377:1912-13.
 - 32 Pirmohamed M. Cited in Mayor S. Fitting the drug to the patient. *BMJ* 2007;334:452-3.
 - 33 Editorial. The human genome at ten. *Nature* 2010;464:649-50.
 - 34 Mackillop WJ, Palmer MJ, O'Sullivan B, et al. Clinical trials in cancer: the role of surrogate patients in defining what constitutes an ethically acceptable clinical experiment. *British Journal of Cancer* 1989;59:388-95.
 - 35 The Psoriasis Association: www.psoriasis-association.org.uk.
 - 36 National Psoriasis Association. Statistics about psoriasis: www.psoriasis.org/netcommunity/learn_statistics.
 - 37 Jobling R. Therapeutic research into psoriasis: patients' perspectives, priorities and interests. In: Rawlins M, Littlejohns P, ed. *Delivering quality in the NHS 2005*. Abingdon: Radcliffe Publishing Ltd, pp53-6.
- 第 11 章 做高质量研究，人人有责
- 1 Oliver S, Clarke-Jones L, Rees R, et al. Involving consumers in research and development agenda setting for the NHS: developing an evidence-based approach. *Health Technology Assessment Report* 2004;8(15).
 - 2 NIHR Guy's and St. Thomas' and King's College London's Biomedical Research Centre. Involving users in the research process: a 'how to' guide for researchers. Version 1, April 2010. Available from: www.biomedicalresearchcentre.org.
 - 3 Cartwright J, Crowe S. *Patient and public involvement toolkit*. London: Wiley-Blackwell and BMJI Books, 2011.
 - 4 European Science Foundation/EMRC. *Implementation of medical research in clinical practice – a growing challenge*. Strasbourg: ESF, 2011.
 - 5 Hanley B, Truesdale A, King A, et al. Involving consumers in designing, conducting, and interpreting randomised controlled trials: questionnaire survey. *BMJ* 2001;322:519-23.
 - 6 Koops L, Lindley RI. Thrombolysis for acute ischaemic stroke: consumer involvement in

design of new randomised controlled trial. *BMJ* 2002;325:415-7.

7 Staley K. Exploring impact: public involvement in NHS, public health and social care research. Eastleigh: INVOLVE, 2009. Available from: www.invo.org.uk.

8 Petit-Zeman S, Firkins L, Scadding JW. The James Lind Alliance: tackling research mismatches. *Lancet* 2010;376:667-9.

9 Patient Partner Project. An EU programme 'Identifying the needs for patients partnering in clinical research': www.patientpartner-europe.eu.

10 Thornton H, Edwards A, Elwyn G. Evolving the multiple roles of 'patients' in health-care research: reflections after involvement in a trial of shared decision-making. *Health Expectations* 2003;6:189-97.

11 Dixon-Woods M, Agarwal S, Young B et al. Integrative approaches to qualitative and quantitative evidence. NHS Health Development Agency, 2004.

12 Kushner R. Breast cancer: a personal history and an investigative report. New York: Harcourt Brace Jovanovitch, 1975.

13 Lerner BH. The breast cancer wars: hope, fear, and the pursuit of a cure in twentieth-century America. New York: Oxford University Press, 2003.

14 Institute of Medical Ethics Working Party on the ethical implications of AIDS: AIDS, ethics, and clinical trials. *BMJ* 1992;305:699-701.

15 Thornton H. The patient's role in research. [Paper given at The Lancet 'Challenge of Breast Cancer' Conference, Brugge, April 1994.] In: Health Committee Third Report. Breast cancer services. Volume II. Minutes of evidence and appendices. London: HMSO, July 1995, 112-4.

16 Concorde Coordinating Committee. Concorde: MRC/ANRS randomised double-blind controlled trial of immediate and deferred zidovudine in symptom-free HIV infection. *Lancet* 1994;343:871-81.

17 Pehudoff K, Alves TL. Patient and consumer organisations at the European Medicines Agency: financial disclosure and transparency. Amsterdam: Health Action International, 2010. Available from www.haieurope.org.

18 Herxheimer A. Relationships between the pharmaceutical industry and patients' organisations. *BMJ* 2003;326:1208-10.

19 Consumers' Association. Who's injecting the cash? *Which?* 2003, April, pp24-25.

20 Koops L, Lindley RI. Thrombolysis for acute ischaemic stroke: consumer involvement in design of new randomised controlled trial. *BMJ* 2002;325:415-7.

21 Donovan J, Mills N, Smith M, et al for the ProtecT Study Group. Quality improvement report: improving design and conduct of randomised trials by embedding them in qualitative research: ProtecT (prostate testing for cancer and treatment) study. *BMJ* 2002;325:766-70.

第 12 章 那么，什么让卫生保健变得更好？

1 Edwards A, Elwyn G, Atwell C, et al. Shared decision making and risk communication in general practice – a study incorporating systematic literature reviews, psychometric evaluation of outcome measures, and quantitative, qualitative and health economic analyses of a cluster randomised trial of professional skill development. Report to Health in Partnership programme, UK Department of Health. Cardiff: Department of General Practice. University of Wales College of Medicine, 2002.

2 Farrell C, ed. Patient and public involvement in health: The evidence for policy implementation. A summary of the results of the Health in Partnership research programme.

London: Department of Health Publications, April 2004. Available from: www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4082332.

3 Adapted from Marshall T. Prevention of cardiovascular disease. Risk and benefit calculator. Available from: www.haps.bham.ac.uk/publichealth/cardiovascular/index.shtml.

4 Evans I, Thornton H. Transparency in numbers: the dangers of statistical illiteracy. *Journal of the Royal Society of Medicine* 2009;102:354-6.

5 Irwig L, Irwig J, Trevena L, et al. Smart health choices: making sense of health advice. London: Hammersmith Press, 2008.

6 Salzburg Global Seminar website: www.salzburgglobal.org.

7 Salzburg statement on shared decision making: Salzburg Global Seminar. *BMJ* 2011;342:d1745. Available from: www.bmj.com/content/342/bmj.d1745.full.

第 13 章 合理的研究：未来更美好的蓝图

1 Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet* 2009;374:86-89.

2 Roberts I, Yates D, Sandercock P, et al; CRASH trial collaborators. Effect of intravenous corticosteroids on death within 14 days in 10008 adults with clinically significant head injury (MRC CRASH trial): randomised placebo-controlled trial. *Lancet* 2004;364:1321-8.

3 Cochrane Injuries Group Albumin Reviewers. Human albumin administration in critically ill patients: systematic review of randomised controlled trials. *BMJ* 1998;317:235-40.

4 Finfer S, Bellomo R, Bryce N, et al (SAFE Study Investigators). A comparison of albumin and saline for fluid resuscitation in the intensive care unit. *New England Journal of Medicine* 2004;350:2247-56.

5 Edwards A, Elwyn G, Atwell C, et al. Shared decision making and risk communication in general practice – a study incorporating systematic literature reviews, psychometric evaluation of outcome measures, and quantitative, qualitative and health economic analyses of a cluster randomised trial of professional skill development. Report to Health in Partnership programme, UK Department of Health. Cardiff: Department of General Practice. University of Wales College of Medicine, 2002.

6 Farrell C, ed. Patient and public involvement in health: The evidence for policy implementation. A summary of the results of the Health in Partnership research programme. London: Department of Health Publications, April 2004. Available from: www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4082332.

7 Donovan J, Mills N, Smith M, et al for the ProtecT Study Group. Quality improvement report: improving design and conduct of randomised trials by embedding them in qualitative research: ProtecT (prostate testing for cancer and treatment) study. *BMJ* 2002;325:766-70.

8 Dickersin K, Chalmers I. Recognising, investigating and dealing with incomplete and biased reporting of clinical research: from Francis Bacon to the World Health Organization. James Lind Library, 2010 (www.jameslindlibrary.org).

9 Chan A-W, Hróbjartsson A, Haahr MT, et al. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004;291:2457-65.

10 Chan AW, Altman DG. Epidemiology and reporting of randomised trials published in

PubMed journals. Lancet 2005;365:1159-62.

11 Glasziou P, Meats E, Heneghan C, Shepperd S. What is missing from descriptions of treatment in trials and reviews? BMJ 2008;336:1472-4.

12 Clarke M, Hopewell S, Chalmers I. Clinical trials should begin and end with systematic reviews of relevant evidence: 12 years and waiting. Lancet 2010;376:20-21.

附加资源

你想要进一步了解关于验证疗效的相关信息吗？

网站

验证疗效交互平台

你可以在这个网站找到免费的本书第二版的电子版本，并且会不断有新的内容补充进去，本书第一版的内容在网站上已经有阿拉伯文、中文、德文、意大利文、波兰文和西班牙文下载。

www.testingtreatments.org.

