针灸治疗慢性膝关节疼痛真的无效吗？

【说明】本文对 JAMA 针灸报告的结论提出质疑，同时指出了文章作者在临床试验研究中的不规范行为，并提供了相关证据。希望有更多的学者参与探讨针灸疗效的真相。成文中参考了多位专家和医生的意见，特此一并感谢。

2014 年 10 月 1 日，美国医学会会刊（JAMA）发表了澳大利亚学者 Rana Hinman 等 14 位作者的文章。题目为“针灸治疗慢性膝关节疼痛：随机临床试验”。文章的结论是：“对于五十岁以上患有中度或重度膝关节慢性疼痛的患者，同假治疗相比，激光针灸或针刺针灸治疗对改善疼痛或功能没有益处。我们的研究发现不支持对这些病人使用针灸治疗。”

此文一发表，就受到了媒体的广泛关注，很多主流媒体和网站都发了消息和评论。Hinman 博士在接受媒体（Reuters）采访时强调，虽然她们的研究证明针灸治疗膝关节疼痛无效，不过还是有其它有效方法可以选择的，比如物理治疗，膝关节托架，身体锻炼等。

此消息在中医针灸界也马上引起的强烈反响，因为这个结果同针灸医生的临床经验，患者的亲身体验，以及近几十年来大量的基础实验和临床试验研究结论不一致。

我在开始读 JAMA 文章时虽然发现多个问题，但对试验设计的初步印象很好，只不过是有点奇怪的感觉。为此还曾征求了其他医生和专家的意见，大多数人都感到实验设计没有问题，甚至很漂亮。但再次认真仔细读，我突然意识到可能被作者误导了。

主要作者们几乎全是科研博士 PhD，没有临床医学专家，也没有针灸的专业人士，以从事运动康复和物理治疗的研究人员为主。我的第一个问题是，他们为什么要研究针灸的疗效呢？

细读一下相关背景资料，发现原来他们是对激光疗法感兴趣。按照大多数国家的法律，没有针灸执照的“治疗师”是不能刺破皮肤的。所以物理治疗师和练身教练们感兴趣的是用“激光针灸”代替金属针具，想证明激光同针刺效果一样或更好。

需要说明，目前针灸医学界对激光针灸还没有普遍认同，不能将激光穴位治疗等同于针刺治疗，或称为“针灸”，我所认识的针灸师中几乎无人使用“激光针灸针”治疗患者。

看出这个作者动机了，下面的分析推理就通畅了。

作者当初一定是设计了这样一个临床治疗慢性膝关节疼痛的试验：

目的：用对照试验证明激光刺激穴位治疗慢性膝关节疼痛有效。共设立四个治疗观察组。

1 激光针灸治疗组：为需要验证的疗法组
2 不治疗组：为阴性对照组
3 假激光针灸治疗组：为激光安慰剂对照，也是阴性对照组
4 针刺治疗组：为阳性对照（因为以前大样本临床试验证明有效，也是激光针灸的比较对象）

特别说明，Hinman 博士文中所称的“激光针灸”（laser acupuncture），其实不是“针”也不是“灸”，本人不能认同此名，正确的称法应该是“激光疗法”（laser therapy）。

科学研究常识：动物实验或人体试验成功的基本条件是，结果数据要显示，阴性对照无效，阳性对照有效（达到预期的指标）。这样才能评价需要验证的疗法，即激光针灸。

但 Hinman 博士及其他作者们在采集试验数据后分析发现，结果同她们预想的不一样：
“针刺阳性对照组”同“假激光阴性对照组”的治疗效果无统计学差异，而她们原来以为一定是针刺优于假激光，并以此差异作为参考基线来评价激光针灸的疗效。但是“针刺阳性对照组”优于“不治疗阴性对照组”，有明显统计学差异。另外，针刺、激光和假激光之间都没有差别，但三者都优于不治疗组，且有统计学差异。上述结果，尤其是第一点，是她们原始试验设计中没有预料到的。

