

[Home](#)[目录页](#)[jameslindlibrary.org](#)

## 卫生保健治疗措施的公平试验

过去半个世纪中，卫生保健对人们延年益寿并免遭严重健康问题困扰的机会产生了实质的影响。据估计，在增加的预期寿命中，有三分之一到一半要归功于卫生保健，并且，卫生保健还使人们免受慢性健康问题困扰的平均时间增加了5年（Bunker等，1994年）。尽管如此，公众本应该，而且仍然可以从投入的意在改进健康的大量资源中，获取更高的价值回报。而且，过去发生的治疗灾难，一些本可以避免，还有一些应在未来加以避免。

创建詹姆斯·林德图书馆（James Lind Library）的目的，是促进公众对卫生保健治疗措施的公平试验及其演变过程的了解。

治疗措施疗效的误导性说法可谓随处可见，所以所有人都应该了解有关疗效的说法是否真实可信。没有这方面的认识，就可能轻易把某些本来无效的治疗措施断定为有效，或者把某些本来有效的治疗措施断定为无效。

治疗措施的公平试验是采取措施，通过减少[偏倚](#)和[机遇因素](#)带来的误导影响，获得有关疗效的可靠信息的试验过程。如果忽视了对治疗措施进行公平试验的必要性，那么人们就会遭受不必要的痛苦和死亡。

詹姆斯·林德图书馆中的解释性文章的撰写目的，是为了帮助人们更广泛地理解治疗措施公平试验的必要性，以及它们包含的内容。你可以通过点击下面标有下划线的词语阅读到每篇文章，或者你也可以从[目录页](#)选择。如果你想下载所有文章，以便把它们一起打印出来脱离屏幕进行阅读，[请点击这里](#)。

之所以[需要进行公平试验](#)，是因为许多例子表明，如果治疗决定不考虑可靠性证据，人们很容易无辜受到伤害。

[公平试验的原则](#)至少有一千年以上的历史沿革——今天还在继续发展。

要判断治疗效果的[真正不确定性](#)，必须进行[对照](#)。公平治疗措施对照须避免[偏倚](#)，无论是[所比较的不同人群之间的差异](#)，还是[治疗结果评估方法之间的差异](#)。对[意料之外疗效](#)的可靠识别尤其具有挑战性。

[对不带偏倚的对照进行解释](#)并不总是直截了当的。有时治疗效果被忽视了，因为[想要接受的治疗和实际接受的治疗之间存在差异](#)。[机遇因素](#)也可能产生误导作用。

治疗措施的公平试验必须考虑到所有相关证据。准备所有相关证据的系统评价，可以最大限度减少[带有偏倚的报告](#)以及[对现有证据带有偏倚的选择](#)所带来的影响。一种称为[Meta分析](#)(meta-analysis)的统计方法可能有助于在系统评价中避免[机遇因素](#)带来的误导。

卫生保健治疗措施的公平试验需要对所有相关的可靠证据进行[最新的系统评价](#)。即便采用最新的系统评价，留神观察偏倚和“润色之辞”也很重要。这些情况导致的结果可能是要进行单独的审查，这种审查可能要处理同样的问题，得出的却是矛盾的结论。

总之，詹姆斯·林德图书馆存有下列文章：

[为什么需要进行公平试验](#)

[为什么对照必不可少](#)

[为什么对照必须解决真正的不确定性](#)

[避免带有偏倚的对照](#)

[被比较对象之间的差异](#)

[治疗结局评价方式的差异](#)

[解释无偏倚对照](#)

[预期接受的治疗和实际接受的治疗之间存在的差异](#)

[考虑机遇因素](#)

[识别治疗措施的非预期效应](#)

[对所有相关证据的系统评价](#)

[处理对现有证据带有偏倚的报告](#)

[避免对现有证据带有偏倚的选择](#)

[利用Meta分析减少机遇因素效应](#)

[对所有相关的可靠证据进行最新的系统评价](#)

这些解释性文章引用了詹姆斯·林德图书馆藏有的丰富说明材料。可以通过点击文章中有下划线的链接或图片获取。

在向“詹姆斯·林德图书馆([www.jameslindlibrary.org](http://www.jameslindlibrary.org))”做出明确致谢的前提下，这些文章中的内容可以复制并用于非商业目的。

文章中的资料也被Evans、Thornton和Chalmers撰写的“《验证治疗措施：高质量研究促进高质量卫生保健》”收录——该书共有100页，由大英图书馆于2006年出版。ISBN 0-7123-4909-X。

## TESTING TREATMENTS BETTER RESEARCH FOR BETTER HEALTHCARE

How do we know whether a particular drug, therapy or operation really works, and how well? How reliable is the clinical evidence? Are clinical trials truly unbiased? And is current research fully focussed on the real needs of patients? Such timely and pressing questions are raised and resolved in this probing inquiry into modern clinical research, with far-reaching implications for daily medical practice and patient care. What emerges is the surprising truth that clinical research is neither as unbiased, nor as relevant as patients have every right to expect, but that everyone – patients, doctors and researchers – can do much to change current practice and achieve better healthcare.

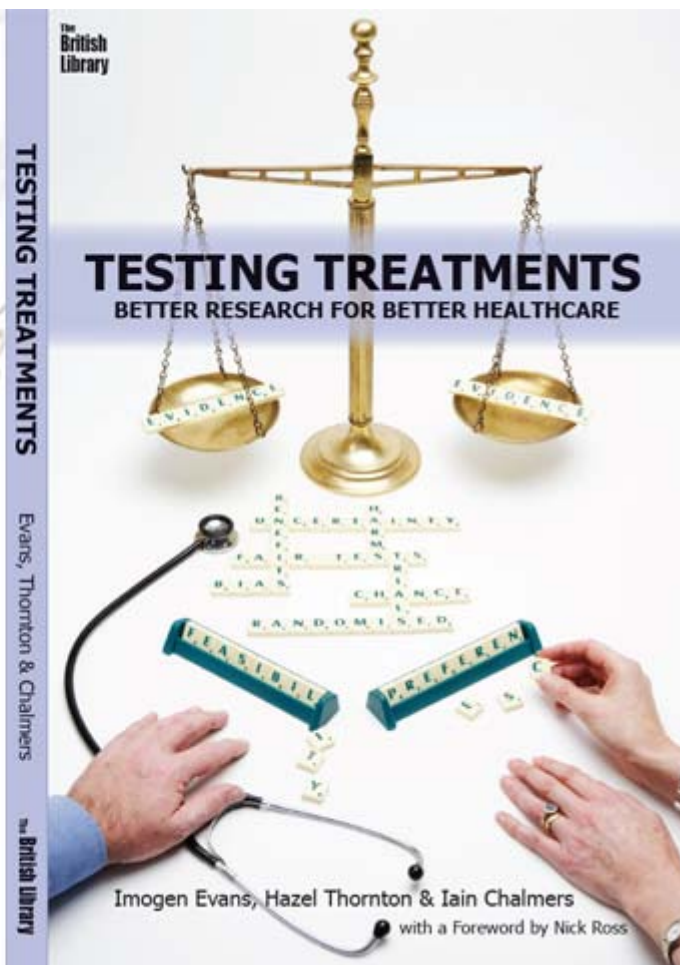
Aimed at both patients and professionals, Testing Treatments builds a lively and thought-provoking argument for better, more reliable, more relevant research, with unbiased or 'fair' trials, and explains how patients can work with doctors to achieve this vital goal.

Expertly and thoroughly researched, but never dry or dull, the fast-moving commentary, spanning the gamut of illness and therapy – from mastectomy to thalidomide – explores a vast range of revealing case-studies, enlivened throughout by entertaining anecdotes and vivid eyewitness accounts drawn from the direct experience of patients, practitioners and researchers.

E, V, I, D, E, N, C, E

The  
British Library

Cover image & design by  
**Iain Chalmers**  
[www.theofolk.com](http://www.theofolk.com)



**Cite as:**

Editorial commentary (2007). 卫生保健治疗措施的公平试验. The James Lind Library ([www.jameslindlibrary.org](http://www.jameslindlibrary.org)).

**Next essay:**

[为什么需要进行公平试验](#)

**Select other essay:**

[为什么需要进行公平试验](#)

[Home](#)

[目录页](#)

[Comments](#)

[Home](#)[目录页](#)[jameslindlibrary.org](#)

## 为什么需要进行公平试验

### 努力做到趋利避害

我们为什么需要对卫生保健的治疗措施进行公平试验？几个世纪以来，医生们难道没有对患者“竭尽所能”？遗憾的是，因为医生和其他卫生专业人员在做出治疗决定时，没有考虑我们如今所认为的关于疗效的可靠证据，因此仍有不少给患者带来伤害的例子。事后分析，在卫生保健领域，卫生专业人员在大多数情况下（并非所有情况下）无意伤害了他们的病人，这种情况发生的规模有时非常之大（[请点击此处查看例子](#)）。事实上，患者本人有时也会伤害其它患者。凭借未经验证的理论和有限的个人经验，他们可能会鼓励他人使用到头来证明有害的治疗措施。现在的问题不是我们是否要责备这些人，而是是否可以减少未经充分验证的治疗措施带来的有害后果。事实上可以在很大程度上减少。



承认某些治疗措施有时候带来的结果可能会是弊大于利，这是减少意外伤害的一个先决条件（Gregory 1772; [Haygarth 1800](#); [Fordyce 1802](#); [Behring 1893](#)）。然后，我们必须有勇气承认治疗效果的不确定性，并推动对治疗措施进行验证，以充分减少不确定性。这种试验就是公平试验。



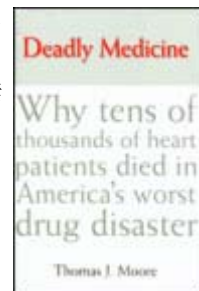
### 为什么治疗效果的理论必须在实践中加以验证

人们受到伤害，常常是因为仅仅根据如何治疗疾病的理论采取了治疗措施，而未对这些理论在实践中的使用效果进行验证。例如，人们数百年来都认为疾病是由于“体液失衡”造成的，让病人放血和腹泻、呕吐并使用鼻吸药，以为这样可以终结假想中的不平衡状态。但早在17世纪，一名独立行医的佛兰德斯医生就大胆挑战了当时的医学权威，建议对他们那些令人毛骨悚然的治疗措施的结果进行公平试验，评估其理论的正确性（[Van Helmont 1662](#)）。

到19世纪初，英国军医已开始披露放血治疗“发烧”带来的有害后果（[Robertson 1804](#); [Hamilton 1816](#)）。几十年后，这种疗法也受到了一名巴黎医生的挑战（[Louis 1835](#)）。然而在20世纪初，美国波士顿的正统医生却由于没有使用放血疗法来治疗肺炎依然被判定为失职（[Silverman 1980](#)）。事实上，一位世界上最具影响力的医学权威威廉姆·奥斯勒（William Osler）爵士对推荐使用未经证实的疗法一向都显得小心谨慎，他却给读者建议说：“在过去几十年中我们放的血肯定太少了。肺炎是可以及时通过及时的静脉切开（放血）术挽救生命的疾病之一。要有效果就应当及早放血。对于一个血液充盈的健康人来说，如果发高烧并且有洪脉，放血二三十盎司从各方面来说都是有益的”（Osler 1892）。



尽管有些人至少在一千年前就认识到有必要在实践中验证理论的正确性（[Ibn Hindu 10-11世纪](#)），但是这个重要的原则往往仍被忽视了。例如，根据未经验证的理论，颇具影响力的美国儿童健康专家本杰明·斯波克（Benjamin Spock）在其畅销书《斯波克育儿经》（Baby and Child Care）一书中告知读者，婴儿仰卧睡眠的缺点之一，就是如果出现呕吐，他们更有可能发生窒息。斯波克医生因此建议他的千百万读者鼓励婴儿以俯卧姿势睡觉（Spock 1966）。我们现在知道，这一在理论上似乎合理的建议，已导致成千上万的婴儿猝死（Gilbert等 2004）。



在有心脏病的人群中使用预防心律异常的药物，这在实践中使用未经验证理论的危险性提供了另一个例子。由于心律异常与心脏病发作后更高的早期死亡风险有关联，理论上这些药物可能减少这种早期死亡。但是，仅仅因为一个理论貌似合理并不意味着它一定是正确的。在这些药物得到批准并在实践中应用数年之后，人们发现它们实际上增加了心脏病发作后猝死的风险。实际上，据估计，在20世纪80年代使用这种药物的高峰期，仅仅在美国每年就可能使得70,000人死亡（Moore 1995）——比美国人在越南战争中死亡的总人数还要多。

另一方面，迷信理论思想将其作为实践指南也造成了一些治疗措施不恰当地遭到拒绝，因为研究人员不相信这些措施具有效用。例如，基于动物研究结果的理论有时可以正确预测人体治疗试验的结果，但事实并非总是如此。根据小白鼠试验的结果，有些研究人员坚信为心脏病发作超过6小时的患者注射溶血栓药物没有任何意义。要不是这类患者参加了关于这类药物的一些公平试验，我们可能不会知道可以从这一治疗中获益（溶栓治疗试验者协作小组，1994）。