詹姆斯·林德图书馆

www.jameslindlibrary.org

Cochrane 协作网

www.cochrane.org

NHS Choices

www.nhs.uk (enter 'research' in search window)

UK Clinical Research Collaboration

www.ukcrc.org

Healthtalkonline

www.healthtalkonline.org

US National Cancer Institute

关于临床试验的教育材料

<http://cancertrials.nci.nih.gov/clinicaltrials/learning>

图书

Ben Goldacre. *Bad science*. London: Harper Perennial, 2009. Bengt D Furberg, Curt D Furberg. *Evaluating clinical research: all the glitters is not gold*. 2nd edition. New York: Springer, 2007.

Steven Woloshin, Lisa Schwartz, Gilbert Welch. *Know your chances: understanding health statistics*. Berkeley: University of California Press, 2008. Available free at www.jameslindlibrary.org.

Les Irwig, Judy Irwig, Lyndal Trevena, Melissa Sweet. *Smart health choices: making sense of health advice*. London: Hammersmith Press, 2008. Available free at www.jameslindlibrary.org.

Trish Greenhalgh. *How to read a paper: the basics of evidence-based medicine*. 4th edition. Oxford and London: Wiley-Blackwell and BMJ Books, 2010.

H Gilbert Welch, Lisa M. Schwartz, Steven Woloshin. *Overdiagnosed: making people sick in the pursuit of health*. Boston: Beacon Press, 2011.

你了解那些我们已知有关疗效的信息？

Cochrane 图书馆

www.thecochranelibrary.com

NHS Evidence

www.evidence.nhs.uk

Informed Health Online

www.informedhealthonline.org

PubMed Health

www.pubmed.gov/health

你想要了解那些对当前疗效还不确定的研究信息吗？

UK Database of Uncertainties about the Effects of Treatments (UK DUETs)

www.evidence.nhs.uk

你了解当前关注于那些疗效不确定的研究信息吗？

WHO International Clinical Trials Registry Platform

www.who.int/trialsearch

US National Institutes of Health Clinical Trials Registry

www.clinicaltrials.gov

EU Clinical Trials Register

<https://www.clinicaltrialsregister.eu>

Australian Cancer Trials

www.australiancancertrials.gov.au

你想参与到与自己相关的和高质量的验证疗效的研究中去吗？

James Lind Alliance

www.lindalliance.org

Promotes working partnerships between patients and clinicians to identify and prioritize important uncertainties about the effects of treatments.

National Institute for Health Research

NIHR Health Technology Assessment

www.nchta.org

Actively involves service-users in all stages of its work.

NIHR Clinical Research Network Coordinating Centre

www.crnc.nihr.ac.uk/ppi

Keen to involve patients, carers, and the public in volunteering for clinical studies and getting actively involved as researchers.

Cochrane Consumer Network

www.consumers.cochrane.org

Promotes patient input to systematic reviews of treatments prepared by the Cochrane Collaboration.

UK Clinical Research Network

www.ukcrn.org.uk

你想练习如何评价研究吗？

Critical Appraisal Skills Programme

www.casp-uk.net

Organizes workshops and other resources to help individuals to develop the skills to find and make sense of research evidence.

US Cochrane Center

Understanding Evidence-based Healthcare: A Foundation for Action <http://us.cochrane.org/understanding-evidence-based-healthcare-foundation-action>

A web course designed to help individuals understand the fundamentals of evidence-based

healthcare concepts and skills.

你想要了解更多关于共同决策的内容吗？

The Foundation for Informed Medical Decision Making

www.informedmedicaldecisions.org

Dartmouth-Hitchcock Medical Center:

Center for Shared Decision Making

http://patients.dartmouth-hitchcock.org/shared_decision_making.html

Salzburg Statement

www.bmj.com/content/342/bmj.d1745.full

www.salzburgglobal.org

你想要了解动物实验的系统评价吗？

www.sabre.org.uk

www.camarades.info

插图列表

引言

p xxi 不要过于确定

Xenophanes, 6th century BCE

Charlie ('Peanuts') Brown, 20th century CE

Susser M. Causal thinking in the health sciences.

Oxford: Oxford University Press, 1983.

第 1 章 新的不一定更好

p 2 轶事仅仅只是轶事

Ross N. Foreword. In: Ernst E, ed. Healing, hype, or harm?

A critical analysis of complementary or alternative medicine.

Exeter: Societas, 2008:vi-vii.

p 4 婴儿流行性失明悲剧

Silverman WA. Human experimentation: a guided step into the unknown. Oxford: Oxford University Press, 1985:vii-viii.

p 11 卷入赫赛汀热潮

Cooper J. Herceptin (rapid response). BMJ.

Posted 29 November 2006 at www.bmj.com.

第 2 章 预期的效果没有变为现实

p 17 难怪她会感到困惑

Huntingford CA. Confusion over benefits of hormone replacement therapy. Lancet 2004;363:332.

第 3 章 越多未必越好

p 22 我们这样做，是因为……

Parmar MS. We do things because (rapid response). BMJ.

Posted 1 March 2004 at www.bmj.com.

p 23 激进的治疗措施并非总是最佳

Brewin T in Rees G, ed. The friendly professional: selected writings of Thurstan Brewin. Bognor Regis: Eurocommunica, 1996.

p 24 传统的根治性乳房切除术/范围更大的根治性乳房切除术

Adapted from Lerner BH. The breast cancer wars: hope, fear and the pursuit of a cure in twentieth-century America.

New York: Oxford University Press, 2003.

p 26 随机分配——一个简单的解释

Harrison J. Presentation to Consumers' Advisory Group for Clinical Trials, 1995.

p 29 努力获取无偏倚的证据

Adapted from Kolata G, Eichenwald K. Health business thrives on unproven treatment, leaving science behind.

New York Times Special Report, 2 October 1999.

第 4 章 越早未必越好

p 31 从普通人到患者

Cochrane AL, Holland WW. Validation of screening procedures.

British Medical Bulletin 1971;27:3-8.

p 35 不要推测早期检测就是有价值的

Morris JK. Screening for neuroblastoma in children.

Journal of Medical Screening 2002;9:56.

p 39 前列腺癌的过度诊断

Chapman S, Barratt A, Stockler M. Let sleeping dogs lie?

What men should know before getting tested for prostate cancer.

Sydney: Sydney University Press, 2010: p25.

p 40 PSA 发现者的话

Ablin RJ. The great prostate mistake. New York Times, 10 March 2010.

p 42 “推销” 筛查

Woloshin S, Schwartz LM. Numbers needed to decide.

Journal of the National Cancer Institute 2009;101:1163-5.

p 44 别和你的基因打赌

Sense About Science. Making sense of testing: a guide to why scans and other health tests for well people aren't always a good idea .

London: Sense About Science 2008, p7.

Available from www.senseaboutscience.org

p 46 一场筛查闹剧

Warlow C. The new religion: screening at your parish church.

BMJ 2009;338:b1940.

第 5 章 应对疗效的不确定性

p 53 科学进展不应该成为新闻头条

Goldacre B. Bad Science. London: Fourth Estate, 2008, p219.

p 56 直面不确定性：一个生死攸关的问题

Chalmers I. Addressing uncertainties about the effects of treatments offered to NHS patients: whose responsibility?

Journal of the Royal Society of Medicine 2007;100:440.

p 57 解决不确定性是门专业

From: Medical Research Council response to Royal College of Physicians consultation on medical professionalism. 2005.

p 59 医生就处方中猜测成分的对话

Adapted from Petit-Zeman S. Doctor, what's wrong? Making the NHS human again. London: Routledge, 2005, pp79-80.

p 61 患者能否解决不确定性的问题?

Evans I. More nearly certain.

Journal of the Royal Society of Medicine 2005;98:195-6.

第 6 章 正确的验证疗效的试验

p 65 误解治疗

James Stuart, King of Great Britaine, France and Ireland. A counterblaste to tobacco. In: The workes of the most high and mightie prince, James.

Published by James, Bishop of Winton, and Deane of his Majesties

Chappel Royall. London: printed by Robert Barker and John Bill, printers to the Kings most excellent Majestie, 1616: pp 214-222.

p 66 所信即所见

Asher R. Talking sense (Lettsomian lecture, 16 Feb, 1959).
Transactions of the Medical Society of London , vol LXXV, 1958-59.
Reproduced in: Jones, FA, ed. Richard Asher talking sense.
London: Pitman Medical, 1972.

p 82 黄卡计划

B owser A. A patient's view of the Yellow Card Scheme. In:
Medicines & Medical Devices Regulation: what you need to know .
London: MHRA, 2008. Available at www.mhra.gov.uk.