如何解释这些结果呢？

正确的解释应该是：试验失败，而不是疗法失败，因为阳性对照没有达到预期的全部效果。原因可能是试验设计上或实施过程有问题，比如样本量不够、针刺治疗不足及试验方法有缺陷等。尤其是作者使用的 Zelen 试验设计方法是非常有争议的和有很多局限性的方法。（关于失败的原因见附文）

接受试验失败的结果是承认原始假设不能得到证实或否定，试验数据没有学术价值，文章也不可能发表在好的杂志上。这样的结局对课题主持人是沉重的打击，等于浪费了数年的时间和大量的经费。

但 Hinman 博士想出了另一个解决办法，不但可以不承认试验失败，还能将结果发表。

即改变原始的试验假设和试验目的，将“针刺治疗”（needle acupuncture）也变成需要被验证的疗法，而不是原来设计的阳性对照。将“假激光针灸”组作为针刺治疗组的安慰剂对照组（虽然这是完全错误的，不可比的，也不在原始计划中），以替补原设计中缺乏“假针刺”对照组之缺陷。

这样比较激光同假激光及针刺同假激光的治疗效果，结论只有一个：
激光针灸和针刺治疗慢性膝关节痛都无效。由试验失败，变成了疗法失败。

其实，如果根据此论文中的数据，比较针刺疗法同不治疗组，结论应该是针刺疗法治疗慢性膝关节疼痛有效。但这个结论并不全面，因为缺乏“假针刺对照组”。而作者最初没设这个对照组的原因是，针刺治疗是阳性对照，没有必要为阳性对照再设对照。

问题的关键：为了挽救验证激光针灸试验的失败，作者采用新增试验目的和替换对照组的方法，“证明”针刺疗法和激光疗法都无效。在试验数据出来后，作者将原设计中的阳性对照针刺治疗改成被需要验证的治疗方法，还错误地使用假激光作为针刺治疗对照，得出针刺治疗慢性膝关节疼痛无效的结论。这样一个失败的试验就变成了失败的疗法。

我们认为，作者在试验结束后，修改原始试验假设和目的，按数据结果再新增错误的治疗组间比较，而且在最后报道时未做说明，这种行为有可能违反了科学研究的诚信规则，应该属于科研行为不规范，甚至可能是科研行为不端（misconduct）。由此产生的结果不但误导公众和患者，伤害针灸业的发展，还有损于医学科学研究及 JAMA 杂志的信誉。如果文章作者不能合理地解释上述问题，此论文应该从杂志撤销，JAMA 杂志应该重新发布新闻公告。作为世界一流的临床医学杂志，JAMA 在审稿和编辑中没有发现上述不端行为，实在令人遗憾。

那么以上分析和推理证据何在？

证据：

A. 2009 年，当作者于申请澳大利亚国家基金以资助此项临床试验时，在研究计划中明确说明，试验的目的有三项：（1）比较“激光针灸”和“假激光针灸”的疗效；（2）比较“激光针灸”与“针刺治疗”的疗效；（3）比较“激光针灸”、“假激光针灸”及“针刺治疗”同“不治疗组”的效果。（见表 1 英文原文）在原始试验设计中根本没有计划要比较“针刺治疗”与“假激光针灸”。而作者 2014 年最后发表论文时的结论“针灸治疗慢性膝关节疼痛无效”，正是通过比较“针刺治疗”和“假激光针灸”得出的。
B. 2012 年，在试验数据出来之前，作者曾发表过整个试验设计及试验有关的基本资料。（见 2012 BMC CAM Journal ，JAMA 文章参考文献 14）文中列出了 9 个试验假设和目的（即要进行的组间比较）。如表 2 显示，所有的 9 项假设试验比较都是比较“激光针灸”同其他治疗组的差异，其中没有一个假设提到要比较“针刺治疗”同“假激光针灸”的效果。可见作者也认为二者不可比，有差异应该是试验成功的前提，不是试验的目的，所以试验设计中也根本没有设计这个比较。在计算样本大小和设计“盲法”时也根本没有考虑到这个比较。此比较和试验目的完全是作者在看到数据结果后再造出来的。
C. 作者对事实的掩盖：作者在 JAMA 文章的讨论部分的最后一段，解释为什么这个临床试验当初没有设立“假针刺对照组”，是因为假针刺不能盲针灸师；假针刺可能产躯体感觉作用；而且过去已有了假针刺研究。这种说法显然与事实不符，因为上述证据显示，原始设计的假设和目的中根本就没要比较针刺同假治疗（安慰剂）的疗效。而整个临床试验研究的主要目的自始至终都是要验证激光疗法的效果，针刺疗法仅仅是个阳性对照。