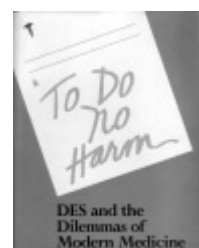


临床实践或实验室观察和动物研究可能显示特定治疗措施会惠及或不会惠及患者；但这类例子和其它例子都表明，有必要利用公平试验来确定这些治疗措施实际上是利大于弊，还是弊大于利。

### 为什么验证医学治疗措施必须用公平试验

未能在实践中检验治疗措施的理论基础并不是导致治疗悲剧的唯一可预防原因。发生这些悲剧还因为用于评估疗效的试验方法不够可靠，或者存有误导。公平试验要求采取措施，降低受到[偏倚](#)和[机遇因素](#)的影响而给我们带来误导的可能性。

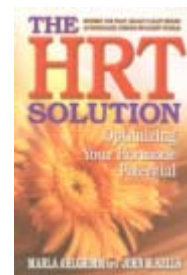
例如，在20世纪50年代，相关理论以及带来不可靠证据的不良对照试验提示，以往具有流产或死胎病史的孕



## 为什么需要进行公平试验

妇服用一种合成性激素己烯雌酚（diethylstilboestrol, DES），可增加随后的怀孕获得成功结果的可能性。尽管公平试验提示使用DES属于无效理论，又缺乏可靠证据，但通过大张旗鼓的市场营销，在接下来的几十年里，医生向数百万孕妇开出了DES处方。由此带来的后果是灾难性的：在服用了DES的妇女中，有些人的女儿患上了阴道癌，另一些孩子遇到了其它健康问题，包括生殖器官畸形和不育症（Apfel和Fisher 1984）。

由于未对治疗措施进行适当验证而导致的问题还在不断发生。再则，由于不可靠的证据和积极的市场营销，成千上万妇女被说服采用激素替代疗法（HRT），不仅因为它可以减缓令人厌烦的更年期综合症，而且还因为据称它还会降低心脏病发作和中风的可能性。当用公平试验来评估这些说法时，结果表明，激素替代疗法不仅没有降低心脏病发作和中风的风险，反而增加了这些危及生命的症状的风险，此外，还可带来其它不良影响（McPherson 2004）。



这些需要关于治疗措施的公平试验的例子，只揭示了众多疗法可能弊大于利的一小部分。必须提高公众对于治疗措施的公平试验的认识，再辅以适量的怀疑主义，从而我们都能够用更具批判性的眼光来评估关于疗效的说法。这样，我们所有人都将变得更有能力判断哪种疗法可能利大于弊。

关于治疗措施公平试验的原则千百年来一直在发生演变——它们会继续演变下去。

**Cite as:** Editorial commentary (2007). 为什么需要进行公平试验. The James Lind Library ([www.jameslindlibrary.org](http://www.jameslindlibrary.org)).

**Next essay:** [为什么对照必不可少](#)

**Select other essay:**

[为什么需要进行公平试验](#)



## 参考文献

Apfel RJ, Fisher SM (1984). To do no harm: DES and the dilemmas of modern medicine. New Haven, Ct: Yale University Press.

Behring, Boer, Kossel H (1893). Zur Behandlung diphtheriekranker Menschen mit Diphtherieheilserum. Deutsche Medicinische Wochenschrift 17: 389-393.

Fibrinolytic Therapy Trialists' Collaborative Group (1994). Indications for fibrinolytic therapy in suspected acute myocardial infarction: collaborative overview of early mortality and major morbidity results from all randomised trials of more than 1000 patients. Lancet 1994; 343: 311-322.

Fordyce G (1802). A second dissertation on fever. London: J Johnson

Gilbert R, Salanti G, Harden M, See S (2005). Infant sleeping position and the sudden infant death syndrome: systematic review of observational studies and historical review of recommendations from 1940 to 2002. International Journal of Epidemiology 34: 874-87.

Gregory J (1772). Lectures on the duties and qualifications of a physician. London: Strahan and Cadell.

Hamilton AL (1816). Dissertatio Medica Inauguralis De Synocho Castrensi (Inaugural medical dissertation on camp fever). Edinburgh: J Ballantyne.

Haygarth J (1800). Of the imagination, as a cause and as a cure of disorders of the body: exemplified by fictitious tractors, and epidemical convulsions. Bath: R. Crutwell.

Ibn Hindu (10th-11th century CE; 4th-5th century AH). Miftah al-tibb wa-minhaj al-tullab [The key to the science of medicine and the students' guide].

Louis PCA (1835). Recherches sur les effets de la saignée dans quelques maladies inflammatoires et sur l'action de l'émétique et des vésicatoires dans la pneumonie. Paris: Librairie de l'Académie royale de médecine

McPherson K (2004). Where are we now with hormone replacement therapy? BMJ 328: 357-358.

Moore TJ (1995). Deadly Medicine. New York: Simon and Schuster.

Osler W (1892). Principles and Practice of Medicine. London: Appleton, p 530.

Robertson R (1804). Observations on the diseases incident to seamen, 2nd edn. Vol. 1, London: for the author.

Silverman W (1980). In: Chalmers I, McIlwaine G (eds). Perinatal Audit and Surveillance. London: Royal College of Obstetricians and Gynaecologists, 1980:110.

Spock B (1966). Baby and Child Care. 165th printing. New York: Pocket Books, pp 163-164.

van Helmont JB (1662). Oriatrike, or physick refined: the common errors therein refuted and the whole are reformed and rectified [translated by J Chandler]. Lodowick-Loyd: London, p 526.

[Home](#)

[目录页](#)

---

[Comments](#)

[Home](#)[目录页](#)[jameslindlibrary.org](#)

## 为什么对照必不可少

### 某种疗法是否比自然和时间更好？

患者及医护专业人士总是希望治疗措施会有所帮助。这种乐观的期待可能会对每个人对医疗保健的满意度产生非常积极的效果，正如英国医生理查德·阿舍（Richard Asher）在他写给医生的一篇文章中所指出的：

“如果你能够坚定地相信你所做的治疗，尽管对照试验表明它一点作用都没有，但是你得到的结果会更好，你的患者会更好，你的收入也会更高。我相信这说明了我们的业界一些天赋不高但更具轻信的一些人所获得的显著成就，也说明了对注重时尚而且成功的医生惯于展示统计数据 and 对照试验的极度厌恶。”（Asher 1972）

人们经常没有接受任何治疗，也能从病中康复。19世纪时有有效的治疗方法还非常少，当时奥利弗·温德·霍姆斯（Oliver Wendell Holmes）表示（[Holmes 1861](#)）：“我坚信，如果现在使用的所有药品都沉入海底，人类将变得更好——而鱼类将陷入悲惨的境地。”

在验证治疗措施时，必须考虑如果不加以治疗疾病会如何发展以及产生何等结果：治疗可能改善病情，也可加重病情。治疗的效果很可能是由于自然的作用而得以改善，在过去几个世纪里，作家们已经注意到有必要对这种说法持有怀疑态度。换种说法就是，“如果你用自然疗法来对付流感，你可能会在一周内康复；但如果你去看医生，那么只要七天时间就可痊愈。”

### 安慰剂效应

由于认识到大多数疾病都可以自愈，医生们有时会开处方采用无作用疗法（inert treatments），希望他们的病人可以从中得到心理效益——也就是所谓的安慰剂效应。相信某种治疗有助于缓解症状的患者，尽管该治疗事实上并没有产生物理效果，病人可能会感觉好了不少。

医生们认识到使用安慰剂的重要性已经有几百年的历史了（[列出相关记录](#)）。例如，威廉·卡伦（William Cullen）早在1772年就提到他使用了一种安慰剂（[Cullen 1772](#)），而在19世纪提到安慰剂的案例不断增加（[Cummings 1805](#)；[内务部, 1832](#)；[Forbes 1846](#)）。由于奥斯汀·弗林特（Austin Flint）认为正统的药物治疗篡夺了“自然”的作用，他向13例风湿病患者提供了“安慰剂治疗”，其中包括高度稀释的苦木树皮提取物。结果是“病人的病程发展十分有利，通常保证病人的整个身心得到了治疗”（[Flint 1863](#)）。在伦敦盖伊医院，威廉·魏希·古尔（William Withey Gull）在治疗21例风湿热患者时，“大部分用的是薄荷水”，进而得出了类似的结论（[Sutton 1865](#)）。20世纪初，威廉·里弗斯（William Rivers）具体探讨了心理介导的治疗效果（[Rivers 1908](#)）。



### 对照的必要性

人们对于自愈力和安慰剂效应的认识已长达多个世纪，同样，人们也很早就认识到有必要进行对照，以评估自然作用和心理介导影响之外的治疗效果。有时候关于疗法的对照是在人们脑海中作出的：他们会有一个印象，与以前对治疗的反应相比，他们或其他人对某种新疗法的反应有所不同。例如，一位法国军医安布罗兹·帕雷（Ambroise Paré）曾做出过结论，用滚烫的油来治疗战争伤口（这是常见的做法）可能是有害的。当时供应的油用光了，他的病人反而比平常康复得更快，他因而得出了这个结论（[Paré 1575](#)）。



大多数时候，这类印象应随后进行正式调查，可能最先从对医护记录的分析开始着手。那么这种印象可能会招致进行仔细对照。如果仅仅把印象本身作为提出治疗建议并做出决定的指导，就会具有危险。

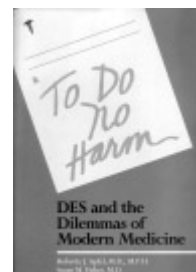
### 治疗的显著效果和适度效果

基于印象或分析相对局限的治疗对照，只在极少数情况下当治疗效果显著时才可获得可靠的信息（[点击这里列出相关记录](#)）。例子包括用鸦片缓解疼痛（[Tibi 2005](#)），做好新生儿卫生以预防破伤风（[Schleisner 1849](#)），氯仿用于麻醉、胰岛素用于治疗糖尿病（[Banting等 1922](#)）、食用肝以治疗恶性贫血（[Minot和Murphy 1926](#)），磺胺类药物用于防止产后感染（[Colebrook和Purdie 1937](#)），链霉素用于治疗结核性脑膜炎（[MRC 1948](#)），肾上腺素用于治疗危及生命的过敏反应（[McLean-Tooke等 2003](#)），以及用经过基因设计的药物治疗某些罕见类型的白血病（[Druker等 2001](#)）。不过，大多数医学治疗都不会产生像这些例子一样的显著效果，除非小心避免带有偏倚的对照，否则很可能导致关于疗效的危险错误结论。

### 对今昔的治疗进行比较

在一定程度上，由于依赖基于以往经验所做的偏倚对照，医生和妇女才相信己烯雌酚（DES）药物可以减少流产和死胎的风险。从来没有任何来自公平（无偏）测试的证据表明己烯雌酚可以实现这个目标，后来事实表明它导致了部分服药孕妇的女儿患上癌症。没有可靠证据表明其有效性，这种疗法就不应推广。

对今天和过去提供的治疗进行对比，只不过是给公平试验提供一个可靠的基础（[Behring等 1893](#)；[Roux等 1894](#)），因为治疗本身之外的其它相关因素会随时间而改变。例如，流产和死胎在第一次怀孕时比在后来怀孕时更常见。因此，如果把后来怀孕时服用DES后流产和死胎的频率和第一次怀孕时没用药的结果进行比较，



就可能构成疗效评估的严重误导。因此，如果可能的话，对照应涉及在大致相同的时间内提供不同的治疗方法。

### 通过患者个体交叉试验比较疗法

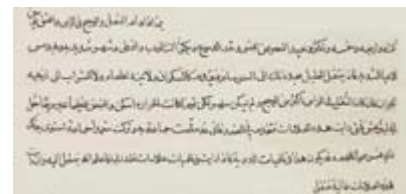
有时在大致相同的时间内提供不同的治疗可能包括为一名患者提供一个接一个不同的治疗——一种所谓的交叉试验（[Martini 1932](#)；[点击这里列出相关记录](#)）。交叉试验的一个早期例子是1786年英格兰巴思的卡莱布·帕里（Caleb Parry）医生报告的。他想找出是否有任何理由要购买昂贵的进口土耳其大黄作为泻药治疗他的病人，而不是用英格兰本地种植的大黄。所以他“交叉”了在不同时间为每个患者提供的大黄类型，然后比较了病人在服用每类大黄时发生的症状（[Parry 1786](#)）。（他没有发现昂贵大黄的任何优点！）



对于单个病人的治疗如果在停止治疗后情况发生逆转，对照就很有意义。在许多情形下，这并不适用。例如，通常不可能以这种方式比较不同的手术，或比较在渐进病况下提供的治疗。

### 比较同时接受不同治疗方法的患者组群

对治疗的测试通常是对接受不同治疗的患者组群进行比较。如果病情相对较好的人接受一种治疗，病情相对较严重的人接受另一种治疗，那么这两种治疗的比较是不公平的，所以必须比较在相同时期接受不同治疗的相似患者组群的经历。阿尔-拉兹（Al-Razi）在一千多年前就认识到了这一点，他希望找到结论，就如何治疗具有早期脑膜炎体征的患者提出看法，他对一组病人进行治疗，同时故意限制治疗对照组的病人（[al-Razi, 9世纪](#)）。



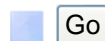
治疗措施的公平试验需要与自然疗法和其它疗法进行比较。要想让这些对照做到公平，必须消除真正的不确定性，避免偏倚和机遇因素，并谨慎地做出解释。

**Cite as:** Editorial commentary (2007). 为什么对照必不可少. The James Lind Library ([www.jameslindlibrary.org](http://www.jameslindlibrary.org)).