第 7 章 考虑偶然因素

p 89 “统计学意义”意味着什么？

Spiegelhalter D, quoted in: Making Sense of Statistics. 2010.
www.senseaboutscience.org.

第 8 章 评价所有相关的可靠证据

p 93 为什么开展研究？

Robinson KA, Goodman SN. A systematic examination of the citation of
prior research in reports of randomized, controlled trials.
Annals of Internal Medicine 2011;154:50-55.

p 94 综合研究信息

R ayleigh, Lord. In: Report of the fifty-fourth meeting of the British
Association for the Advancement of Science; held at Montreal in August
and September 1884. London: John Murray, 1884: pp3-23.

p 95 系统评价的重要性

Moher D, Liberati A, Tetzlaff, Altman DG. The PRISMA Group. Preferred
reporting items for systematic reviews and meta-analyses: The PRISMA
statement (www.equator-network.org), 2009.

p 97 营销医学

Spielmanns GI, Parry PI. From Evidence-based Medicine to Marketing-based
Medicine: Evidence from Internal Industry Documents . Journal of Bioethical
Inquiry 2010;7(1):13-29. Available online: <http://tinyurl.com/Spielmanns>.

p 100 科学知识日积月累，但科学家并未科学地积累证据

Goldacre B. Bad Science: How pools of blood trials could save lives. The
Guardian , 10 May 2008, p16.

p 102 事先查找证据能否避免死亡？

Perkins E. Johns Hopkins Tragedy. Information Today 2001;18:51-54.

p 103 医学期刊《柳叶刀》投稿须知：将研究结果和以往证据对比

Clark S, Horton R. Putting research in context – revisited.
Lancet 2010;376:10-11.

第 9 章 对正确疗法的试验监管：是推动还是阻碍？

p 105 谁说医学研究有害健康？

Hope T. Medical ethics: a very short introduction.
Oxford: Oxford University Press, 2004, p99.

p 108 一个理想世界

Goldacre B. Pharmaco-epidemiology would be fascinating enough even if

society didn't manage it really really badly. The Guardian, 17 July 2010.
Available online: www.badsience.net/2010/07/pharmaco-epidemiology-would-be-fascinating-enough-even-if-society-didnt-manage-it-really-really-badly.

p 109 偏倚伦理学

Lantos J. Ethical issues – how can we distinguish clinical research from innovative therapy?

American Journal of Pediatric Hematology/Oncology 1994; 16: 72-75.

p 110 反思知情同意

Manson NC, O'Neill O. Rethinking informed consent in bioethics.

Cambridge: Cambridge University Press, 2007, p200.

p 111 在高质量的医疗实践中进行知情同意的一般方法

Gill R. How to seek consent and gain understanding. BMJ 2010;341:c4000.

p 113 学术争论，还是明智选择？

Harrison J. Testing times for clinical research. Lancet 2006; 368:909-910.

p 114 研究监管系统应该做什么

Ashcroft, R. Giving medicine a fair trial. BMJ 2000;320:1686.

第 10 章 高质量、低质量和不必要的研究

p 117 我参加 Magpie 试验的经历

MRC News Release. Magnesium sulphate halves risk of eclampsia and can save lives of pregnant women. London: MRC, 31 May 2002.

p 124 “Me-too 药”在加拿大的影响

Morgan SG, Bassett KL, Wright JM, et al. 'Breakthrough' drugs and growth in expenditure on prescription drugs in Canada. BMJ 2005;331:815-6.

p 125 医生和制药公司

Angell M. Drug companies & doctors: a story of corruption.

New York Review of Books, 15 January 2009.

p 126 骗子、傻瓜和坏蛋？

Sackett DL, Oxman AD. HARLOT plc: an amalgamation of the world's two oldest professions. BMJ 2003;327:1442-5.

p 127 竭尽全力只为找基因

Iannucci A. The Audacity of Hype. London: Little, Brown, 2009, pp270-1.

p 128 科学研究未能使银屑病患者受益

Jobling R. Therapeutic research into psoriasis: patients' perspectives, priorities and interests. In: Rawlins M, Littlejohns P, eds. Delivering quality in the NHS 2005. Abingdon: Radcliffe Publishing Ltd, pp53-6.

第 11 章 做高质量研究，人人有责

p 131 患者的选择：大卫和歌利亚

Refractor. Patients' choice: David and Goliath. Lancet 2001;358:768.

p 132 至关重要的合作关系

Professor Dame Sally Davies. Foreword to Staley K. Exploring impact: public involvement in NHS, public health and social care research.

Eastleigh: INVOLVE, 2009.

p 134 外行帮助反思艾滋病（AIDS）

Epstein S. Impure science: AIDS, activism and the politics of knowledge . London: University of California Press, 1996.

p 137 儿童消费力和新药

Wilson PM, Booth AM, Eastwood A. Deconstructing media coverage of trastuzumab (Herceptin): an analysis of national newspaper coverage. *Journal of the Royal Society of Medicine* 2008;101:125-132.

p 138 公众参与改善卫生保健

Moynihan R. Power to the people. *BMJ* 2011;342:d2002.

第 12 章 那么，什么让卫生保健变得更好？

p 144 共同决策

Adapted from Thornton H. Evidence-based healthcare. What roles for patients? In: Edwards A, Elwyn G, eds. *Shared decision-making in health care. Achieving evidence-based patient choice . Second edition.* Oxford: Oxford University Press, 2009, p39.

p 151 别被吸引眼球的数值愚弄！

Goldacre B. *Bad Science.* London: Fourth Estate, 2008, pp239-40.

p 157 谁才有糖尿病？

Welch HG, Schwartz LM, Woloshin S. *Overdiagnosed: making people sick in the pursuit of health.* Boston: Beacon Press, 2011, pp17-18.

要点列表

第1章 新的不一定更好

- 验证新的治疗方案是必要的，因为其比现有治疗更好或更差的可能性是一样的
- 对治疗方案有偏倚（不科学）的试验可导致患者遭受痛苦甚至死亡
- 事实上，被批准的治疗并不能保证就是安全的
- 治疗的副作用通常需要经过一段时间才会出现
- 治疗的好处常常会被夸大，而坏处则会被淡化

第2章 预期的效果没有变为现实

- 医学理论或专家意见对安全有效的治疗都不是可靠的指导
- “已被使用”的治疗并不意味着利大于弊
- 即使患者没有遭受未经充分验证的治疗带来的伤害，使用它们仍会对个人和社会的资源造成浪费

第3章 越多未必越好

- 治疗越多未必越好，有些情况可能弊大于利。

第4章 越早未必越好

- 早期诊断未必能带来更好的结果；有时反而使问题更糟
- 只有在良好证据显示其有效的情况下，才能开展筛查项目
- 不开展筛查也许是最好的选择
- 受邀参加筛查的人群需获得全面的关于利弊的信息
- 筛查的益处往往被夸大
- 筛查的弊端常未予以重视或被忽视
- 讨论筛查的利弊及风险很必要

第5章 应对疗效的不确定性

- 效果显著的治疗很少见
- 治疗的不确定性普遍存在
- 通常情况下，不同治疗的效果都有微小的差别，合理地辨别这些差别很重要
- 当没有人知道该如何回答疗效中重要的不确定性时，需要采取行动来减少这种不确定性
- 在帮助患者参与到减少疗效不确定性方面，仍有很大的提升空间

第6章 正确的验证疗效的试验

- 验证疗效的科学性必不可少，否则我们会将无效的治疗认为有效，反之亦然
- 对照是所有治疗性科学试验的根本
- 比较不同治疗时（或比较治疗与不治疗时），组间可比性至关重要
- 评估治疗结局必须尽量减少偏倚

第7章 考虑偶然因素

- 在评价现有证据的可信度（基于证据质量和数量）时，必须考虑偶然因素的影响

第8章 评价所有相关的可靠证据

- 单个研究很难提供充足的证据指导卫生保健中的治疗决策
- 评估不同疗法的优劣需基于针对所有相关可靠证据的系统评价
- 在单个验证疗效的研究中，必须采取方法减少偏倚和偶然因素的误导
- 不考虑系统评价的研究结果会使患者遭受不必要的伤害，还会浪费医疗卫生和研究的资源

第9章 对正确疗法的试验监管：是推动还是阻碍？

- 研究监管过于复杂

- 当前研究监管系统阻碍了那些能改善医疗卫生水平的科学试验
- 尽管有繁琐的监管准则约束研究者,但监管系统几乎未能保证申报的研究是真正必要的
- 研究监管并未有效地监控和随访已批准的研究