为什么“假激光”不能作为针刺治疗的对照组？
“假针刺治疗”同“假激光治疗”根本不是一会事。安慰剂在临床上的作用有很多种，效果和程度也可以不同，不能将所有的安慰剂都视为具有同样效果。比如“假开颅手术”治疗一些疾病有效，但临床上不能用“安慰剂药片”（sugar pill）来取代。尚没有试验证明假针灸同假激光效果相同。而且，此试验中使用的激光治疗仪是特殊定制的，由专用计算机操作的，对患者的安慰剂作用完全是未知数。

以上质疑主要针对作者科研行为不规范，或为学术不端。关于论文在设计、疗法、患者、治疗师、疗效观察、及数据处理等方面的问题，见附文：“针灸治疗慢性膝关节疼痛试验失败的原因分析”。

李永明 MD PhD
2014 年 10 月 7 日草，10 日修改于 NJ

附文：

表 1“针灸治疗慢性关节疼痛计划”在澳大利亚和新西兰临床试验注册摘要(11/2009)

表 2“针灸治疗慢性膝关节疼痛”试验设计中的 9 个假设（Hinman, BMC CAM Journal, 2012）

“针灸治疗慢性膝关节疼痛试验”失败的原因分析

表 1.“针灸治疗慢性关节疼痛计划”在澳大利亚和新西兰临床试验注册摘要：(11/2009)

The study firstly aims to compare the effectiveness of laser acupuncture to sham laser acupuncture in people suffering from chronic knee pain. Secondly, the study aims to compare laser acupuncture to needle acupuncture in people suffering from chronic knee pain. And thirdly, the study aims to compare laser acupuncture, sham laser acupuncture and needle acupuncture to no treatment for chronic knee pain. The study hypothesis is that laser acupuncture will be more effective than sham laser acupuncture or no treatment and will be equal or more effective than needle acupuncture in treating people suffering from chronic knee pain.

表 2.“针灸治疗慢性膝关节疼痛”试验设计中的 9 个假设(Hinman, BMC CAM Journal, 2012）

试验假设	疗效比较目的	时间	指标	试验结果*
1.主要 H1	激光针灸优于假激光针灸	12 周	关节疼痛和功能	否
2.主要 H2	激光针灸和针刺治疗优于不治疗组	12 周	关节疼痛和功能	是
3.主要 H3	激光针灸优于针刺治疗	12 周	关节疼痛和功能	否
4.次要 H4	假激光针灸优于不治疗组	12 周	关节疼痛和功能	是
5.次要 H5	激光针灸、假激光针灸和针刺治疗优于不治疗组（同时激光针灸优于针刺治疗）	12 月	关节疼痛和功能	否
6.次要 H6	激光针灸、假激光针灸和针刺治疗优于不治疗组	12 月	站立痛和功能	否
7.次要 H7	激光针灸优于针刺治疗	12 月	站立痛和功能等	否
8.次要 H8	激光针灸、假激光针灸和针刺治疗优于不治疗组	12 月	整体健康	否
9.次要 H9	激光针灸优于针刺治疗	12 月	整体健康	否