**Next essay:** [为什么对照必须解决真正的不确定性](#)

### Select other

**essay:** [为什么需要进行公平试验](#)



## 参考文献

al-Razi (10e siècle CE; 4e siècle AH). Kitab al-Hawi fi al-tibb [The comprehensive book of medicine].

Asher R (1972). Talking sense. Londres : Pitman Medical.

Banting FG, Best CH, Collip JB, Campbell WR, Fletcher AA (1922). Pancreatic extracts in the treatment of diabetes mellitus. Canadian Medical Association Journal 12: 141-146.

Behring, Boer, Kossel H (1893). Zur Behandlung diphtheriekranker Menschen mit Diphtherieheils Serum. Deutsche Medicinische Wochenschrift 17: 389-393.

Colebrook L, Purdie AW (1937). Treatment of 106 cases of puerperal fever by sulphanilamide. Lancet 2: 1237-1242 & 1291-1294.

Cullen W (1772). Clinical lectures. Edimbourg, février-avril, 218-9.

Cummngs R (1805). Medical and Physical Journal, page 6.

Druker BJ, Talpaz M, Resta DJ, Peng B, Buchdunger E, Ford JM, Lydon NB, Kantarjian H, Capdeville R, Ohno-Jones S, Sawyers CL (2001). Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. New England Journal of Medicine 344: 1031-1037.

Flint A (1863). A contribution toward the natural history of articular rheumatism; consisting of a report of thirteen cases treated solely with palliative measures. American Journal of the Medical Sciences 46: 17-36.

Forbes J (1846). Homeopathy, allopathy and 'young physici.' British and Foreign Medical Review 21: 225-265.

Holmes OW (1861). Currents and countercurrents in medical science. In: Works, 1861 Vol ix, p 185.

Martini P (1932). Methodenlehre der Therapeutischen Untersuchung. Berlin: Springer.

McLean-Tooke APC, Bethune CA, Fay AC, Spickett GP (2003). Adrenaline in the treatment of anaphylaxis: what is the evidence? *BMJ* 327:1332-1335.

Medical Research Council (1948). Streptomycin treatment of tuberculous meningitis. *Lancet* 1:582-596.

Ministère des Affaires intérieures (1823). [Conclusion du Conseil médical sur le traitement homéopathique]. *Zhurnal Ministerstva Vnutrennih del*, 3:49 - 63.

Minot GR, Murphy WP (1926). Treatment of pernicious anaemia by a special diet. *JAMA* 87:470-476.

Paré A (1575). *Les oeuvres de M. Ambroise Paré conseiller, et premier chirurgien du Roy avec les figures & portraits tant de l' Anatomie que des instruments de Chirurgie, & de plusieurs Monstres*. Paris: Gabriel Buon.

Parry CH (1786). Experiments relative to the medical effects of Turkey Rhubarb, and of the English Rhubarbs, No. I and No. II made on patients of the Pauper Charity. *Letters and Papers of the Bath Society III*: 431-453.

Rivers WHR (1908). *The influence of alcohol and other drugs on fatigue*. Londres: Edward Arnold.

Roux E, Martin L, Chaillou A (1894). Trois cent cas de diphtérie traité par le sérum antidiphtérique. *Annales de l' Institut Pasteur* 8:640-661.

Schleisner PA (1849). *Island fra et lægevidenskabeligt Synspunkt*. København: Boghandler Iversen.

Sutton HG (1865). Cases of rheumatic fever, treated for the most part by mint water. Compilé dans les livres cliniques du Dr Gull, avec des remarques sur l' histoire naturelle de cette maladie. *Rapport hospitalier de Guy* 11:392-428.

Tibi S (2005). *The medicinal use of opium in ninth-century Baghdad*. Leiden: Brill.

[Home](#)

[目录页](#)

[Comments](#)

[Home](#)[目录页](#)[jameslindlibrary.org](http://www.jameslindlibrary.org)

## 为什么对照必须解决真正的不确定性

大量研究甚至是在没有真正不确定性的情况下开展的。那些没有对以往治疗措施试验进行系统评价就开始进一步研究的研究人员，有时不承认（或选择忽视事实）已经令人信服地解决了关于疗效的不确定性。这意味着参加研究的人有时候无法获得有益治疗，或者接受了可能殃及自身健康的疗法。

本文所载图表和下段文字表明了为评估抗生素（相对于无活性安慰剂）是否可以降低直肠手术病人的术后死亡的风险所做的公平试验的累积证据（Lau 等 1995）。据报告，首次公平试验于1969年进行。这次小型研究的结果就抗生素是否具有效果留下了不确定性——表示结果的水平线跨过了划分抗生素有利于治疗和不利于治疗的竖线。这种不确定性适时地在20世纪70年代的进一步测试中得到了解决。

不过，随着证据的积累，到70年代中期人们已经很明确抗生素可降低术后死亡的风险（横线显然落在有利于治疗的竖线这一侧）。不过研究者们继续深入研究，直到20世纪80年代末。在这些后来的研究中得到安慰剂的一半患者因此也就无法获得一种表明可降低他们术后死亡风险的医疗方式。这怎么会发生呢？这可能是由于研究人员在继续研究的过程中没有系统评价现有的证据。这种行为在学术界仍然司空见惯，从一定程度上讲，这是由于商业和学术研究领域中的某些激励机制没有把病人的利益放在首位所致（Chalmers 2000）。

此外，还可能由于研究人员没有对动物研究的相关证据进行系统评价就开始在人体上试验疗法，而让参加研究的患者和参加者遭受痛苦。一个荷兰的研究小组审查了7000多名患者的经历，他们参加了向中风患者提供的某种新型钙离子阻断药物试验。他们找不到可以支持在实践中增加使用的证据（Horn和Limburg 2001）。这使他们怀疑哪些导致在患者身上做研究的动物研究的质量和结果。他们对动物研究的审查显示，这些研究从未表明该药物对于人体有用（Horn等 2001）。

研究者未解决真正的不确定性的最常见原因是没有受过足够的训练，使得在开始新研究之前能够系统评价所有相关的现有证据。不过，有时也有更险恶的原因。研究者可能知道现有证据，但是他们想设计研究活动，来确保他们自己的研究可以产生对某些特定疗法有利的结果。通常情况下，但并非总是如此，这是出于商业利益考虑（Djulgovic等 2000； Sackett和Oxman 2003）。这些研究都刻意设计成不公平的治疗措施试验。为了实现这一目的，采取的措施是压制一种明知对患者有效的对照疗法（如上文所述例子），或者以不适当的低剂量作为比较疗法（从而让它们不是那么有效），或者以不适当的高剂量（从而让它们产生有害的副作用）（参见Mann和Djulgovic的评论）。这还可能是由于追踪患者的时间过短（从而错过治疗的延缓效果），或者是采用与对患者重要的结果有很少相关性或根本没有相关性的结果衡量标准（“替代品”）。

可能让本文读者吃惊的是，近几十年来建立的旨在确保研究符合伦理道德的研究伦理委员会，在影响这种研究中的不端行为方面，几乎没有什么作为。大多数这类委员会都让本该受他们保护的民众失望，因为他们并没有要求那些寻求批准新试验的研究人员和赞助者事先系统评价现有证据（Savulescu等 1996； Chalmers 2002）。研究伦理委员会未能以这种方式有效保护患者和公众，这突显了提高公众认识的重要性，他们应当了解关于医学治疗的公平试验的特点。

**Cite as:** Editorial commentary (2007). 为什么对照必须解决真正的不确定性. The James Lind Library ([www.jameslindlibrary.org](http://www.jameslindlibrary.org)).

**Next essay:** [避免带有偏倚的对照](#)

**Select other**

**essay:** [为什么需要进行公平试验](#)

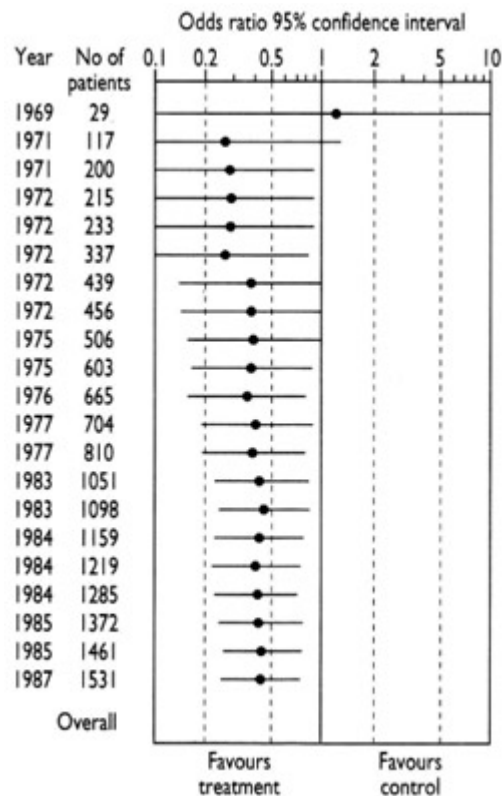


## 参考文献

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.

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

Chalmers I (2002). Lessons for research ethics committees. *Lancet* 359: 174.



Djulgovic B, Lacey M, Cantor A, Fields KK, Bennett CL, Adams JR, Kuderer NM, Lyman GH (2000). The uncertainty principle and industry-sponsored research. *Lancet* 356:635-638.

Horn J, Limburg M (2001). Calcium antagonists for acute ischemic stroke (Cochrane Review). In: *The Cochrane Library*, Issue 3, Oxford: Update Software.

Horn J, de Haan RJ, Vermeulen M, Luiten PGM, Limburg M (2001). Nimodipine in animal model experiments of focal cerebral ischaemia: a systematic review. *Stroke* 32:2433-38.

Lau J, Schmid CH, Chalmers TC (1995). Cumulative meta-analysis of clinical trials builds evidence for exemplary clinical practice. *Journal of Clinical Epidemiology* 48:45-57.

Mann H, Djulgovic B. Why comparisons must address genuine uncertainties. James Lind Library ([www.jameslindlibrary.org](http://www.jameslindlibrary.org)).

Sackett DL, Oxman AD (2003). HARLOT plc: an amalgamation of the world's two oldest professions. *BMJ* 2003;327:1442-1445.

Savulescu J, Chalmers I, Blunt J (1996). Are research ethics committees behaving unethically? Some suggestions for improving performance and accountability. *BMJ* 313:1390-1393.

[Home](#)

[目录页](#)

[Comments](#)

[Home](#)[目录页](#)[jameslindlibrary.org](http://www.jameslindlibrary.org)

## 避免带有偏倚的对照

有时候治疗具有显著的效果 ([点击这里列出相关记录](#))。这些可能是无意和特定的情况, 例如, 有人对某种抗生素药物会产生过敏反应。治疗也可能有令人惊异的良好效果, 例如用肾上腺素治疗危及生命的过敏反应 (McLean-Tooke et al. 2003)。不过这种惊人的疗效很少见。通常有些治疗的效果更为适度, 不过非常值得了解, 例如, 使用阿司匹林以减少心脏病发作的风险 ([Elwood 2004](#))。

例如, 阿司匹林无法避免所有心脏病发作后的过早死亡, 但是它的确能使死亡概率降低约百分之二十, 这对于这类常见病症具有重要意义。要想可靠检测到大多数疗法的这些适度但有很重要的效果, 必须小心谨慎, 以确保不让带有偏倚的对照使得我们对某些实际上无效或有害的疗法误信为有效, 或者实际上有效时误认为无效。

在治疗试验中造成的偏倚, 就是那些可能导致疗效的结论系统偏离真实情况所带来的影响作用和因素。尽管有许多种偏倚可能扭曲卫生研究的结果 (Sackett 1979), 但詹姆斯·林德图书馆侧重于研究在治疗措施的公平试验中必须尽可能减少的那些偏倚。它们是:

- [由于被比较对象的差异而产生的偏倚](#);
- [由于治疗结果评估方式的差异而产生的偏倚](#);
- [对现有证据带有偏倚的报告](#);
- [对现有证据带有偏倚的选择](#)。

无视这些偏倚 (或有时肆无忌惮地利用它们), 可能会导致人们相信某种新的疗法优于某种现有疗法, 但事实却并非如此。这可能是由于所得出的结论基于:

- 对病情较好但接受某种新疗法的人和病情较重但接受标准疗法的人的各自病程发展进行比较的研究 ([分配偏倚](#))。
- 关于治疗结果的评估可能偏向于某种新疗法的研究, 例如, 对知道自己采用了某种昂贵的新疗法的人和那些可能失望于将继续采用某种普通的标准疗法的人的观点进行比较 ([观察者或衡量偏倚](#))。
- 只从有利角度体现某种新疗法的研究, 而不是表明它可能有害的研究, 往往也不报告 ([报告偏倚](#))。
- 对现有证据的带有偏倚的选择和解释, 旨在支持某个特定的观点 ([审阅者偏倚](#))。

通常, 由于这些偏倚导致的关于疗法的不公平检验不会被如实辨认出来。但是, 既得利益者有时会利用这些偏倚, 使疗法显得似乎优于真实情形 (Sackett和Oxman 2003)。

无论偏倚是无心还是有意的, 其后果是相同的: 除非验证治疗措施的试验是公平的, 否则一些无用或有害的治疗方法会被认为有益, 而一些有效的治疗方法却会被认为无益或有害的。

**Cite as:** Editorial commentary (2007). 避免带有偏倚的对照. The James Lind Library ([www.jameslindlibrary.org](http://www.jameslindlibrary.org)).