第 10 章 高质量、低质量和不必要的研究

- 不必要的研究是对时间、精力、金钱及其他资源的浪费,也会对患者造成潜在的伤害,有悖于伦理
- 只有在对过去所有相关研究做最新的系统评价,且结果显示有必要时,才应开展一项新的研究,而且还要在注册之后方可进行
- 最新研究产出证据后应对以往的系统评价及其他相关证据进行更新
- 很多研究是低质量的,并且其出发点让人怀疑
- 商业和学术界会扭曲研究
- 和患者息息相关的问题常常得不到解决

第 11 章 做高质量研究,人人有责

- 患者和研究人员合作有助于发现并减少治疗的不确定性
- 患者的参与能够提高研究质量
- 有时患者会无意中破坏验证疗效的科学试验
- 患者组织和药厂的关系可以歪曲治疗效果的信息
- 想要做出有效贡献,患者需更好地了解研究的基础知识,才能可容易获得无偏倚的信息
- 没有一个“正确的方法”来实现研究中的协同参与
- 患者参与应基于特定的研究目的
- 患者参与研究的方法学在不断更新

Index 索引

- abdominal aortic aneurysm screening 腹主动脉瘤筛查 36, 46
- Ablin, R.J. 40
- absolute risk reduction 绝对危险度降低率 150
- academic medicine 学术医学 126-7
- ‘active monitoring’, preferred to ‘watchful waiting’ “观察等待”更应该描述为“主动监测” 140
- adverse effects see also side-effects; unexpected bad effects 不良反应，也见“副作用”；预期之外的有害作用
- dramatic effects 显著的不良反应 70
 - hunches about 预期不良反应 81-4
- advertising 广告 97, 125-6
- advocacy groups (patients’) see patient organizations (患者) 支持小组，见“患者团体/组织”
- AIDS see HIV/AIDS 艾滋病，也称为“人类免疫缺陷综合征”
- allocation to treatment groups 分配到治疗组 73-6
- al-Razi (Persian physician) al-Razi (波斯医生) 69
- alternate allocation procedure 交替分配流程 73
- anecdotes, pitfalls of 轶事的陷阱 2, 16
- Angell, M. 125
- animal magnetism, testing 动物磁性（催眠）试验 79
- antibiotics in pre-term labour 抗生素预防早产 58-9
- aortic aneurysms 主动脉瘤
- screening 主动脉瘤的筛查 36, 46
 - surgery 主动脉瘤的手术 148
- apnoea of prematurity 早产儿呼吸暂停 57-8
- applicability of research evidence 研究证据的适用性 151-3
- aprotinin 抑肽酶 121-2
- arrhythmias, treatment of 心律失常的治疗措施 14-15, 97
- arthritis 关节炎 65-6
- Ashcroft, R. 114
- Asher, R. 66
- aspirin 阿司匹林
- finding the right dosage 确定阿司匹林的最适剂量 21, 54
 - systematic review of heart attack evidence 心脏病证据的系统评价 88-9, 92
- Atkins, H. 27
- Australia, neuroblastoma screening in 澳大利亚神经母细胞瘤筛查 34-5
- Avandia 文迪雅 7-8
-
- B12 to treat pernicious anaemia 维生素 B12 治疗恶性贫血 51, 53
- babies 婴儿
- antibiotics and bilirubin processing 抗生素和胆红素代谢 83
 - caffeine for breathing problems in premature babies 咖啡因治疗早产儿呼吸困难 57-8, 92, 111-12
 - sleeping positions, babies’ 婴儿睡姿 13-4, 84
 - unnecessary research into respiratory distress 关于呼吸窘迫的不必要研究 120

'back to sleep' campaign “躺着睡”运动 13-4, 84
bad research, examples of 低质量研究的实例 118-20
bai lin tea diet 白琳减肥茶 67
Balfour, T.G. 73
Barratt, A. 39
'believing is seeing' “境由心生” 66-7, 80
belladonna and scarlet fever 颠茄和猩红热 73
Bendectin 盐酸双环胺 xxi
benefits which don't materialize 未能达到疗效
 babies' sleeping position 婴儿睡姿 13-4
 diethylstilboestrol (DES) 己烯雌酚(DES) 15-6
 evening primrose oil 月见草油 18-20
 heart arrhythmia treatments 心律失常治疗措施 14-5
 HRT (hormone replacement therapy) 激素替代疗法(HRT) 16-8
 human albumin solution 人血白蛋白溶液 162
beta-carotene and cancer β -胡萝卜素与癌症 154-5
bias see also fair tests 偏倚, 也见“科学试验”
 confirmation bias 证实性偏倚 66
 education about concepts of 关于偏倚概念的教育 164
 ethics cause bias towards clinical use vs setting up a trial 伦理引起临床应用与开展试验的偏倚 109
 lead-time and length-time bias 领先时间偏倚与病程长短偏倚 33-4
 like with like, comparing 同类对照/相似对照 69-84
 publication bias 发表偏倚 96-7, 163, 166
 reducing bias in systematic reviews 减少系统评价的偏倚 95-6
 'salami slicing' “切香肠” 99
 spin 编造 97, 98-9
 under-reporting of research 研究的报告不全 96-7, 163
 and vested interests 偏倚和既得利益 98-9
bilirubin processing 胆红素代谢 83
birthmarks 胎记
 portwine stains 鲜红斑痣 51, 54
 strawberry birthmarks 草莓胎记 52-3
Björk-Shiley heart valves Björk-Shiley 心脏瓣膜 8-9
bleeding and tranexamic acid study, as exemplary fair test 科学试验的典型范例——出血和氨甲环酸的研究 90-1
blinding/ masking 盲法或隐藏 79-1
blood transfusions 输血 50
blood-letting tests 放血试验 69
blueprint for a better future 未来更美好的蓝图 164-7
bone marrow transplantation 骨髓移植术 28-0
Bosk, C. 62
Bowser, A. 82
Bradford Hill, A. 92

breast cancer 乳腺癌

- bone marrow transplantation 骨髓移植术 28-0
- challenging the 'more is better' mentality 挑战“越多越好”的理念 21-30
- ductal carcinoma in situ (DCIS) 乳腺导管原位癌(DCIS) 30
- Herceptin 赫赛汀 9-12, 30, 136-7
- HRT 激素替代疗法 18
- mutilating surgery 破坏性手术 23-8
- patient-researchers 作为研究人员的乳腺癌患者 133-4
- professional uncertainty 乳腺癌治疗的不确定性 59
- screening 乳腺癌筛查 30, 36-8
- systematic reviews 乳腺癌的系统评价 98
- as systemic disease 乳腺癌作为一种系统性疾病 25, 27

Brewin, T. 23

caffeine, for breathing problems in premature babies 咖啡因治疗早产儿呼吸暂停 57-8, 92, 111-12

calcium antagonists 钙拮抗剂 121

Canada 加拿大 124

cancer see also breast cancer; leukaemia; lung cancer; neuroblastoma; prostate cancer 癌症, 见“乳腺癌; 白血病; 肺癌; 神经母细胞瘤; 前列腺癌”

- and beta-carotene 癌症和β-胡萝卜素 154-5
- DES (diethylstilboestrol) causing 己烯雌酚导致癌症 16, 68, 83
- enrolment in trials 在癌症试验中登记 154
- and the 'more is better' myth 癌症治疗“越多越好”的无稽之谈 25-6

CancerBACUP CancerBACUP 组织 134

cardiovascular disease see heart problems 心血管疾病, 见“心脏问题”

carotid endarterectomy surgery 颈动脉内膜切除术 115-16

Ceox 万络的别称 5-7, 83

Chalmers, I. 54-6

chance, play of 偶然因素的影响 85-91, 98

Chapman, S. 39

chest X-rays for screening 胸部X线筛查 41-2

childbirth see pregnancy and birth 分娩, 见“怀孕和分娩”