*试验结果由本文作者根据 2014 JAMA 文章结果加入。Hinman 等在 JAMA 论文中没有提及这些假设和结果。

附文： “针灸治疗慢性膝关节疼痛”试验失败的原因分析

Hinman 博士文章所报告的数据显示，激光针灸和针刺治疗同假激光组比较没有统计学差异。阳性对照（针刺）同阴性对照（假激光）之间也没有差异，没有达到试验成功的基本标准，其原因可能是多方面的，主要同下列因素有关。

1. Zelen 试验方法有一定局限性，并不适合针灸临床试验。由于患者进入试验前没有被告知针灸或激光试验，导致分组后 19%（13/70）的病人拒绝针灸治疗，而这些未经治疗的病人还被留在针灸治疗组观察，参加最后针灸治疗组的统计分析。此方法稀释了针刺的疗效，增加了假阴性的可能，还造成同其他组的不公平的比较（假激光组中有 13%（9/70）拒绝治疗）。另外，研究人员在病人最初参加临床试验时对病人提供了虚假的信息，此研究方法不但在医学伦理上存在问题，还可能产生“反安慰剂作用”（nocebo effect）。关于 Zelen 方法的局限性和在此试验中使用的问题，我们曾咨询过临床试验统计学专家，得到相似的意见。
2. 非专业针灸师从事治疗，专业培训和经验可能不足或过于混杂。此试验选择 8 位受过针灸训练的西医师（GP）操作针灸治疗，受过的训练包括传统中国针灸和西方针灸。这些医生通常忙于西医临床工作，用在针灸治疗的时间远远低于专业针灸师（acupuncturist）。每次治疗仅用 20 分钟也低于专业针灸师的治疗时间。这些因素都难免影响临床疗效和试验结果。
3. 针刺方法不规范和治疗剂量不足。在 282 位进入试验的患者中，仅有 54 位患者（19%）完成了 12 周的针刺治疗。其中平均治疗次数不到每周一次（8-12 次），每次仅针刺治疗 20 分钟。而且还允许病人中途因假期等原因停止治疗达 4 周，对患者用药情况没有限制或详细记录。病人大部分身体超重肥胖，病程久，但使用的针灸针相对较短和细，对针灸师没有手法和患者“得气”的要求，而且没有使用电针。设计中的穴位处方混乱，医生随意性很大（中式和西式针灸混合）。因此，针刺刺激剂量明显低于常规临床使用剂量，每次治疗时间和疗程也短于发表过的同类研究报告。
4. 针灸治疗师过多且分散，难以保证统一治疗方案和稳定疗效。54 名患者的针刺治疗由 8 名针灸治疗师分别在 9 个诊所，在 2 年期间完成。平均每位针灸师治疗 6.8 人。这种分散的治疗队伍很难保证统一的质量，如有 2-3 位针灸师不按规则操作或治疗技能欠佳，可能会影响到整个试验结果。
5. 样本量不足，不能检测小幅度但具有临床意义的差别。根据已经发表的大量针灸临床研究报告，针刺治疗同不治疗组比较优势一般比较稳定，但针刺治疗同安慰剂比通常较差异较小，需要较大的样本量才能显示真实疗效。此试验中只有 54 人完成了 12 周的针灸治疗，而 81%的患者分在对照组和激光针灸组。尤其是在最后分析数据时，作者将所要验证的疗法从一项增加到两项（激光针灸和针刺疗法）。这好比要用一个 282 人的样本测试两个新药的疗效，而样本量原来是按一个新药而设计的。
6. 评估时间及方法不符合临床实际。患者自我评估仅有两次，12 周和 12 个月。而在临床实践中，针灸师对患者临床疗效的评估频率远远高于此。针刺治疗的疗效有短期疗效和长期疗效，而短期疗效，如几天到几周，也会有益于患者的生活质量。尤其是对于大多数中重度膝关节慢性疼痛的患者，在病因没有消除的情况下，期望 10 来次针刺治疗就能产生长达一年以上的疗效是不现实的。目前也没有任何西药能达到如此疗效。