**Next essay:** [被比较对象之间的差异](#)

**Select other essay:**



## 参考文献

Elwood P (2004). The first randomised trial of aspirin for heart attack and the advent of systematic overviews of trials. The James Lind Library ([www.jameslindlibrary.org](http://www.jameslindlibrary.org)).

McLean-Tooke APC, Bethune CA, Fay AC, Spickett GP (2003). Adrenaline in the treatment of anaphylaxis: what is the evidence? *BMJ* 327:1332-1335.

Sackett DL (1979). Bias in analytic research. *Journal of Chronic Diseases* 32:51-63.

Sackett DL, Oxman AD (2003). HARLOT plc: an amalgamation of the world's two oldest professions. *BMJ* 2003;327:1442-1445.

[Home](#)[目录页](#)

[Comments](#)

[Home](#)[目录页](#)[jameslindlibrary.org](#)

## 避免带有偏倚的对照

### 被比较对象之间的差异

#### 比较提供给不同患者组群的不同疗法

疗法比较通常需要对照接受了不同治疗方法的患者组群的经历。要使对照做到公平，不同组群的患者构成必须类似——因而，同类患者可以与同类患者进行比较。如果接受某种治疗的患者很可能无论如何都要比接受另一种治疗的患者做得好（或坏），这就带来了分配偏倚，使我们无法确信疾病的结局反映了治疗措施的不同疗效，而不是由于自然和时间流逝带来的效果。

18世纪的医生威廉·切塞尔登（William Cheselden）认识到了“不相似组群”问题，当时医生们正在分别比较摘除膀胱结石术后的患者死亡率。切塞尔登指出，考虑不同医生治疗的病人的年龄很重要。他注意到，事实上死亡率随着患者的年龄不同而发生变化（[Cheselden 1740](#)）——年老的患者比年轻的患者更可能死亡。这意味着，如果人们希望比较进行不同类手术的患者组群之间的死亡率，就必须考虑对照组中患者的年龄差异。



对过去无意中接受过不同治疗措施的患者的经历和结局进行比较，今天仍然被用作一种尝试评估疗效的方法。面临的挑战是要了解清楚用于对照的组群在接受治疗前是否十分相似。通过比较采用激素替代疗法（HRT）的妇女和未采用该疗法的妇女之间的患病经历，设法评估激素替代疗法产生的效果就是一个例证。正如随后对激素替代治疗措施的公平试验所作的分析所显示的那样，试图以这种方式回顾评价治疗措施的效果，有时可能带有危险的误导性（McPherson 2004）。

几乎不太可能完全相信，从过去接受过某种疗法的人群中选出的对照组与最近接受过另一种疗法的群组在所有重要方面都可以类比。即使我们能够获取接受过不同治疗方法的患者的某些信息（例如他们的年龄，或者他们以往的病史），结果也是如此。其它具有重要意义的信息（例如自愈可能性）可能根本无法获取。

较好的做法是在开始治疗之前就筹划疗法的对照。例如，詹姆斯·林德1747年在皇家海军军舰“索尔兹伯里”号开始比较治疗坏血病的六种方法之前，他注意选择的病人都处于这种往往致命的疾病的类似阶段。他还确保他们相同的基本饮食，并被安置在相似的居住环境中。这些都是治疗方法之外有可能影响他们康复可能性的因素（[Lind 1753](#)）。必须付出同等努力，设法确保治疗对照组由类似的人员构成。



#### 使用交替或随机的方法确定无偏倚治疗对比群组

尽管林德注意保证他的六个对照组的水手具有相似性，但是他们并没有叙述他如何决定哪些水手应接受六种疗法中的哪一种。只有一种方式能够确保治疗对照组的设置方式在已知和未知的各个重要方面都相似。这就是通过利用某种形式的随机程序来组织治疗对照组，从而避免了在开始治疗之前针对不同疗法的偏倚选择。

在林德时代过去一百年之后，一名军医格雷厄姆·鲍尔弗（Graham Balfour）在一次试验中示范了如何做到这一点，该试验的目的是观察颠茄制剂是否能预防儿童患猩红热。在他负责的部队福利院里，他用交替的方法——“以避免选择的问题”——来决定哪些男孩得到颠茄制剂，哪些不给（[Balfour 1854](#)）。交替是在提供被比较的治疗方法之前组织相似的治疗对照组所用的集中无偏方法之一。在20世纪上半叶，有许多利用交替或轮换方法（例如[Hamilton 1816](#)；[MRC 1944](#)）或抽签（[Colebrook 1929](#)）来组织治疗对照组的例子——例如，掷骰子（[Doull et al. 1931](#)）、彩色珠子（[Theobald 1937](#)）或随机抽取号码（[Bell 1941](#)；[MRC 1948](#)；[MRC 1950](#)；[MRC 1951](#)）。这种“随机分配”是这类被称为“随机化”的公平试验唯一的但极为重要的特征。一个随机（有别于随意）分配意味着，已知发生某些事件的可能性，但任何特定情况下的结果都无法预料。所以举例来说，如果用一枚硬币进行随机选择，头朝上的可能性是50%，但不可能知道某次抛硬币的结果如何。



正如点击[这里](#)可以阅读到的文章所述，掷签或抽签是做出公平决定的一个确立已久的办法。这些方法有助于确保各个对照组不会由不同类型的人组成。可以检查已知并已测量的重要因素，如年龄。然而，可以预计一些可能影响康复的未测量因素，例如饮食、职业和焦虑，平均来说可以相抵。如果你想观察随机分配如何产生相似的人员分组（[请点击此处查看示范](#)）。

随着使用交替和随机分配方法对患者群体进行无偏分组，以比较不同疗法的进一步普及，人们清楚地认识到，必须严格遵守分配计划，以避免设立有偏倚的治疗对照组（[MRC 1934](#)）。如果不让那些决定谁参加治疗对照的人看到治疗分配计划，就可以避免分配偏倚带来的风险——简言之，防止他们作弊从而使比较产生偏倚（[MRC 1944](#)；[MRC 1948](#)；[MRC 1950](#)；[MRC 1951](#)）。



#### 避免治疗对照组的偏倚损失

在想方设法使治疗对照组的设立方式实现了同类比较之后，很重要的一点是避免由于患者选择性退出对照组而导致的偏倚。应当尽可能保持分组的相似性，使分配到治疗对照组的所有人员都得到随访并纳入到试验结果的

主要分析之内，这就是所谓的“意向治疗”分析（[Bell 1941](#)）。

不这么做就有可能导致验证治疗措施的试验不公平。举例来说，对于脑部血管部分堵塞而导致晕眩的病人，具有两种截然不同的治疗方法。对这种病况的治疗非常重要，因为由于这种原因出现短暂晕眩的病人患中风的风险有所增大，中风有可能导致残疾，甚至死亡。针对晕眩的一种疗法，包括服用阿司匹林来阻止阻塞恶化；另一种疗法包括一种外科手术，以清除血管中的血栓。

对这两种治疗晕眩的方法进行公平比较，将包括利用某种无偏倚的分配方法（如随机选择）设立两个患者组群。那么比较可以从两组相似的患者开始，接着继续比较他们各自随后发生中风的发生率。但如果通过手术治疗组的中风发生率只记录术后存活的病人，那么就会忽视手术本身会导致中风和死亡这一重要事实。这将导致对两种治疗措施的不公平比较，由此导致对手术效果的一种带有偏倚的盲目乐观描述。这就说，没有进行相似者之间的比较。

随机试验的比较重点，必须尽可能建立在指派的所有人都接受了要比较的各自治疗措施，不能有例外，而且这些人要属于他们原先被指派的组。如果不遵循这一原则，人们就可能得到总体疗效的偏倚信息。

**Cite as:** Editorial commentary (2007). 被比较对象之间的差异. The James Lind Library ([www.jameslindlibrary.org](http://www.jameslindlibrary.org)).

**Show JLL records:** illustrating [control of allocation bias](#) and [crossover test](#)

**Next essay:** [治疗结局评价方式的差异](#)

**Select other essay:**

为什么需要进行公平试验



## 参考文献

Balfour TG (1854). Cité dans West C. Lectures on the Diseases of Infancy and Childhood. Londres, Longman, Brown, Green and Longmans, p 600.

Bell JA (1941). Pertussis prophylaxis with two doses of alum-precipitated vaccine. Public Health Reports 56: 1535-1546.

Cheselden W (1740). The anatomy of the human body. 5th edition. Londres: William Bowyer.

Colebrook D (1929). Irradiation and health. Medical Research Council Special Report Series No.131.

Doull JA, Hardy M, Clark JH, Herman NB (1931). The effect of irradiation with ultra-violet light on the frequency of attacks of upper respiratory disease (common colds). American Journal of Hygiene 13:460-77.

Hamilton AL (1816). Dissertatio Medica Inauguralis De Synocho Castrensi (Inaugural medical dissertation on camp fever). Edimbourg: J Ballantyne.

Lind J (1753). A treatise of the scurvy. In three parts. Containing an inquiry into the nature, causes and cure, of that disease. Together with a critical and chronological view of what has been published on the subject. Edimbourg: Imprimé par Sands, Murray et Cochran pour A Kincaid et A Donaldson.

McPherson K (2004). Where are we now with hormone replacement therapy? BMJ 328: 357-358.

Medical Research Council Therapeutic Trials Committee (1934). The serum treatment of lobar pneumonia. BMJ 1: 241-245.

Medical Research Council (1944). Clinical trial of patulin in the common cold. Lancet 2: 373-5.

Medical Research Council (1948). Streptomycin treatment of pulmonary tuberculosis: a Medical Research Council investigation. BMJ 2: 769-782.

Medical Research Council (1950). Clinical trials of antihistaminic drugs in the prevention and treatment of the common cold. BMJ 2: 425-431.

Medical Research Council (1951). The prevention of whooping-cough by vaccination. BMJ 1: 1463-1471

Parry CH (1786). Experiments relative to the medical effects of Turkey Rhubarb, and of the English Rhubarbs, No. I and No. II made on patients of the Pauper Charity. Letters and Papers of the Bath Society III:407-422.

Silverman WA, Chalmers I. Casting and drawing lots. The James Lind Library ([www.jameslindlibrary.org](http://www.jameslindlibrary.org)).

Theobald GW (1937). Effect of calcium and vitamin A and D on incidence of pregnancy toxæmia. Lancet 2:1397-1399.

[Home](#)

[目录页](#)

---

[Comments](#)

[Home](#)[目录页](#)[jameslindlibrary.org](#)

## 避免带有偏倚的对照

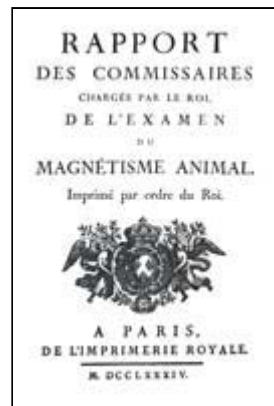
### 治疗结局评价方式的差异

#### 利用盲法来减少治疗效果评估的偏倚

对于用来评估治疗的某些结局——例如存活——不太可能使用带偏倚的评价，因为几乎没有提出意见的空间。18世纪一些外科手术方法的试验就是如此，在这种情况下存活是手术成功与否的主要衡量方式（[Faure 1759](#)）。不过，其它大多数结局的评估要么始终涉及主观性（如病人的症状），要么可能涉及主观性。导致这些错误知觉的偏倚被称为观察者偏倚。当人们相信他们已经“知道”某种治疗的效果时，或者当他们由于特殊原因偏好被比较的治疗措施中的某一种时，就会带来某种问题。如果没有采取措施来减少治疗措施的比较中带有偏倚的结局评估，治疗效果往往会被高估（[Schulz等 1995](#)）。在评估结局时，主观性的成分占得越大，就越需要减少相应的观察者偏倚，以确保治疗措施试验的公平性。

在这些常见的情形中，针对患者和医生的“盲法”是进行公平试验的理想方法。对于某种疗法的最早的盲法（蒙面）评估是由路易十六所委任的一个调查委员会来实施的，目的是调查安东·梅斯梅尔

（Anton Mesmer）声称的“动物磁力说”（[皇家委员会 1784](#)）。该委员会负责评估这种新的治疗方法传说中的疗效，到底是由于任何“真正”的力量，还是由于“精神幻觉”。被蒙住眼睛的人会被告知他们正在接受或没有接受磁力，但有些时候事实上发生的情况却恰恰相反。受试者只有当得知他们正在接受治疗时才会感觉到“动物磁力”，而不是其他情形（[Kaptchuk 1998](#)；[Schulz等 2002](#)）。