Clark, S. 103

Clement-Jones, V. 134

clinical trials 临床试验 64-84

- cluster randomized trials 群随机试验 76
- single patient trials 单病例试验 152

Clinical Trials Directive 《临床试验准则》 107

clot-busting drugs 溶栓药 56

cluster randomized trials 群随机试验 76

Cochrane, A.L. 31

Cochrane Collaboration Cochrane协作网 131

Cochrane Database of Systematic Reviews Cochrane系统评价数据库 156

codes of practice 实践规范 107

collaboration 协作

- between patients and doctors - 'shared decision making' 患者和医生之间的协作——“共同决策” 143-7, 158-9
- between patients and researchers 患者和研究人员之间的协作 133-4, 136-41

commercial considerations see profit motive 商业考虑, 见“利益动机”

comparisons in clinical trials 临床试验的比较

- against placebos 与安慰剂比较 73, 79, 124
- areas of comparison 不同部位的比较 152
- comparing like with like 同类对照/相似对照 69-84
- against current best treatment, even in life-threatening situations 即使在性命攸关的情况下也要与当前最佳的治疗措施比较 147-8
- 'level playing field' in Lind's scurvy tests 林德坏血病试验中的“公平竞争环境” 2
- setting up comparison groups 设立对照组 72-3
- should be against current most effective treatment 临床试验应该和当前最有效的治疗措施比较 124, 147-8
- 'compassionate' release of drugs “善意地”投放药物 136

computed tomography (CT) scans 计算机断层(CT)扫描 43, 47-8

confidence intervals 可信区间 86-7

confirmation bias 证实性偏倚 66

conflicts of interests see also profit motive 利益冲突, 见“利益动机”

- call for transparency 利益冲突需要透明化 166-7
- in information sources 信息资源中的利益冲突 156
- patient organizations and the pharmaceutical industry 患者组织与药厂的利益冲突 137-9
- in systematic reviews 系统评价中的利益冲突 98-9

consent, informed 知情同意 110-13, 166

Cooper, J. 11

co-researchers, patients and clinicians as 患者和临床医生是研究合作者 133-4

cot deaths due to sleeping position 睡姿导致的婴儿猝死 13-14, 84

co-trimoxazole 复方新诺明 117-18

Crile, G. 25-6

crossover trials 交叉试验 74-5

CT scans (computed tomography) CT扫描(计算机断层扫描) 43, 47-8

cumulative meta-analysis 累积Meta分析 100, 122 *see also* systematic reviews 见“系统评价”

cystic fibrosis 囊性纤维化 128

data protection law 数据保护法 107, 108

Davies, S. 132

depression, drugs for and suicidal thoughts 给抑郁症和有自杀念头患者的药物 84, 99

DES (diethylstilboestrol) 己烯雌酚 15-16, 68, 83

diabetes 糖尿病

- Avandia 文迪雅 7-8
- insulin used for 用于糖尿病的胰岛素 50-1
- overdiagnosis 糖尿病的过度诊断 156-7

diethylstilboestrol (DES) 己烯雌酚 15-16, 68, 83

diets 饮食 67

‘disease clock’ “疾病钟” 34

diuretics 利尿剂 21

DNA testing (screening) see genetics DNA检测（筛查），见“遗传基因”

dosages, finding the right balance 找到最适的平衡剂量

- aspirin 阿司匹林 21, 54
- diuretics 利尿剂 21
- morphine 吗啡 70

double blind tests 双盲试验 79-81

dramatic treatment effects 显著的疗效 50-2, 70

drug licensing authorities 药物审批机构 124

drug pricing 药物定价 124

drug trials 药物试验

- domination of drugs over other treatment types 相对于其他类型的治疗措施，药物居主导地位 122-3
- regulation of 药物试验的监管 107

ductal carcinoma in situ (DCIS) 导管原位癌(DCIS) 30

eclampsia 子痫 56, 116-17

eczema, GLA supplements for 补充γ-亚麻酸治疗湿疹 18-20

effectiveness of treatments 治疗措施的有效性

- doubt about 质疑治疗措施的有效性 50-63
- effective treatments denied due to lack of meta-analysis 由于缺少Meta分析，有效的治疗措施遭到否认 120-1
- known effective treatments should not be withheld 不应该隐瞒已知的有效治疗措施78, 101, 120
- offering uncertain treatments as part of a trial to find out 将疗效不确定的治疗措施作为试验的一部分 60-3
- safety vs effectiveness, evening primrose oil 月见草油安全性和有效性 18-20

Eichenwald, K. 29

envelope system of random allocation 随机分配的信封系统 74

epidural analgesia 硬膜外麻醉 119-20

Epstein, S. 134

ethics 伦理学

- ethics cause bias towards clinical use vs setting up a trial 伦理引起临床应用与开展试验的偏倚 109
- ethics committees 伦理委员会 106
- proportionate review 合理的伦理审查 110
- and regulation 伦理学和监管 105
- and uncertainty 伦理学和不确定性 60-1

European Medicines Agency study 欧洲药品管理局的研究 139

Evans, I. 61

evening primrose oil 月见草油 18-20

fair tests 科学试验

- principles of 科学试验的原则 64-84
- comparing like with like 科学试验的同类对照/相似对照 69-84
- fair allocation to treatment groups 公平分配到治疗组 73
- patient involvement jeopardizing fair tests 患者参与会破坏科学试验 136-9
- people more at risk when not taking part in one 当患者不参与科学试验时，人们将承担更大的风险 59
- providing treatment as part of 将治疗措施作为科学试验的一部分 60-3
- tranexamic acid after injury, good example of fair test 科学试验的典型范例——损伤后使用氨甲环酸 90-1
- uncertain treatments - clinical use vs proper fair testing 不确定的治疗措施——临床应用和适当的科学试验 59, 62-3, 107, 108-9, 166
- why needed 为什么需要科学试验 1-3

false-positives see also overdiagnosis 假阳性，见“过度诊断” 37-8, 43

Fisher, B. 25-6

follow up, in treatment comparisons 治疗措施比较中的随访 76-7, 83-4

Franklin's Law 富兰克林法则 xx

frequencies, as best way to present statistics 反映数据的最好方式——频率 149

gamma-linolenic acid (GLA) supplements for eczema 补充γ-亚麻酸(GLA)治疗湿疹 18-20

genetics 遗传学

- genetic research, (misguided) hopes for 遗传学研究的（被误导的）期望 127-8
- genetic testing 基因检测 43-4, 153

getting better without treatment 不治疗会好转 32, 64-5, 158

ghostwriting 代笔 124-5

Gill, R. 111

Goldacre, B. 53, 108, 151, also author of Foreword page ix 也是第 ix 页序的作者

Good Medical Practice (GMC) 《医学实践规范》(GMC) 60

good research, examples of 高质量研究的实例 115-18

Goodman, S.N. 93

graphs for presentation of statistics 反映统计数据的图表 149

group randomized trials 群随机试验 76

guesswork, doctors and 医生和推测 xiv, 59

haemangioma (strawberry birthmarks) 血管瘤（草莓胎记） 52-3

Halsted mastectomies Halsted 乳房切除术 24

Harrison, J. 26, 113

healthcare IT systems 卫生保健信息技术系统 62

healthy volunteers, risks to 健康志愿者的风险 101

heart problems 心脏问题

- accidental discovery of treatment for birthmarks 胎记治疗方法的偶然发现 52-3
- aspirin 阿司匹林 54
- Avandia 文迪雅 8
- blood substitutes 血液代用品 100
- defibrillators 除颤仪 50
- heart attacks and systematic reviews 心脏病和系统评价 99-100

heart rhythm abnormalities, treatment of 心律失常的治疗措施 14-15, 97
Herceptin 赫赛汀 10, 12
HRT 激素替代疗法 17, 18
mechanical heart valves 人工心脏瓣膜 8-9
screening for risk factors 心脏病危险因素的筛查 35
synthetic oestrogen 合成雌激素 14-15
use of oxygen in heart attacks 心脏病氧疗的使用 50
vascular screening 血管筛查 46-7
Vioxx 万络 5-6, 83
heel-prick test 足跟血试验 36
Herceptin 赫赛汀 9-12, 30, 136-7
hexamethonium 六烃季铵 102
Hill, A.B. 92
hip replacements 髋关节置换 53-4, 70
'historical controls' in tests 试验中的“历史对照” 71
HIV/AIDS 人类免疫缺陷综合征或艾滋病
 HIV in children 艾滋病患儿 117-8
 patient involvement in studies 参与研究的艾滋病患者 134-5
Holland, W.W. 31
Homoeopathy 顺势疗法 xi
Hope, T. 105
hormone replacement therapy (HRT) 激素替代疗法(HRT) 16-8, 72
Horton, R. 103
HRT (hormone replacement therapy) 激素替代疗法(HRT) 16-8, 72
human albumin solution 人血白蛋白溶液 162
Huntingford, C.A. 17

Iannucci, A. 127
Imatinib 伊马替尼 51
information, balanced, need for 适度的信息需求 38, 148, 154-5
informed consent 知情同意 110-3, 166
Institutional Review Boards (IRBs) 机构审查委员会(IRBs) 106
insulin 胰岛素 50-1
'intention to treat' analysis 意向性分析 76-7
interferons 干扰素 136, 139
IQWiG website 德国药物评审机构(IQWiG)网站 156