#### 使用安慰剂实现盲法

在进行动物磁力效果试验后数年，约翰·海加斯（John Haygarth）利用一种假冒装置（一种安慰剂）进行了一次实现盲法的实验（[Haygarth 1800](#)）。本段所配的漫画显示了一名医生用一种由艾莉莎·珀金斯（Elisha Perkins）申请专利并销售的装置来治疗一名有钱的患者。珀金斯声称他的“牵引器”——小金属杆——通过“电物理作用力”治愈了各种疾病。在一本题为《导致身体不适和作为治疗方法的想象力：以假牵引器为例》的小册子中，约翰·海加斯报告了他如何对珀金斯的一些说法进行公平试验。在不知道其评估细节的众多患者身上，他采用一种交叉研究方法，比较了已获得专利的金属牵引器（这意味着通过“电物理作用力”起作用）和看起来完全相同的木制“牵引器”（“安慰物牵引器”）。他没有发现金属牵引器有任何益处（[Haygarth 1800](#)）。



约翰·海加斯关于珀金斯牵引器的公平试验是使用安慰剂实现盲法，以减少治疗结局评估中发生偏倚的一个早期实例。在关于顺势疗法的辩论中，安慰剂成为一种研究工具，该疗法是19世纪非传统疗法的另一个主要形式。顺势疗法常常用盲法评估和安慰剂对照来进行“检验”，在健康的志愿者中间试验其疗法的效果（[Löhner 1835](#)；[Kaptchuk 1998](#)）。一个最复杂的安慰剂对照试验，于1879-1880年在密尔沃基医学院（Milwaukee Academy of Medicine）进行。该试验采用了“双盲”法：患者和实验人员都不知道所做的治疗是真正的顺势疗法，还是服用了一种糖丸（[Storke等 1880](#)）。

直到此后很久，主流医学界才持有更加怀疑的态度，使得人们承认有必要采用盲法评估和安慰剂来评估自我说法的真实性。德国的研究人员主要受到药理学家们的启发，逐渐采用了盲法评估。例如，1918年，阿道夫·宾格（Adolf Bingel）报告说，他在比较白喉的两种不同疗法时努力做到“尽可能客观”（[Bingel 1918](#)）。他对自己或同事是否能够猜到哪些病人接受了哪种治疗进行了评估：“我没有仅仅依靠自己的判断，而是在没有告知试验血清本质的情况下，征求了白喉病室助理医师的意见。因此他们的判断完全不带任何偏见。我渴望看到我的观察结论能够得到独立审查，因此最热烈地推荐使用这种‘盲’法。”（[Bingel 1918](#)）。事实上，在两种疗法之间没有检测到任何差异。德国确立了强有力的盲法评估传统，并由临床药理学家保罗·马提尼（Paul Martini）进行了编纂（[Martini 1932](#)）。



在现代英语国家，药理学家在德国传统以及本地采用盲法评估的“惩治庸医（quackbuster）”运动影响下，首次开始采用盲法评估（[Kaptchuk 1998](#)）。到20世纪30年代，他们在临床试验中使用安慰剂对照方面已经处于领先地位。例如，英国医学研究理事会两例最早的公平试验是针对普通感冒的治疗方法。要是没有用“双盲”法来避免患者和医生知道哪些患者得到的是新药以及哪些患者得到的是安慰剂，那么就很难解释他们的结果（[MRC 1944](#)；[MRC 1950](#)）。哈里·古尔德（Harry Gold）对于盲法评估的奋力宣传似乎对美国带来了特别重要的影响（[疗法会议 1954](#)）。

#### 如果不能对患者和临床医师设盲，可对观察者设盲

在比较治疗措施时，就疗法的本体对患者和医生进行盲法处理是不可能的，例如，当对手术治疗与药物治疗或不治疗作比较时就是如此。不过，即使在这种情况下，也能采取步骤来减少治疗效果的偏倚评价。可不使独立的观察员了解患者采用了哪种疗法。例如，在20世纪40年代初，曾有一项试验，对接受了当时的标准疗法（卧床休息）的肺结核患者和另外还接受了链霉素药物注射的患者进行比较。研究者认为，向仅仅卧床休息的分组患者注射无活性安慰剂，达到对患者和诊疗医生设盲的目的是不道德的

([MRC 1948](#))，但是他们采取了替代预防措施，减少对结果带来的偏倚评价。尽管主要结果（存活）几乎没有发生偏倚评价的危险，但主观性可能导致对胸部X光检查的评价产生偏倚。因此，评阅X光片的医生一直不知道他们评估的结果归属于使用链霉素治疗的患者，还是归属于仅采取卧床休息治疗的患者。

与随机化做法结合在一起，在可能情况下使用安慰剂的盲法评估，如今已成为于验证治疗措施公平试验的关键方法之一。

**Cite as:** Editorial commentary (2007). 治疗结局评价方式的差异. The James Lind Library ([www.jameslindlibrary.org](http://www.jameslindlibrary.org)).

**Show JLL records:** illustrating [control of observer bias](#)

**Next essay:** [解释无偏倚对照](#)

**Select other essay:**

[为什么需要进行公平试验](#)



## 参考文献

Bingel A (1918). Über Behandlung der Diphtherie mit gewöhnlichem Pferdeserum. Deutsches Archiv für Klinische Medizin 125:284-332.

Commission Royale (1784). Rapport des commissaires chargés par le roi du magnétisme animal. Paris : Imprimerie royale.

Conference on Therapy (1954). How to evaluate a new drug. American Journal of Medicine 17:722-727.

Faure (1759). Recueil des pièces qui ont concouru pour le prix de L' Académie Royale de Chirurgie. Vol 8. Paris, P.AI Le Prieur.

Haygarth J (1800). Of the imagination, as a cause and as a cure of disorders of the body: exemplified by fictitious tractors, and epidemical convulsions. Bath : R. Crutwell.

Kaptchuk TJ (1998). Intentional ignorance: a history of blind assessment and placebo controls in medicine. Bulletin of the History of Medicine 72:389-433.

Löhner G (1835), on behalf of a Society of truth-loving men. Die Homoöopathischen Kochsalzversuche zu Nürnberg [The homeopathic salt trials in Nuremberg].

Martini P (1932). Methodenlehre der Therapeutischen Untersuchung. Berlin :Springer.

Medical Research Council (1944). Clinical trial of patulin in the common cold. Lancet 2:373-375.

Medical Research Council (1948). Streptomycin treatment of pulmonary tuberculosis: a Medical Research Council investigation. BMJ 2:769-782.

Medical Research Council (1950). Clinical trials of antihistaminic drugs in the prevention and treatment of the common cold. BMJ 2:425-431.

Schulz KF, Chalmers I, Hayes RJ, Altman DG (1995). Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 273:408-412.

Schulz KF, Chalmers I, Altman D (2002). The landscape and lexicon of blinding. Annals of Internal Medicine 136:254-259.

Storke EF, Martin R, Rosenkrans EM, Ford J, Schloemilch A, McDermott GC, Carlson OW (1880). Final report of the Milwaukee test of the thirtieth dilution. Homeopathic Times: A Monthly Journal of Medicine, Surgery and the Collateral Sciences 7:280-281.

[Comments](#)

[Home](#)[目录页](#)[jameslindlibrary.org](#)

## 解释无偏倚对照

一个公平的治疗对照应避免带有偏倚的比较。这要求采取措施，尽量减少由于被比较患者之间差异所造成的偏倚，以及由于治疗结果评估方法上的差异而造成的偏倚。

不过，即便已经避免了这些偏倚，但解释无偏倚对照并不总是直截了当的。例如，是否已考虑了预期接受的治疗和实际接受的治疗之间存在的差异，是否已考虑机遇因素效应？

有时，一项新的研究会提供关于某项治疗效果的非常强有力的证据。例如，数以万计的人参加了一项大型研究，该研究表明阿司匹林药片能够大大减少心脏病发作病人死亡的风险（ISIS-2 1988）。不过，单项研究提供这种强有力证据的情形极为罕见。因此，在阅读大多数研究报告时，很重要的一点是要询问，是否已将新证据纳入到所有其它相关证据的系统评价之内。如果答案是肯定的，在汇总过程中是否已采取措施，尽量减少对现有证据带有偏倚的报告及现有证据带有偏倚的选择带来的影响？是否已考虑利用Meta分析方法减少机遇因素影响的可能性？

**Cite as:** Editorial commentary (2007). 解释无偏倚对照. The James Lind Library (www.jameslindlibrary.org).

**Next essay:** [预期接受的治疗和实际接受的治疗之间存在的差异](#)

**Select other essay:**

[为什么需要进行公平试验](#)



## 参考文献

ISIS-2 Second International Study of Infarct Survival Collaborative Group (1988). Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17187 cases of suspected acute myocardial infarction: ISIS-2. Lancet 2: 349 60.

[Home](#)[目录页](#)[Comments](#)

[Home](#)[目录页](#)[jameslindlibrary.org](http://www.jameslindlibrary.org)

## 解释无偏倚对照

### 预期接受的治疗和实际接受的治疗之间存在的差异

医学治疗措施的公平试验必须仔细规划。制定这些规划的文件称为试验方案。除其他外，该方案详细说明了所要对照的治疗措施的细节。然而，计划制定得再好，最后并非总能如愿以偿。患者在试验中实际接受的治疗有时会与他们应接受的治疗不同。在解释治疗对照的结果时，必须考虑这些偏离意图的现象。

在医学治疗措施的公平试验的发展过程中引入安慰剂的原因之一，就是为了减少对预期治疗的偏离（Kaptchuk 1998）。但是，即便是在以安慰剂作为对照的试验中，也可能步入歧途。在第二次世界大战期间，为感冒患者使用了一种称为棒曲霉素的药液，这与只使用了溶有药物的液体的其他患者进行了对比（MRC 1944）。结果分析显示，该药物不具有任何有益效果。这就出现了一个问题，即可能是用来溶解药物的液体使得药物失去了效力。换句话说，1000多名患者可能参加了两种无效疗法的对照！庆幸的是，通过试验证实，尽管没有发现棒曲霉素对感冒具有疗效，在试验中使用的这种药物确实具有活性（Chalmers and Clarke 2004）！

由于种种原因，实际接受的治疗可能与预期的治疗有所不同。例如，医生可能判定原本分配到某个正式治疗对照中的患者不应接受该项治疗；患者可能拒绝接受分配来的治疗措施，或者认为治疗措施带有蓄意性；使用的治疗剂量可能与治疗意图有所不同；或者某种治疗措施的供给可能会耗尽。

例如，当显然相同的白血病治疗措施在英国和美国的儿童身上得出的结果出现差异时，通过调查显示，英国得出的结果之所以较差是由于化学疗法渐渐显示可怕的毒性效应之后，英国医生不愿坚持进行化疗（医学研究理事会儿童白血病工作组，1986）。

由于这些原因，对公平试验的解释必须考虑实际接受的治疗与预期治疗出现差异的可能性。如果预期和实际之间出现了差别，重要的是考虑对解释证据带来的影响。

**Cite as:** Editorial commentary (2007). 预期接受的治疗和实际接受的治疗之间存在的差异. The James Lind Library ([www.jameslindlibrary.org](http://www.jameslindlibrary.org)).

**Next essay:** [考虑机遇因素](#)

**Select other essay:**

[为什么需要进行公平试验](#)



## 参考文献

Chalmers I, Clarke M (2004). The 1944 Patulin Trial: the first properly controlled multicentre trial conducted under the aegis of the British Medical Research Council. *International Journal of Epidemiology* 32:253-260.

Kaptchuk TJ (1998). Intentional ignorance: a history of blind assessment and placebo controls in medicine. *Bulletin of the History of Medicine* 72:389-433.

Medical Research Council (1944). Clinical trial of patulin in the common cold. *Lancet* 2:373-375.

Medical Research Council Working Party on Leukaemia in Children (1986). Improvement in treatment for children with acute lymphoblastic leukaemia. *Lancet* 1:408-11.

[Home](#)[目录页](#)[Comments](#)

[Home](#)[目录页](#)[jameslindlibrary.org](http://www.jameslindlibrary.org)

## 解释无偏倚对照

## 考虑机遇因素

当比较两种疗法时，结局的任何差异都可能由机遇因素造成的。例如，对一种新疗法和一种标准疗法进行比较，其中前者有4人病情出现好转，后者有6人好转。如果颇有信心得出结论，新疗法比标准疗法差，显然就会犯错：这种结果可能只是机遇因素造成的。如果重复进行比较，好转的患者人数可能出现相反情况（6对4），或相同（5对5），或出现其它比率。

不过，如果40名患者在新疗法的治疗下出现好转，而60名患者在标准疗法的治疗下出现好转，那么就不大可能利用机遇因素解释这种差异了。而如果400名接受新疗法治疗的患者出现好转，而600名接受标准疗法治疗的患者出现好转，那么显然新疗法的确实比不上标准疗法。因此，在疗法比较中减少机遇因素带来误导可能性的办法，就是确保公平试验要有足够多的人员参与进来，体验人们希望影响到的结果，例如出现好转或恶化。

在某些情况下，要在公平试验中获得关于疗效的可靠估计，需要有大量人员（数以千计，有时数以万计）参与。例如，当我们感兴趣的治疗结果十分罕见时（例如，在健康的中年妇女中采用激素替代疗法（HRT）预防心脏病发作和中风），就有必要涉及大量参与者。如果要可靠找到治疗措施具有的适度但又重要的疗效（例如，将心脏病发作患者早期死亡的风险降低20%），同样需要大量人员。

为了评估机遇因素在公平试验的结果中可能带来的影响作用，研究人员采用了“统计学意义检验”。当统计学家和其他人谈到疗法之间的“显著性差异”时，他们通常是指统计学意义。治疗方法之间在统计学意义上的显著性差异不一定具有任何实际意义。但是统计学意义检验是很重要的，因为这能帮助我们避免得出错误的结论，即当疗法之间不存在真正差异时却认为存在差异（有时称为第一类错误）。

另外同样重要的是考虑足够大数量的治疗结果，以避免更为常见的危险——当疗法之间实际上有差异时却得出没有差异的结论。这类错误有时称为第二类错误。当托马斯·格雷厄姆·鲍尔弗（Thomas Graham Balfour）对颠茄制剂能够使他所诊治的孤儿免患猩红热这种说法的测试结果进行解释时，他认识到了后一种危险（[Balfour 1854](#)）。被分配获得颠茄制剂的76名男孩中有2人患了猩红热，而没有获得该药物的75名男孩中则有2人患猩红热。鲍尔弗指出“人数太少，无法证实患病率降低是由于颠茄制剂的预防力量造成的”。如果更多的男孩患有猩红热，鲍尔弗或许能够得出关于颠茄疗效的一个更有信心的结论。相反，他只是指出151名男孩中仅有4例猩红热病人，人数太少无法得出一个令人信服的结论。

一种降低由于机遇因素带来误导的可能性的办法，涉及到对治疗差异的范围进行估计，真实差异很可能就存在其中（[Gavarret 1840](#)；[Huth 2006](#)）。这些估计范围被称为可信区间。正如本文首段所述，重复一项治疗对照，有可能在结果方面出现治疗措施具有差异效果的不同估计情况，特别是当估计数字只是以少数结果为依据得来时更是这样。可信区间顾及到了这种变化。可信区间能够比单一的统计学意义检验含有更多的信息，因此更有助于降低被机遇因素误导的可能性。

统计学检验和可信区间（无论是为了单项研究的分析，还是在关于若干独立但相似的研究的 [Meta分析](#) 中）帮助我们考虑机遇因素，并当疗效和差异实际上不存在时，避免得出存在疗效和差异的结论，或者当它们实际上存在时得出不存在的结论。

**Cite as:** Editorial commentary (2007). 考虑机遇因素. The James Lind Library ([www.jameslindlibrary.org](http://www.jameslindlibrary.org)).

**Show JLL records:** illustrating [taking account of the play of chance](#)

**Next essay:** [识别治疗措施的非预期效应](#)

**Select other essay:**

[为什么需要进行公平试验](#)

Go

## 参考文献

Balfour TG (1854). Quoted in West C. Lectures on the Diseases of Infancy and Childhood. Londres, Longman, Brown, Green et Longmans, p 600.

Gavarret LDJ (1840). Principes généraux de statistique médicale: ou développement des règles qui doivent présider à son emploi. Paris: Bechet jeune & Labé.

Huth EJ (2006). Jules Gavarret' s Principes Généraux de Statistique Médicale: a pioneering text on the

statistical analysis of the results of treatments.

[Home](#)

[目录页](#)

---

[Comments](#)

[Home](#)[目录页](#)[jameslindlibrary.org](#)

## 识别治疗措施的非预期效应

可以预计，在更广泛使用新疗法时，可能出现始料未及的治疗效果。例如，注册新药所需的最初测试，在几个月时间里至多覆盖几百或几千人。在这一阶段，只能在短期内碰到相对经常出现的非预期效应。

一些罕见的治疗效果，或者那些需要经过一段时间才能逐渐显现出来的效果，只有当治疗试验已持续足够长时间或者该疗法已广为应用之后才会被发现。此外，应用新疗法的患者往往在许多重要方面不同于那些参与最初测试的人。他们可能更加年长或更年轻，所属性别不同，病得更严重或更轻，生活在不同的环境之中，或者除了治疗相应的疾病之外还同时面临其它健康问题。这些差异可能影响疗效，并可能出现新的始料未及的效果（参见 [《英国医学杂志》\(BMJ\)2004年7月3日](#) 特刊）。

非预期效应（无论是**不利**还是**有利**的）的检测和确认，通常与用于评估新疗法预期效果的方法截然不同。有时卫生专业人员或患者最早对非预期效应产生怀疑。要确定这些初始直觉中有哪些属于治疗措施效果的真实反映，就构成了挑战。对此，读者们在看到前面的本系列文章时已经很熟悉，即避免由于**偏倚**和**机遇因素**带来的误导。

如果出现了非常惊人的非预期效应，并且在实施治疗之后十分频繁地出现，那么很可能会被卫生专业人员或患者自然而然地注意到。举例来说，出生的婴儿缺少四肢几乎是闻所未闻，所以当20世纪60年代发生这种病症的次数突然增加，就会自然引起人们的关注。这类婴儿的母亲都曾在妊娠早期服用过一种新上市的抗恶心药物——**酞胺哌啶酮** (thalidomide)，所以这很可能就是其中的原因，几乎没有必要进一步作出评估。意料之外的药物有利效果往往也是通过类似方式发现的。例如，人们发现某种治疗精神错乱的药物还能降低胆固醇 (Goodwin 1991)。

当人们注意到这种显著关联时，往往会将其确定为真正的非预期效应 (Venning 1982)。不过，关于非预期治疗效应的许多直觉都是基于远远无法令人信服的证据得来的。因此，就像设计试验用来检测预期疗效一样，对试验进行筹划，以确认或排除不太显著的疑似非预期效应，就涉及到**避免偏倚对照**。

验证疑似的非预期效应是否真实的研究，必须遵循“同类”比较的原则。治疗措施的随机分配是满足这一要求的理想方式。不过，只有在极少情况下，疑似疗效会通过进一步的分析或人员随访做出调查，在接受治疗之前这些人的治疗方式就随机进行了就被随机分配 (Hemminki和McPherson 1997)。因此面临的挑战是以其它方式设立无偏倚对照组，这往往要利用治疗期间例行收集的信息。

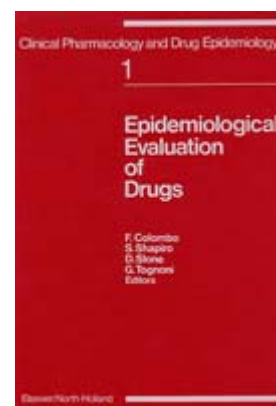
在这些研究中，做出治疗决定时没有预料到疑似效果实际上会带来帮助。因为这意味着在选择患者进行治疗时，尚未考虑到疑似病况带来的风险：非预期效应通常是指与该疗法治疗的病况或疾病有所不同的某种病况或疾病 (Vandenbroucke 2004a)。

例如，当推广激素替代疗法 (HRT) 治疗更年期症状时，不大可能考虑到一名女性形成静脉血栓的风险，因为大多数医生和女性认为两者是不相关的。因此没有理由预料采用激素替代疗法的女性与未服用该药物的人相比形成静脉血栓的风险会有所不同。由此确立了公平试验的基础，这些测试表明激素替代疗法会增加静脉血栓的风险。

当某种疑似非预期效应与针对常见健康问题（例如心脏病发作）的一项疗法有关，但并不经常伴随新疗法出现（或无法完全由其缓解）时，需要对接受该疗法的患者进行大规模监测，以检测这种非预期效应。例如，尽管有些人认为**阿司匹林**可能降低心脏病发作的风险，并于20世纪60年代末开始在患者身上对该理论进行公平试验 (Elwood et al. 1974)，但是当时大部分人都认为该理论非常难以置信。当实施一项大型研究，探测药物始料未及的负面作用时，就带来了突破：研究人员发现，由于心脏病发作入院接受治疗的患者中，近期服用过阿司匹林的病人患病的可能性低于明显相似的患者 ([波士顿药物监测合作组, 1974](#))。这些调查结果与一次公平试验的结果相符合。在该试验中，患者在心脏病发作后，以随机分配的方法，决定服用阿司匹林与否。这两份报告背靠背发表在同一期的《英国医学杂志》 (British Medical Journal, BMJ) 上 ([BMJ 1974](#))。

探测并调查非预期效应的基本准则，是20世纪70年代末首次明确提出的 ([Jick 1977](#); [Colombo等 1977](#))。在**酞胺哌啶酮**灾难之后，他们吸取了由此积累的调查非预期效应的共同经验。基于波士顿和牛津研究人员的经验 ([Jick and Vessey 1978](#))，对于一种重要类型的研究（即关于治疗可能带来的负面作用的病例对照研究），各种要求应写在论文之中。从此之后，虽然推出了许多强有力的治疗方案，但是治疗措施公平试验这一环节，今天仍然像当时一样具有挑战性和重要意义 (Vandenbroucke 2004b; Vandenbroucke 2006; Papanikolaou 等 2006)。

正如再前面的系列文章中所强调的那样，重要的是要认识到，单个报告中或是暗示或排除对于非预期效应的怀疑，都有可能起到误导作用。像所有其它的治疗措施公平试验一样，必须利用所有相关证据的**系统评价方法**，调查可能的非预期效应。例如，使用那些证实激素替代疗法与心脏病、中风和乳腺癌之间存在关系的证据 (Hemminki and McPherson 1997; 乳腺癌激素因素协



作小组, 1997)。

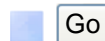
**Cite as:** Editorial commentary (2007). 识别治疗措施的非预期效应. The James Lind Library (www.jameslindlibrary.org).

**Show JLL records:** illustrating [identification of unanticipated effects](#)

**Next essay:** [对所有相关证据的系统评价](#)

**Select other essay:**

[为什么需要进行公平试验](#)



## 参考文献

Boston Collaborative Drug Surveillance Group (1974). Regular aspirin intake and acute myocardial infarction. *BMJ* 1:440-443.

Collaborative Group on Hormonal Factors in Breast Cancer (1997). Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. *Lancet* 350:1047-1059 .

Colombo F, Shapiro S, Slone D, Tognoni G, eds (1977). *Epidemiological Evaluation of Drugs*. Amsterdam: Elsevier/North Holland Biomedical Press, 1977.

Elwood PC, Cochrane AL, Burr ML, Sweetnam PM, Williams G, Welsby E, Hughes SJ, Renton R (1974). A randomised controlled trial of acetyl salicylic acid in the secondary prevention of mortality from myocardial infarction. *BMJ* 1:436-440.

Goodwin JS (1991). The empirical basis for the discovery of new therapies. *Perspectives in Biology and Medicine* 35:20-36.

Hemminki E, McPherson K (1997). Impact of postmenopausal hormone therapy on cardiovascular events and cancer: pooled data from clinical trials. *BMJ*;315:149-153.

Jick H (1977). The discovery of drug-induced illness. *New England Journal of Medicine* 296:481-485.

Jick H, Vessey M (1978). Case-control studies in the evaluation of drug-induced illness. *American Journal of Epidemiology* 107:1-7.

Papanikolaou PN, Christidi GD, Ioannidis JPA (2006). Comparison of evidence on harms of medical interventions in randomized and nonrandomized studies. *CMAJ* 174:635-641.

Vandenbroucke JP (2004a). When are observational studies as credible as randomised trials? *Lancet* 363:1728-1731.

Vandenbroucke JP (2004b). Benefits and harms of drug treatments. *BMJ* 329:2-3.

Vandenbroucke JP (2006). What is the best evidence for determining harms of medical treatment? *CMAJ* 174:645-646.

Venning GR (1982). Validity of anecdotal reports of suspected adverse drug reactions: the problem of false alarms. *BMJ* 284:249-254.