James Lind Alliance 詹姆斯·林德联盟 131
Japan 日本
 mutilating breast surgery 破坏性乳房切除术 28
 neuroblastoma screening 神经母细胞瘤筛查 33-5
Jobling, R. 128
Johns Hopkins tragedy 约翰·霍普金斯的悲剧 102
journal supplements 杂志附录 125

Kocher, T. 82-3
Kolata, G. 29
Kushner, R. 27, 133-4
'labelling' of people with illnesses 带有疾病“标签”的人 156-7
laboratory studies, focus on at expense of people-based studies 实验室研究，关注基于人群的研究的支出 126-7,160
labour, epidural analgesia in 分娩时行硬膜外麻醉 119-20
Lantos, J. 109
large numbers, law of 大数定律 85-6
laser treatment of portwine stains 鲜红斑痣的激光疗法 51, 54
laws, and regulation of medical research 医学研究的规范与法律 107
lead-time bias 领先时间偏倚 34
length-time bias 病程长短偏倚 33
Lerner, B.H. 24
leukaemia 白血病
 chronic myeloid leukaemia and imatinib 慢性粒细胞性白血病和伊马替尼 51
 comparison between US and British children 美国和英国白血病患儿的比较 77-8
 randomized trials 白血病的随机试验 61
life-threatening conditions, trying 'anything' 性命攸关时尝试“任何方法” 147-8
like with like, comparing 同类对照/相似对照 69-84
Lind, J. 1-3, 69-70, 131
long-term side-effects 长期副作用 81
lorcainide 劳卡胺 97
lumpectomies 乳房肿瘤切除术 25
lung cancer 肺癌
 doctors would not participate in trials 医生将不参与肺癌试验 129
 lung cancer screening 肺癌的筛查 41-3
 researching treatments using historical controls 使用历史对照研究治疗措施 71
lymphoedema 淋巴水肿 30, 59

magnesium sulphate for pre-eclampsia 硫酸镁治疗先兆子痫 56, 116-17
Magpie trial Magpie 试验 117
Mammography 乳房 X 线摄影 36-8
Manson, N.C. 110
marketing-based medicine 遵循市场的药物 97,125
masking/blinding 隐藏或盲法 79
mastectomies 乳房切除术 23-8
Mastectomy Association 乳房切除术协会 134
mechanical heart valves 人工心脏瓣膜 8-9
media, role in medical misinformation 媒体在医疗误导中的作用 x
Medicines and Healthcare products Regulatory Agency (MHRA) 药品和保健品管理局(MHRA)
20, 82
mental disorders, research into 精神障碍的研究 118-19

mesmerism 催眠术 79
meta-analysis *see also* systematic reviews Meta 分析, 见“系统评价” 91, 98, 100
‘me-too’ drugs “me-too”药 123-4
Moher, D. 95
‘more is better’ myth “越多越好”的无稽之谈 21-30
Morgan, S.G. 124
morphine 吗啡 70
Morris, J.K. 35
mother’s kiss 妈妈的吻 52
motives, researchers’ 研究人员的动机 x,106
Moynihan, R. 138
multi-centre research 多中心研究 89-91
multinational trials 跨国试验 91
multiple sclerosis drugs 多发性硬化症药物 136, 139
mutilating surgery for breast cancer 乳腺癌的破坏性手术 23-8
natural progression of illnesses 疾病的自然进程 64-5
neuroblastoma screening 神经母细胞瘤的筛查 32-5
‘new’ treatments, not necessarily improved 新的治疗措施并不一定更好 1,78-9
New Zealand 新西兰
 drug pricing 新西兰的药品定价 124
 Herceptin 赫塞汀 10-11
NHS Choices NHS Choices 网站 156
NICE (UK National Institute for Health and Clinical Excellence) NICE (英国国家卫生医疗质量标准署) 英国国家卫生与临床优化研究所 10
nimodipine 尼莫地平 121
n-of-1 trials 单病例随机对照试验试验 152
NSAIDs (non-steroidal anti-inflammatory drugs) 非甾体抗炎药 (NSAIDs) 5-7,122

oestrogen, synthetic 合成雌激素 15-16
offer of screening is a medical intervention in itself 提供筛查本身就是一种医疗干预 32
O’Neill, O. 110
optimism, beneficial effects of 乐观主义效应, 见“安慰剂效应” 66-7
orthopaedics, to plaster or not to plaster 打石膏或不打石膏的骨科手术 54-6
osteoarthritis of the knee 膝关节骨关节炎
 example of shared decision making consultation 共同决策咨询的实例 145-6
 studies into things patients consider important 患者关注的重要内容的研究 122-3
overdiagnosis 过度诊断
 breast cancer 乳腺癌的过度诊断 37
 diabetes 糖尿病的过度诊断 156-7
 generally 过度诊断普遍存在 156-7
 lung cancer 肺癌的过度诊断 43
 neuroblastoma 神经母细胞瘤的过度诊断 32, 33
 prostate cancer 前列腺癌的过度诊断 39
over-reporting of studies 研究的过度报告 99

Oxman, A.D. 126

Parmar, M.S. 22

paroxetine (Seroxat) 帕罗西汀 (赛乐特) 99

patient objectives, importance of 患者意愿的重要性 119-20, 122-3, 126-7, 128-9, 165

patient organizations 患者组织

- conflicts of interests with the pharmaceutical industry 患者组织与药厂的利益冲突 137-9
- patient organizations in clinical trials 临床试验中的患者组织 132, 137-9
- patient organizations used to lobby by drug companies 患者组织被药厂利用为其游说 139
- pushing for better research 患者组织推动更高质量的研究 27

patients 患者

- coping with uncertainty 患者处理不确定性 61
- and creating better research 患者开展高质量研究 130-5
- demanding bone marrow transplantation 患者要求骨髓移植 28-30
- overdiagnosis turns people into 过度诊断使正常人变成患者 31, 33, 37, 44, 156-7
- patient activists 积极的患者 27, 133-4
- patient involvement jeopardizing fair tests 患者参与破坏科学试验 136-9
- patient-researchers 作为研究人员的患者 133-4
- shared decision making with doctors 患者与医生共同决策 143-7, 158-9
- should be offered participation in proper trial 患者应参与适当的试验 59, 62-3, 107, 108-9, 166
- people-focused research *see* patients 基于人群的研究, 见“患者”

Perkins, E. 102

pernicious anaemia 恶性贫血 51

‘pester power’ “儿童消费力” 137

Petit-Zeman, S. 59

Peto, R. 28

PHARMAC (New Zealand Pharmaceutical Management Agency) 新西兰药物管理局 10-12

pharmaceutical industry 药厂

- and doctors 药厂与医生 125
- driving research 药厂操控研究 123-9
- suppression of research results 药厂对研究结果的隐瞒 96-7

phenylketonuria (PKU) screening 苯丙酮尿症 (PKU) 36

Pirsig, R.M. ix

Placebos, role in fair tests 安慰剂在科学试验中所扮演的角色 73, 79-81, 124

‘placebo effect’ “安慰剂效应” ix, 67

plaster casts, dispute over 关于石膏固定的争论 54-6

play of chance, reducing 减少偶然因素的作用

- in studies 在研究中减少偶然因素的作用 85-91
- in systematic reviews 在系统评价中减少偶然因素的作用 98

pneumothorax (collapsed lung) treatment 气胸 (肺塌陷) 的治疗措施 70

portwine stains, laser treatment of 鲜红斑痣的激光疗法 51, 54

positive thinking, benefits of 积极思考的益处 66-7

pre-eclampsia, research into 先兆子痫的研究 56, 116-17

pregnancy and birth 怀孕与分娩

- antibiotics and bilirubin processing for newborn babies 新生儿抗生素和胆红素代谢 83
- antibiotics in pre-term labour 抗生素预防早产 58-9
- Bendectin 盐酸双环胺 xxi
- caffeine for breathing problems in premature babies 咖啡因治疗早产儿呼吸困难 57-8, 92, 111-12
- epidural analgesia 硬膜外麻醉 119-20
- pre-eclampsia 先兆子痫 56, 116-17

premature babies 早产儿

- antibiotics in pre-term labour 抗生素预防早产 58-9
- and blood oxygen levels 早产儿与血氧水平 91
- caffeine for breathing problems 咖啡因治疗呼吸困难 57-8, 82, 111-12
- unnecessary research into respiratory distress 关于呼吸窘迫的不必要研究 120

previous studies, taking account of *see also* systematic reviews 考虑先前的研究, 见“系统评价” 93-4

PRISMA statement PRISMA 声明 95

probabilities, importance of understanding 理解的可能性和重要性 xxii, 89

profit motive 利益动机

- calls for transparency 利益动机需要透明化 166-7
- drug trials more likely than e.g. physiotherapy trials 比起像理疗这样的试验, 药物试验更有可能受利益驱使 122-3
- as main driver for research 利益动机作为研究的主要驱使因素 123-9
- and pharmaceutical research 利益动机与药物研究 97, 122-3
- profit-making from intensive treatments 从过度治疗中盈利 28-9
- profit-making from screening 从筛查中盈利 40, 42, 46-7
- sponsorship of research 研究的赞助 125-6, 139
- vs patient choice 利益动机与患者的选择 131

proportionate review 合理审查 110

propranolol 普萘洛尔 52-3

prostate cancer 前列腺癌

- collaborative approach to researching treatments 研究前列腺癌治疗措施的协作方式 140-1
- relative vs absolute risk reduction statistics 相对危险度降低率与绝对危险度降低率 150-1
- screening 前列腺癌的筛查 38-41

prostate-specific antigen (PSA) 前列腺特异性抗原 38-9

proving a negative, impossibility of 阴性结果验证的不可能性 xxi

psoriasis 银屑病 128, 129

psychiatric disorders, research into 精神障碍的研究 118-20

psychology of responding to treatment 对治疗措施反应的心理 66-7

publication bias 发表偏倚 97, 163, 166

publication of results, importance of 结果发表的重要性 163, 166

PubMed Health PubMed Health 网站 156

pus, drainage of 脓液引流 50

radical mastectomies 根治性乳房切除术 24-8, 134

randomized trials 随机试验

- and breast cancer surgery 随机试验和乳腺癌手术 26-7
- as key feature of fair tests 随机试验作为科学试验的主要特点 74-6
- cluster randomized trials 群随机试验 76
- methods of randomization 随机的方法 74-6
- not used in material supporting adverts 随机试验不作为支持广告的材料 125-6
- random allocation ('randomization') 随机分配 ("随机化") 26, 73-6
- 'randomized cross-over trials' "随机交叉试验" 74-5
- telephone/computer randomization 电话或计算机随机化 74

Rayleigh, Lord Rayleigh 勋爵 93-4

recording treatments, standardization of 治疗措施记录的标准化 62

recovery without treatment (natural progression) 未经治疗就痊愈 (自然过程) 32, 64-5, 158

recruitment of trial participants 试验参与者的招募 29-30, 111-12, 154

regulation of medical research 医学研究的规范 105-14

relative risk reduction 相对危险度降低率 150

reliability of information 信息的可靠性 154-6

research design, importance of good 高质量研究设计的重要性 162-3

Research Ethics Committees (RECs) 研究伦理委员会(RECs) 106

research priorities, distorted 扭曲的研究优先性 122-9

research protocols, departures from 偏离研究方案 77-8

- and regulation of 研究方案的规范 106
- and systematic reviews of 研究方案的系统评价 96

respiratory distress in premature babies 早产儿呼吸窘迫 120

right questions, asking 提出正确的问题 161-2

risk 风险

- concept not taught in schools 风险是在学校没被讲授的概念 xx
- probabilities, importance of understanding 风险理解的重要性和可能性 xxii, 89
- relative and absolute risk reduction 相对和绝对危险度降低率 150
- risk ratios 风险率 85-6

Robinson, K.A. 93

rofecoxib (Vioxx) 罗非昔布 (万络) 5-7, 83

rosiglitazone (Avandia) 罗格列酮 (文迪雅) 7-8

Ross, N. 2

rosuvastatin 罗素伐他汀 81

Sackett, D.L. 126

safety vs effectiveness, evening primrose oil 月见草油的安全性和有效性 18-20

'salami slicing' "切香肠" 99

Salzburg Statement 萨尔茨堡宣言 158-9

scarlet fever and belladonna 猩红热和颠茄 73

schizophrenia 精神分裂症 118-9

Schwartz, L.M. 42, 157
scoring systems, deficiencies in 评分系统的缺陷 129
screening 筛查
 generally 普遍筛查 31-49
 abdominal aortic aneurysm screening 腹主动脉瘤的筛查 36, 46
 breast cancer screening 乳腺癌的筛查 30, 36-8
 causing overdiagnosis 筛查造成过度诊断 32, 33, 37, 39, 43
 chest X-rays for screening 胸部 X 线筛查 41-2
 clearly beneficial 筛查确实有益 35-6
 criteria for assessing value of 评估筛查价值的标准 45
 false-positives 假阳性 37-8, 43
 genetic testing 基因测试 43-4
 lung cancer screening 肺癌筛查 41-3
 neuroblastoma screening 神经母细胞瘤筛查 32-5
 offering screening is a medical intervention in itself 提供筛查本身就是一种医疗干预
31-2
 phenylketonuria (PKU) screening 苯丙酮尿症(PKU)筛查 36
 profit-making from screening 从筛查中获益 40, 42, 46-7
 prostate cancer screening 前列腺癌筛查 38-41
 vascular screening 血管筛查 46-7
 weighing benefits and harms 权衡利弊 35-47
scurvy tests 坏血病试验 1-3, 69-70
self-limiting diseases 自限性疾病 32, 64-5, 158
'selling' screening “推销”筛查 42, 46-7
Seroxat 赛乐特 99
shared decision making 共同决策 143-7, 158-9
side-effects 副作用
 fair tests needed to find out about 找出副作用所需要的科学试验 67-8
 long-term side-effects 长期不良反应 81
 prostate cancer treatments 前列腺癌治疗措施的副作用 38
 synthetic oestrogen 合成雌激素的副作用 16
 unexpected discoveries (Viagra) 预期之外的发现(万艾可) 68
significant differences (statistical concept) 显著差异(统计概念) 88-9
sildenafil (Viagra) 西地那非(万艾可) 68
Silverman, W.A. xix, 4
single patient trials 单病例试验 152
single studies, rarely sufficient 单个研究远远不够 92
sleeping positions, babies' 婴儿睡姿 13-4, 84
Smart Health Choices 《聪明的健康选择》 156
social scientists, included in clinical research teams 包括社会学家的临床研究团队140-1
Spiegelhalter, D. 89
spin 编造 97, 98-9
spiral CT scans 螺旋CT扫描 43
Spock, Dr. Benjamin Benjamin Spock博士 13, 84

sponsorship of research 研究的赞助 125-6, 139

spontaneous regression/remission 自然消退或缓解 32, 64-5

statistics 统计学

- confidence intervals 可信区间 86-7
- law of large numbers 大数定律 85-6
- patients' understanding 患者对统计学的理解 148-1
- statistical significance 统计学意义 88-9

stem cell rescue 干细胞移植治疗 28-30

steroids 类固醇

- and premature labour 类固醇与早产 93, 120
- for traumatic brain injury 类固醇治疗创伤性脑损伤 56, 90-1, 92, 103-4, 161-2

stillbirths, synthetic oestrogen doesn't prevent 合成雌激素不能预防死产 15-6

Stockler, M. 39

strawberry birthmarks 草莓胎记 52-3

strokes 卒中

- aspirin 阿司匹林 54
- Avandia 文迪雅 8
- carotid endarterectomy surgery 颈动脉内膜切除术 115-16
- clot-busting drugs 溶栓药 56
- collaborative approach to researching emergency treatments 研究紧急治疗措施的协作方式 139-40
- drugs tests (wasted resources) 药物试验 (浪费的资源) 102-3
- example of good research 高质量研究的实例 115-16
- example of unnecessary research 不必要研究的实例 120-1
- HRT 激素替代疗法 18
- screening for risk factors 危险因素的筛查 35
- surgical unblocking of blood vessels 血管疏通手术 76
- vascular screening 血管筛查 46-7
- Vioxx 万络 5-6

subjectivity in assessing outcomes 评估结局的主观性 80

suppression of research results 研究结果的扣留 96-7

surgery 手术

- aneurysm surgery 动脉瘤手术 148
- and blind tests 手术和盲法试验 80
- carotid endarterectomy surgery 颈动脉内膜切除手术 115-16
- mutilating surgery for breast cancer 破坏性乳腺癌手术 23-8
- surgical unblocking of blood vessels 血管疏通手术 76
- systematic reviews of breast cancer surgery studies 乳腺癌手术研究的系统评价 28