[Home](#)

[目录页](#)

[Comments](#)

[Home](#)[目录页](#)[jameslindlibrary.org](http://jameslindlibrary.org)

## 对所有相关证据的系统评价

二十世纪治疗措施公平试验的先驱者之一奥斯汀·布拉德福德·希尔 ([Austin Bradford Hill](#)) 指出, 研究报告的读者需要得到四个问题的答案: “你为什么开始研究呢?”, “你研究了什么?”, “你有什么发现?” 和 “它究竟意味着什么?” (Hill 1965)。回答希尔最后一个问题的质量尤其重要, 因为一份研究报告的这部分内容最可能影响治疗措施的实际选择和做出的决定。

单次治疗措施公平试验很难产生足够有力的证据, 为 “它意味着什么?” 这一问题提供一个令人信服的答案。某项治疗措施的一次公平试验, 通常是解决相同问题的一系列测试其中的一个。因此, 要想对 “它意味着什么?” 这一问题给出一个可靠的答案, 必须在对解决相关问题的公平试验所得出的全部证据进行审慎评估的背景下, 解释某次特定公平试验获得的证据。

早在一个多世纪前, 英国科学促进协会主席就表示, 有必要遵守这个原则:

“如果像有时所假设的那样, 科学包含的只是辛苦积累的事实, 那么它就像以往一样, 很快就会停滞不前, 被其自身重量压得粉碎……因此, 有两个过程同时起作用, 新材料的接纳和旧材料的消化和吸收……值得最大赞誉, 但是恐怕并没有总是得到最大赞誉的工作, 就是发现和解释工作的齐头并进, 不仅要陈述新的事实, 而且要指出与旧事实之间的关系。” ([Rayleigh 1885](#))

或许这是因为, 该原则在实践中的应用几乎仍然吸引不到学术界的赞誉, 治疗措施公平试验报告很少在针对所有其它相关证据做系统评估的背景下讨论得出的结果 (Clarke等 2002)。因此, 读者常常难以从新研究的报告中获得 “它意味着什么?” 这一问题的可靠答案。

正如在先前的解释性文章中所述, 没有先系统评价可以从现有研究中学到什么, 就开始实施医学治疗措施的新试验, 这带有危险性, 浪费了资源并且是不道德的 (参见 “[为什么比较的过程必须解决真正的不确定性](#)”)。同样具有危险的是, 报告了新测试的结果, 而不根据对其它相关证据的系统评估来解释新的证据, 因为它会延误对有益和有害疗法的鉴定 (Antman等 1992)。例如, 在20世纪60年代至90年代初之间, 对于旨在减少心脏病患者心律异常的药物进行了50多次公平试验之后, 人们才意识到这些药品简直是在杀人。如果每份报告在所有相关证据的背景下评估了新测试的结果, 那么药物的致命作用很可能在十年前就已经鉴别出来了, 许多不必要的过早死亡就可以得到避免。

在电子出版时代, 应该可能处理大多数新研究报告中带来的局限性 (Chalmers and Altman 1999; Smith and Chalmers 2001)。不过, 使用研究证据的用户已不再根据单个或几个单独研究的治疗措施作出结论, 而是日益依靠对所有相关可靠证据的[最新系统评价](#)来获取可靠信息, 这是因为人们日益认识到, 系统评价为医学治疗效果方面的结论提供了最好的依据。

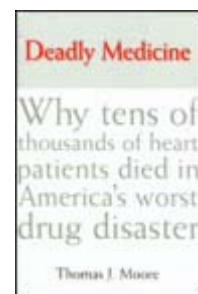
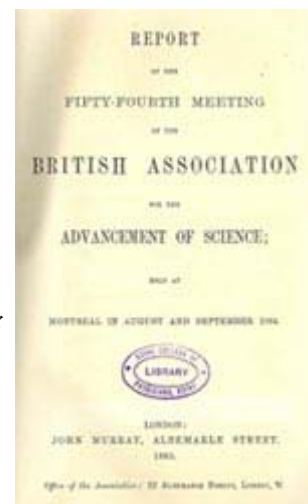
就像在规划、实施、分析和解释关于治疗措施的单个公平试验时, 必须采取措施避免[偏倚](#)和[机遇因素](#)的误导一样, 在规划、实施、分析和解释系统评价时, 也必须采取类似的步骤。这就要求:

- 明确说明由系统评价处理的问题
- 确定应包含研究的合格标准
- 找出 (全部) 可能符合条件的研究内容
- 以限制偏倚的方式应用合格标准
- 尽可能多地搜集各项研究的相关信息
- 在适当并且可能的情况下, 利用Meta分析和各种分析手段来分析这类信息
- 编写一份结构化报告

人们日益认识到, 通过系统评价对疗效进行评估至关重要, 其表现之一是用来提高回顾本身可靠性的方法得到快速发展。一本名为《系统评价》(Systematic Reviews)的书的第一版不到100页 ([Chalmers and Altman 1995](#)): 仅6年后, 第二版就厚达500页, 并包括了快速演化的加强研究信息获取的策略 (Egger等 2001)。

目前用来准备系统评价的方法有了重大进展, 包括确定非预期效应所需的那些方法 (Glasziou等 2004), 以及整合研究结果的那些方法, 这些研究描述并分析了提供和接受治疗者的经验 (Thomas 2004)。一旦有相关资料, 将会补充到詹姆斯·林德图书馆。

**Cite as:** Editorial commentary (2007). 对所有相关证据的系统评价. The James Lind Library ([www.jameslindlibrary.org](http://www.jameslindlibrary.org)).



**Show JLL records:** illustrating [systematic review](#)

**Next essay:** [处理对现有证据带有偏倚的报告](#)

**Select other essay:**

为什么需要进行公平试验



## 参考文献

Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC (1992). A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. *JAMA* 268:240-48.

Chalmers I, Altman DG (1995). *Systematic Reviews*. Londres: BMJ Publications.

Chalmers I, Altman DG (1999). How can medical journals help prevent poor medical research? Some opportunities presented by electronic publishing. *Lancet* 353:490-493.

Egger M, Davey Smith G, Altman D (2001). *Systematic Reviews in Health Care: meta-analysis in context*. 2nd Edition of *Systematic Reviews*. Londres: BMJ Books.

Glasziou P, Vandenbroucke J, Chalmers I (2004). Assessing the quality of research *BMJ* 328:39-41.

Hill AB (1965). Cited in 'The reasons for writing'. *BMJ* 4:870.

Rayleigh (1885). *Discours de son Excellence Lord Rayleigh*. Dans : Report of the fifty-fourth meeting of the British Association for the Advancement of Science ; à Montréal en août et septembre 1884, Londres: John Murray.

Smith R, Chalmers I (2001). Britain's gift: a 'Medline' of synthesized evidence. *BMJ* 323:1437-1438.

Thomas J, Harden A, Oakley A, Oliver S, Sutcliffe K, Rees R, Brunton G, Kavanagh J (2004). Integrating qualitative research with trials in systematic reviews *BMJ* 328:1010-1012.

[Home](#)

[目录页](#)

[Comments](#)

[Home](#)[目录页](#)[jameslindlibrary.org](http://www.jameslindlibrary.org)

## 对所有相关证据的系统评价

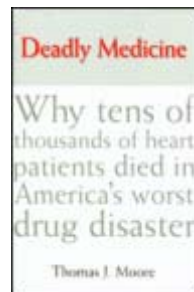
### 处理对现有证据带有偏倚的报告

要**避免偏倚对照**，就要在系统评价中识别并考虑所有相关的可靠证据。这在许多方面都具有挑战性，特别是由于提交哪些研究结果并获得发表，人们可能做出偏颇的决定，从而使得一些相关证据未能得到发表。与其它研究相比，带来“令人失望”或“负面”结果的那些研究得以报告的可能性较小。这常常被称为“发表偏倚”或“报告偏倚”。

人们认识到这些报告偏倚的时间已达数个世纪（Dickersin 2004a）。例如，1792年，詹姆斯·费里亚（James Ferriar）曾经强调，记录治疗失败案例和治疗成功案例一样重要（Ferriar 1792）。一个世纪多之后，这一原则在《波士顿医学与外科期刊》发表的一篇社论中做出了重申（[社论，1909](#)）。

如今有大量证据证实，报告偏倚是一个很大的问题。还有证据表明，报告偏倚主要是由于研究人员没有写成或提交用于发表的研究报告，而不是由于期刊编辑偏颇地拒绝提交的研究报告造成的（Dickersin 2004b）。近期研究还发现另外一个问题：经过研究，如果估计某些结果的治疗效果，与研究者的结论不相符合，这些数据有时也不会报告出来（Chan et al. 2004）。

例如，关于向心脏病发作患者提供药物，减少心律异常的发生，如果药物效果的所有研究都已报告出来，那么由于这些药物而导致的数以万计的死亡病例本可避免。1993年，考利（Cowley）博士及其同事就指出，13年前完成的一项未公布的研究可能已经为“将来的麻烦提供了早期预警”。在被指定服用抗心律失常药物（劳卡胺（lorcainide））的49名患者中，有9人死亡，而服用安慰剂的类似患者人数中，只有1人死亡。“当我们在1980年开展我们的研究时，”他们报告说，“我们认为增加的死亡率是由偶然因素造成的……由于商业原因，放弃了劳卡胺的开发，因此该研究从未发表过；现在，就成了‘发表偏倚’的一个很好的例子”（Cowley等 1993）。



报告偏倚往往导致做出医学疗法比真实情况更有效果的结论。因此，这可能带来不必要的痛苦和死亡，使得花费在毫无效果或具有危险性的治疗措施上的资源白白浪费（Chalmers 2004）。当研究人员请求应试者参与疗法测试时，应试者假定自身的参与会导致知识的增加。这意味着，如果研究人员不公布研究成果，就违反了研究者和参与者之间的契约。研究情况的偏颇少报是科学上的不正当行为，也是不道德的（Chalmers 1990）。制药业赞助研究的选择性报告是一个特殊问题（Melander等 2003），尽管该问题并不仅限于那些商业既得利益相关者。研究伦理委员会、医学伦理学家和研究资助者，到目前为止在保护患者和公众不受报告偏倚负面影响方面还做得不够（Savulescu等 1996）。只要政府和应当保护公众利益的其他方面，继续容忍这种形式的科研不当行为，那么疗法（尤其是那些存在商业利益的疗法）的公平试验就会继续受到损害。

世界卫生组织已提出了一些解决方案，来解决无法识别的研究和发表（或传播）偏倚所带来的问题：首先，该组织将建立标准，用于试验登记数据的登记和交换。其次，还将提议在患者招募工作开始之前在满足上述标准的数据库中登记研究方案。最后，将提议建立一个开放式门户网站（[www.who.int/ictpr](http://www.who.int/ictpr)），搜集整理所有登记者的数据，使得人们了解即将实施、正在实施和已经完成的研究方案。

我们所有人都必须支持世界卫生组织发挥的领头作用，要求所有治疗措施公平试验在开始阶段就进行登记，并坚持研究结果应予公布，从而减少报告偏倚。

**Cite as:** Editorial commentary (2007). 处理对现有证据带有偏倚的报告. The James Lind Library ([www.jameslindlibrary.org](http://www.jameslindlibrary.org)).

**Show JLL records:** illustrating [reporting bias](#)

**Other material:** view [Monty Python's take](#) on reporting bias - YouTube video

**Next essay:** [避免对现有证据带有偏倚的选择](#)

**Select other essay:** [为什么需要进行公平试验](#)



### 参考文献

Chalmers I (1990). Under-reporting research is scientific misconduct. JAMA 263: 1405-1408.

Chalmers I (2004). In the dark: drug companies should be forced to publish all the results of clinical trials. New Scientist 181: 19.

Chan A-W, Hróbjartsson A, Haahr M, Gøtzsche PC, Altman DG (2004). Empirical evidence for selective reporting of outcomes in randomized trials: Comparison of protocols to publications. *JAMA* 291:2457-2465.

Cowley AJ, Skene A, Stainer, Hampton JR (1993). The effect of lorainide on arrhythmias and survival in patients with acute myocardial infarction. *International Journal of Cardiology* 40:161-166.

Dickersin K (2004a). Publication bias: recognising the problem, understanding its origins and scope, and preventing harm. Dans : Rothstein H, Sutton A, Borenstein M, eds. *Handbook of publication bias*. New York : Wiley.

Dickersin K (2004b). How important is publication bias? A synthesis of available data. *AIDS Educ Prev* 1997;9 (1 Suppl):15-21.

Editorial (1909). The reporting of unsuccessful cases. *Boston Medical and Surgical Journal* 161:263-264.

Ferriar J (1792). *Medical histories and reflexions*. Vol 1. Londres : Cadell et Davies, 1792.

Melander H, Ahlqvist-Rastad J, Meijer G, Beermann B (2003). Evidence b(i)ased medicine - selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. *BMJ* 326:1171-3.

Savulescu J, Chalmers I, Blunt J (1996). Are research ethics committees behaving unethically? Some suggestions for improving performance and accountability. *BMJ* 313:1390-1393.

[Home](#)

[目录页](#)

[Comments](#)

[Home](#)[目录页](#)[jameslindlibrary.org](http://www.jameslindlibrary.org)

## 对所有相关证据的系统评价

### 避免对现有证据带有偏倚的选择

偏倚会扭曲医学疗法测试，并导致错误的结论。还会扭曲对证据的审查。应当在研究方案中列出系统评价计划，例如[Cochrane 协作网](#)（Cochrane Collaboration）公布的那些研究方案，明确将采取什么措施来减少偏倚。

其中包括明确说明：

- 在评价中将处理疗法方面的什么问题；
- 使研究符合纳入条件的标准；
- 用于搜寻可能符合条件的研究的策略；
- 为了尽量减少研究选择带来的偏倚，将要采取的步骤；以及纳入评价的数据（Berlin 1997）。

在处理似乎相同的医学疗法效果问题时，不同的系统评价往往会得出不同的结论。有时这是因为针对的问题存在着细微的差别。有时它反映了评价者所用材料和方法上的差异，在这种情况下必须判断那些评价最可能最成功地减少了分配偏倚。

评价者是否存有其它利益，可能影响到他们所作的评价工作或作出的解释，也值得加以考虑。例如，与月见草油制造商相关的人员评价了该药物对湿疹的疗效（Morse等 1989）。他们对该药价值得出的结论，远远比与商业利益无关的调查人员所做的评价更加积极，后者在他们的评估中包括了未发表研究的结果（Williams 2003）。

导致对包含在评价中的现有证据进行偏倚选择，不仅仅是由商业利益造成的。我们每个人都带有偏见性，这可能导致证据的偏倚选择，而研究人员、卫生专业人员、患者和其他评估疗效的人均不能免受其影响。人们已认识到利益冲突的重要性，正在采取一些措施加以考虑。

**Cite as:** Editorial commentary (2007). 避免对现有证据带有偏倚的选择. The James Lind Library (www.jameslindlibrary.org).

**Show JLL records:** illustrating [reporting bias](#)

**Other material:** view [Monty Python's take](#) on reporting bias - YouTube video

**Next essay:** [利用Meta分析减少机遇因素效应](#)

**Select other essay:**

### 参考文献

Berlin JA (1997). Does blinding of readers affect the results of meta-analyses? University of Pennsylvania Meta-analysis Blinding Study Group. *Lancet* 350: 185-186.

Morse PF, Horrobin DF, Manku MS, Stewart JC, Allen R, Littlewood S, Wright S, Burton J, Gould DJ, Holt PJ, et al (1989). Meta-analysis of placebo-controlled studies of the efficacy of Epogam in the treatment of atopic eczema. Relationship between plasma essential fatty acid changes and clinical response. *British Journal of Dermatology* 121: 75-90.

Sackett DL, Oxman AD (2003). HARLOT plc: an amalgamation of the world's two oldest professions *BMJ* 327: 1442-1445.

Williams HC (2003). Evening primrose oil for atopic dermatitis. *BMJ* 327: 1358-1359.

[Home](#)[目录页](#)[Comments](#)

[Home](#)[目录页](#)[jameslindlibrary.org](http://www.jameslindlibrary.org)

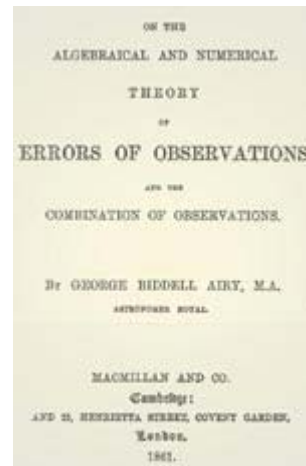
## 对所有相关证据的系统评价

### 利用Meta分析减少机遇因素效应

医学治疗措施的公平试验需要对所有相关的可靠证据进行[系统评价](#)。为避免关于疗效的误导性结论，准备系统评价的人必须采取措施以避免各种类型的[偏倚](#)，例如，[考虑所有相关证据](#)，并[避免对现有证据的偏倚选择](#)。

即使可能已采取措施来减少评价中的偏倚，但关于疗效的误导性结论也可能是由于[机遇因素](#)造成的。在系统评价中每次讨论单独但相似的研究，可能也会由于机遇因素作用而留下困惑的印象。如果可能并合适，可利用一种现在称为“Meta分析”的统计程序统合所有相关研究的数据来减少这个问题。

今天在Meta分析中使用的大多数统计技术都来自德国数学家卡尔·高斯（Karl Gauss）和法国数学家皮埃尔-西蒙·拉普拉斯（Pierre-Simon Laplace）在19世纪上半叶的工作。天文学是他们的方法具有实际用途的领域之一：在多个场合测量恒星的位置往往导致稍有不同的估计，那么需要一定的技术来统合估计，从汇集的结果中得到一个平均值。1861年，英国皇家天文学家乔治·艾利（George Airy）出版了一本针对天文学家的“教材”（[Airy 1861](#)），在书中叙述了这种定量合成过程所采用的方法。一个多世纪之后，美国社会科学家吉恩·格拉斯（Gene Glass）将此过程命名为“Meta分析”（Glass 1976）。



Meta分析的一个早期医学实例由卡尔·皮尔森（Karl Pearson）于1904年发表在《英国医学杂志》上（[Pearson 1904](#); O'Rourke 2006），他应政府要求，负责评价关于某种伤寒疫苗效果的证据。尽管统计学家在随后70年间对Meta分析方法进行了发展，但直到20世纪70年代这些方法才开始得到更广泛的应用，最早是从社会科学家开始的（Glass 1976），然后得到医学研究者的应用（Stjernswärd J 1974; Stjernswärd等 1976; Cochran等 1977; Chalmers等 1977; Chalmers 1979; Editorial 1980）。

Meta分析可以利用[Cochrane协作网](#)的标识予以说明。该标识展示了利用来自七次公平试验的数据所做的一次Meta分析。每条横线代表一次测试的结果（线越短结果越确定）；而钻石代表它们的综合结果。竖线表示试验中比较的两种疗法具有相似效果，则横向将聚集在它附近；如果一条横线与竖线相交，是指特定测试在两种疗法之间没有发现明确的（“统计学意义显著的”）差异。当个别横线与“无差异”竖线相交时，意味着该疗法可能增加或减少婴儿死亡率。不过，两者结合起来可以发现，横线趋向于落在“无差异”线的有益侧（左侧）。钻石代表了这些测试的综合结果，利用Meta分析统计过程生成。钻石的位置显然落在“无差异”线的左侧，这表示该疗法是有益的。



这个图表显示了一种疗程短并且价格低廉治疗方法的公平试验的系统评价结果，该方法是向预计早产的孕妇提供某种类固醇药物。这些测试的首例报告时间是1972年。图表总结了要是十年后（1981年）对可用测试进行系统评价本来可以揭示的证据：它有力地表明类固醇降低了婴儿死于早产并发症的风险。到1991年，还报告了其它七次试验，标识中的图画变得更为清晰。

在1989年（Crowley 1989）之前没有发表关于这些试验的任何系统评价，所以大多数产科医师、助产士和孕妇都没有意识到这种疗法如此有效。毕竟，一些测试并没有显现出“统计学意义显著”益处，而或许只有这些测试被人察觉到了。由于没有进行系统评价，数以万计的早产儿受苦并毫无必要地死去，而资源则浪费在不必要的研究上。这仅仅是给人类带来损失的许多例子中的一个，它未能在公平试验的最新系统评价中评估疗效，利用Meta分析降低[机遇因素](#)造成误导的可能性。

到20世纪末，人们广泛认为Meta分析是治疗措施公平试验的一个重要组成部分。当某种疗法实际上有益或有害时，它有助于避免得出治疗措施没有任何效果的不正确结论。

**Cite as:** Editorial commentary (2007). 利用Meta分析减少机遇因素效应. The James Lind Library ([www.jameslindlibrary.org](http://www.jameslindlibrary.org)).

**Show JLL records:** illustrating [meta-analysis](#)

**Next essay:** [对所有相关的可靠证据进行最新的系统评价](#)

**Select other essay:** [为什么需要进行公平试验](#)



## 参考文献

Airy GB (1861). On the algebraical and numerical theory of errors of observations and the combination of observations. Londres : Macmillan.

Chalmers I (1979). Randomized controlled trials of fetal monitoring 1973-1977. Dans : Thalhammer O, Baumgarten K, Pollak A, eds. Perinatal Medicine. Stuttgart : Georg Thieme, 260-265.

Chalmers TC, Matta RJ, Smith H, Kunzler A-M. (1977). Evidence favoring the use of anticoagulants in the hospital phase of acute myocardial infarction. New England Journal of Medicine 297:1091-1096.

Crowley P (1989). Promoting pulmonary maturity. Dans : Chalmers I, Enkin M, Keirse MJNC, eds. Effective care in pregnancy and childbirth. Oxford : Oxford University Press, pp 746-762.

Editorial (1980). Aspirin after myocardial infarction. Lancet 1:1172-3.

Glass GV (1976). Primary, secondary and meta-analysis of research. Educational Researcher 10, 3-8.

O'Rourke K (2006). An historical perspective on meta-analysis: dealing quantitatively with varying study results. The James Lind Library.

Pearson K (1904). Report on certain enteric fever inoculation statistics. BMJ 3:1243-1246.

Stjernswärd J (1974). Decreased survival related to irradiation postoperatively in early operable breast cancer. Lancet 2:1285-1286.

Stjernswärd J, Muenz LR, von Essen CF (1976). Postoperative radiotherapy and breast cancer. Lancet 1:749.

[Home](#)

[目录页](#)

[Comments](#)

[Home](#)[目录页](#)[jameslindlibrary.org](http://www.jameslindlibrary.org)

## 对所有相关的可靠证据进行最新的系统评价

### 卫生保健治疗措施的公平试验

很少有人把医学疗法个别公平试验的结果系统地置于其它相似研究的背景之中，用相关方法来减少**偏倚**和**机遇因素**。未能对关于治疗效果的研究进行系统评价，带来了大量本可以避免的痛苦。[卫生保健治疗措施的公平试验](#)，还要求不带偏倚地准备针对被评估疗法所有相关可靠研究的系统评价。

关于这一程序的一些例子可以追溯到200多年前。例如，1753年，詹姆斯·林德在回顾关于坏血病预防和治疗的报告时，指出：

“摒除成见是一件很不容易的事情，……必须全面而公正地展示迄今关于坏血病的已发表观点……事实上，在明确而适当地定位这一课题之前，必须清除大量的垃圾。” ([Lind 1753](#))

人们日益看到，对涉及疗效问题的所有相关研究的系统评价，可提供关于疗效结论的最可靠的依据。有时系统评价会表明不存在可靠证据，而这正是它最重要的功用之一。相仿，系统评价有时可能证实可靠证据仅限于某一单个研究，而这里弄清楚这个情况也很重要。



提供治疗措施公平试验需要系统评价这一认识，体现在以纸质和电子方式发表的系统评价报告数量的迅猛增长上（疗效评价文摘数据库(DARE)：[Cochrane协作网](#)）。系统评价正被用于（1）为临床实践提供信息，例如，通过英国医学杂志出版集团（BMJ）的出版物《[临床证据](#)》（Clinical Evidence）和[苏格兰学院间指南网络](#)（Scottish Intercollegiate Guidelines Network）；（2）评估哪种医学疗法最具成本效益，例如通过[英国国家卫生与临床研究院](#)（National Institute for Health and Clinical Excellence）；（3）满足患者有关提供疗效可靠信息的需求，例如，通过[Informed Health Online网站](#)和[英国国家卫生图书馆](#)（National Library for Health）。

### 未竟之业

所有这些及类似进展表明，那些努力改善医疗保健领域信息选择所需证据的获取途径的人已认可了系统评价的重要性。然而，前面还有很长一段路要走：据估计，Cochrane协作网目前产出的几千项系统评价需要增加到超过10,000项才能覆盖现有证据（Mallett和Clarke 2002）。然后，随着新证据的涌现，还应不断加以更新。事实上，一位杂志主编提出，在我们掌握现有证据能够告诉我们的内容之前，应当暂停一切新的研究活动（Bausell 1993）。

那些负责拨付研究经费的人，必须确保提供资源以应对这种积压局面，并且只有当关于现有证据的系统评价表明有必要进行更多研究，并且这些研究的设计考虑了以往研究的经验教训时才支持新的研究。如果杂志主编们想要更好地为读者服务，那他们就必须效仿《柳叶刀》所作的引领，确保新的研究报告明确，获得的新证据对所有相关证据的最新系统评价有何作用（Young和Horton 2005）。

人们越来越容易得到最新的系统评价，这改善了疗效信息的质量，但是系统评价的结论不应不加批判地接受。专门针对相同治疗问题的不同评价有时可能得出不同的结论。这些评价的作者是人，我们必须意识到他们可能以支持各自偏见和利益的方式来选择、分析和出示证据。随着编写并维护系统评价的可靠方法的不断演变，会有助于处理这个问题，但是不能指望靠它们来彻底解决问题。

尽管系统评价数目的增长提高了卫生保健领域主要治疗措施公平试验的可得性，但是这些评价往往也会揭示出许多疗效研究的质量不高且毫不相干。正如一位社论作者关于“不良医学研究丑闻”的评论所述，我们需要更少、更好的研究和具备正当理由的研究（Altman 1994）。没有公众对治疗措施公平试验的逻辑依据及其特征的更多了解，没有公众在治疗措施公平试验各个阶段的更大影响和更广泛参与，这一目标似乎不可能实现。这一议程的推动取决于患者和临床医生组成的新联盟所面对治疗效果的不确定性（Chalmers 2004；[www.duets.nhs.uk](http://www.duets.nhs.uk)；[詹姆斯·林德联盟](#)）。

如果公众和卫生专业人员能够更方便地获取关于所有相关可靠证据（涉及关于疗效的重要不确定因素）的最新系统评价情况，以及针对这些不确定因素正在进行的研究的相关信息，那他们就获得了良好的服务（Smith和Chalmers 2001）。

**Cite as:** Editorial commentary (2007). 对所有相关的可靠证据进行最新的系统评价. The James Lind Library ([www.jameslindlibrary.org](http://www.jameslindlibrary.org)).

**Select other essay:**

[为什么需要进行公平试验](#)



## 参考文献

Altman (1994). The scandal of poor medical research. *BMJ* 308:283-284.

Bausell BB (1993). After the meta-analytic revolution. *Evaluation and the Health Professions* 16:3-12.

Bunker JP, Frazier HS, Mosteller F (1994). Improving health: measuring effects of medical care. *Milbank Quarterly* 72:225-258.

Chalmers I (2004). Well informed uncertainties about the effects of treatments: how should clinicians and patients respond? *BMJ* 328:475-476.

Lind J (1753). A treatise of the scurvy. In three parts. Containing an inquiry into the nature, causes and cure, of that disease. Together with a critical and chronological view of what has been published on the subject. Edimbourg : imprimé par Sands, Murray et Cochran pour A Kincaid et A Donaldson.

Mallett S, Clarke M (2002). The typical Cochrane Review. *International Journal of Technology Assessment in Health Care* 18:820-823.

Smith R, Chalmers I (2001). Britain's gift: a 'Medline' of synthesized evidence. *BMJ* 323:1437-1438.

Young C, Horton R (2005). Putting clinical trials into context. *Lancet* 366:107-8.

[Home](#)

[目录页](#)

---

[Comments](#)