survival times, skewed by date of diagnosis 因为诊断时间而歪曲的存活时间 34

Susser, M. Xxi

systematic reviews 系统评价 92-104

- generally 系统评价很常见 xi, 94-104
- of bone marrow transplantation treatments 骨髓移植治疗的系统评价 30
- of breast cancer screening studies 乳腺癌筛查研究的系统评价 37

of breast cancer surgery studies 乳腺癌手术研究的系统评价 28
Cochrane Database of Systematic Reviews Cochrane 系统评价数据库 156
current practice in using 系统评价在当前临床实践的应用 164
disseminating 系统评价的传播 165
if done properly reduce unnecessary research 如果恰当完成系统评价，可减少不必要的研究 120-2
of nimodipine trials 尼莫地平试验的系统评价 121
systemic therapies 系统疗法 25

tardive dyskinesia 迟发性运动障碍 118-19
telephone/computer randomization 电话或计算机随机化 74
test, fair (see fair tests) 科学试验，见“科学试验”
thalidomide 沙利度胺 4-5, 70, 83
Thornton, H. 144
thyroid removal, unanticipated effects 甲状腺切除术预期之外的作用 82-3
tobacco 烟草 65
toxaemia, research into 毒血症的研究 56, 116-17
tranexamic acid, as exemplary fair test 科学试验的典型范例——氯甲环酸 90-1
trastuzumab (Herceptin) 曲妥珠单抗（赫赛汀） 9-12, 30, 136-7
treatment groups, setting up 设立治疗组 73
trials, as treatments 治疗性试验 60-1

uncertainty 不确定性 50-63
about treatments 治疗措施的不确定性 54-6
confidence intervals 可信区间 86-7
encouraging honesty about 鼓励诚实面对不确定性 165
facing up to 勇于面对不确定性 56, 60, 165
inevitability of 不确定性的必然性 xx
professionalism of addressing 处理不确定性具有专业性 57, 63, 112
uncertain treatments - clinical use vs proper fair testing 不确定的治疗措施——临床应用和恰当的科学试验 59, 62-3, 107, 108-9, 166
under-reporting of research 研究的报告不全 96-7, 163, 166
unexpected bad effects see also side-effects 预期之外的有害作用，见“副作用”
Avandia 罗格列酮 7-8
blindness in babies caused by incorrect use of supplemental oxygen 供氧不正确导致婴儿失明 4
hunches about 预期不良反应 81-4
mechanical heart valves 人工心脏瓣膜 8-9
thalidomide 沙利度胺 4-5
Vioxx 万络 5-7

unfair tests, schizophrenia studies as examples of 精神分裂症研究是不科学试验的实例 119
universities and clinical research 大学和临床研究 126-7
unnecessary research 不必要的研究 120-2
unnecessary treatments, avoiding see also overdiagnosis 避免不必要的治疗措施，见“过度诊

断”156-7

untested theories, harm of 未经验证的理论造成的伤害

babies' sleeping position 婴儿睡姿 13-4

Diethylstilboestrol (DES) 己烯雌酚(DES) 15-6

evening primrose oil 月见草油 18-0

heart arrhythmia treatments 心律失常的治疗措施 14-5

HRT (hormone replacement therapy) 激素替代疗法(HRT) 16-8

human albumin solution 人血白蛋白溶液 162

USA 美国

AIDS drugs trials 艾滋病药物试验 135

Avandia 文迪雅 8

bone marrow transplantation 骨髓移植术 28-9

Institutional Review Boards (IRBs) 机构审查委员会 (IRBs) 106

patient activists 积极的患者 133-4

and prostate cancer screening 前列腺癌的筛查 39-40

treatment for heart arrhythmias 心律失常的治疗措施 15, 97

US doctors' opinion of plaster casts compared with British 英美医生对于打石膏观点的比较 54-6

ventricular fibrillation 心室纤维性颤动 50

Veronesi, U. 27

vested interests 既得利益 98-9

Viagra 万艾可 68

Vioxx 万络 5-7, 83

Vitamin B12 to treat pernicious anaemia 维生素B12治疗恶性贫血 51,53

Warlow, C. 47

wastage in medical research 医学研究中的浪费 160-1, 166

'watchful waiting', replaced with term 'active monitoring' 将“观察等待”替换为“主动监测” 140

Weatherall, D. 127

weight reduction diets 减肥饮食 67

Welch, H.G. 157

Westgate, B. 134

Wikipedia 维基百科 155

Wilson, P.M. 137

wishful thinking 心怀希望 66-7

withholding of effective treatments 有效的代扣疗法 78, 101,120

Woloshin, S. 42, 157

World Health Organization, and screening test criteria 世界卫生组织和筛查试验标准 45

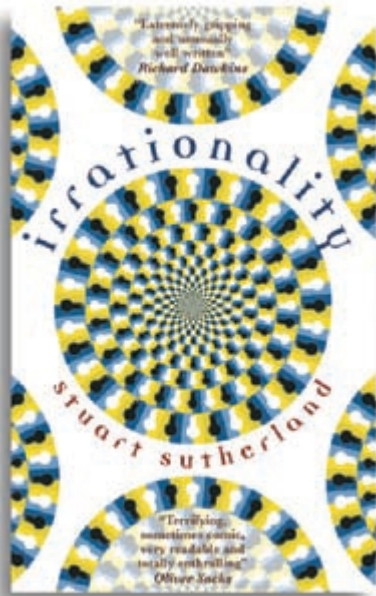
World Medical Association Declaration of Helsinki 世界医学协会的《赫尔辛基宣言》 107

Yellow Card Scheme 黄卡计划 82

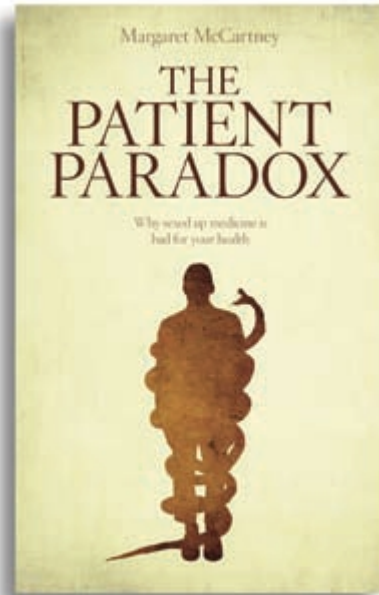
zidovudine 齐多夫定 135



PINTER & MARTIN
PUBLISHERS
freedom to think



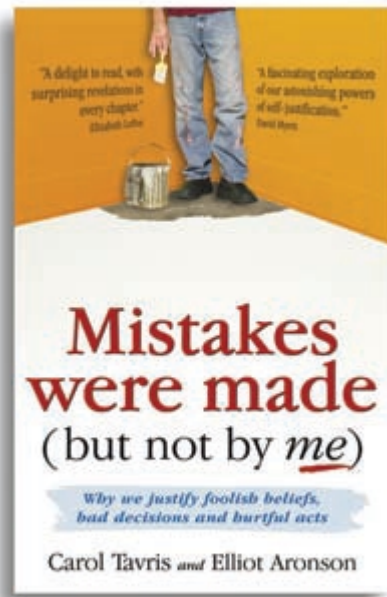
Irrationality
《非理性》
Stuart Sutherland



The Patient Paradox
《患者的矛盾》
Margaret McCartney



Childbirth in the Age of Plastics
《塑料时代的分娩》
Michel Odent



Mistakes were made (but not by me)
《不是自己犯的错》
Carol Tavris & Elliot Aronson

验证疗效的科学试验

高质量研究促进高质量卫生保健

第二版

如何知道某项治疗措施是否有效？证据有多么可靠？如何确保验证治疗措施的研究最大程度地满足了患者的需求？《验证疗效的科学试验》一书即以生动、丰富的形式回答了诸如此类的问题。本书附有大量形象的实例，无论患者还是专业人员阅读本书后都将深受启发。

基于第一版的成功经验，第二版内容进行了全面的修订和更新。该版加入了疾病的筛查的内容，解释了为什么早期诊断未必更好，其内容发人深省；并有一个新章节揭示了监管下的研究如何仍能违背患者的最大利益；另有一新章节介绍了充足的研究证据如何能使卫生保健实践模式做到让患者与临床医生共同决策。

《验证疗效的科学试验》一书鼓励每一个人加入改进现有研究、促进未来医疗的行列，并提出了切实可行的办法，愿患者和医生携手并进。

封面设计：Klor

登陆 www.pinterandmartin.com 查看更多精彩图书

卫生/医疗	
www.pinterandmartin.com	
建议零售价：	
	
英国 9.99 英镑	联系本书
美国 17.95 美元	
Pinter&Martin 有限公司出版	

ISBN 978-1-905177-48-6



9 781905 177486