

国际期刊科学出版理念的更迭

中华医学会杂志社 姜永茂 刘冰中华医学杂志英文版 季媛媛中华医学会杂志社市场部 包靖岭慢性疾病与转化医学 魏佩芳



International Congress on Peer Review and Scientific Publication

Enhancing the quality and credibility of science

September 10-12, 2017 | Chicago, USA

会议概述

四年一次,会期3天

主办方: 美国医学会杂志(JAMA)、英国医学杂志(BMJ)

主旨: 促进对同行评审和科学出版质量及可信度的研究。

参与者: 40多个国家,500多名代表

期刊编辑,出版商,图书馆情报专业人员,科研人员,

医学生,以及临床医生。

老主题:

同行评议、文献质量、报告规范、出版伦理、临床试验注册、数据共享,等等

新方向:

同行评议:多种同行评议形式的比较(单盲/双盲/Open/Inline)

报告规范: 已从"向作者介绍、推广"转为研究"推广后的初步成效"

临床试验注册:注册平台中的结果与文章发表结果的一致性

利益冲突: 关注"披露的程度"

撤稿: 撤稿后的效应

数据共享:初步实施情况

Peer Review: Are We Looking at the Right Things?

--- 悉尼大学的Lisa A. Bero

1. 披露目前同行评议存在的问题:

- ① 执行起来有难度
- ② 花费时间
- ③ 并不能避免研究偏倚
- ④ 同行评审的意见可能带来更多的偏倚
- ⑤ 让读者更具批判性
- ⑥ 是 "必要的"但并不"足够"

Lisa认为,目前:

(1)我们**信任**同行评议,但我们**并不确定**,同行评议到底应该是什么样的?

目前还有很多种的类型:

- □ 预印本
- □ 出版后同行评议(或组合型)
- □ 开放型
- □盲法评议
- □ 级联型
- □ 外包型

AQ:到底哪种类型好?充满争议

(2) 我们对同行评议的过程了解很多,但对同行评议的结局知道得很少:

- □ 审稿人建议了多少大修改,多少小修改
- □ 审稿的质量如何
- □ 审稿结果不一致的发生率如何
- 审稿的时间花费如何
- □ 盲法审稿的效果如何
- 同行评议还有哪些不足

(3)同行评议也可能导致非常错误的结果

Siler et al PNAS 2015 的研究显示:

采用和拒稿的文章引用次数分布:虽然拒稿的文章可能引用最少,但引用次数到最高时的分布也往往是拒稿的文章。采用的文章在引用次数上分布属于中庸段。

2. 资助情况和利益冲突存在的偏倚

- 证据如何
- ② 需要更多的披露
- ③ 即便是披露了也不一定减少偏倚
- ④ 很多"高调的"研究是由公司或与公司有经济关联的研究者进行的。

为减少资助型利益冲突,期刊和同行评议能做些什么?

- ① 停止出版公司赞助的研究(在国内可行么?)
- ② 惩罚那些未准确披露经济关联的作者和公司(如何惩罚?)
- ③ 向审稿人提供所有研究资助来源以及其他利益相关披露
- ④ 要求审稿人运用工具来评价资金来源及作者利益冲突 作为偏倚危险因素
- ⑤ 对于有利益关联的研究,招募专业审稿人来专门分析研究设计和数据分析

3. 介绍了一个概念 "Spin"

- ① 什么是Spin ? 有偏倚的报告呈现和解释 , 以使结果看起来更有利 (夸 大或选择性的解读结果) 。
- ② Spin如何导致的?(1)调查者和研究的个性特征(2)很少重视的系统性问题
- ③ 同行评议过程能捕捉住Spin吗?编辑和审稿人可以先介绍Spin.

为减少Spin的影响,期刊和同行评议能做些什么?

- □ 开发一个可适用于所有研究种类的针对Spin的清单
- 审稿人是否花费时间审理:方法部分,结果部分,以及补充材料了?
- □ 去掉讨论部分
- □ 出版多元化讨论



1. 利益冲突和同行评议相关偏倚

(1)生物医学研究中披露COI普遍性和程度 — 悉尼大学药学院&哈佛医学院

结果

- □总共只有<1/4的文章披露了COI,这与2003年的调查结果相似。
- □披露COI的文章主要集中在某些类型的文章
- ※所有文章类型中:评论性文章(31.4%)>原创性文章(19.8%)>系统综述或荟萃分析(10%)
- ※原创性文章类型中:药物相关(31.5%)>仪器设备/外科相关(22.3%)>非药物或仪器(15.4%)
- □目前COI的实践情况中,COI声明没有结构化,披露的形式和程度也不一致。
- ※130多种方式表达"无利益冲突"
- ※约40%的COI声明包含了传记式信息或目前的资助信息
- ※COI声明内容/形式高度不一致,很难定位到
- □ 发现IF、Altmetric Score越高的期刊,关于COI披露详细度程度也越高。
- □需要规范化对COI的披露,应遵循ICMJE中关于COI的建议

结论

目前生物医学研究中对利益冲突应**披露的程度**,还缺乏理解:

- □披露的COI只是实际存在COI的"代理品"
- ■还只依赖于"自我报告"
- □文章中披露COI实践过程中"形式多样化"、 "程度深浅不一"

1. 利益冲突和同行评议相关偏倚

(2)公司资助及其他经济上的利益冲突对系统综述结果和质量的影响——丹麦Odense大学循证医学中心

结果显示很奇妙:

*与没有利益冲突的系统综述相比,有药物或设备相关利益冲突的系统综述得到的结论显得"更有利",且方法学质量在一定程度上"更低"。

*并不清楚是否是经济利益冲突对综述的结果造成了影响。

1. 利益冲突和同行评议相关偏倚

(3)Nature系列杂志中作者选择"单盲"vs."双盲"的情况和最终决策结果—Springer Nature出版集团

结果很有意思:

- Nature系列杂志各层次的刊(Nature/姐妹刊/Nature Communications)都显示作者选择"单盲"比例远多于"双盲"
- 倾向选择双盲审稿的有典型特征:
 - 1. 投给杂志层次越高的杂志,选择双盲的越多;
 - 2.作者来源于声望越低的研究机构,选择双盲的越多;
 - 3.某些特殊的地区作者(印度第一,中国第二),选择双盲的会偏多一些。
- □ "双盲" 审稿与更低的稿件通过率相关。(可能因审稿偏倚?可能因审稿质量?)

PS: "指导"中国作者选择单盲

(1)纳入ClinicalTrials注册平台中的信息对系统性综述结果的影响

—美国Brown大学/Johns Hopkins/Stanford 临床循证中心

结果显示:

- ◎分析了五个临床学科领域,只有24%的研究在ClinicalTrials上注册了,有38%的注册平台上注册的结果并未在同行评议期刊上发表结果。这些信息的缺失对系统综述得到的结果造成的潜在影响还不清楚。
- ◎既在ClinicalTrials注册了,又在期刊上出版了的研究中,有37%的研究结果是不相符的。这就引发了因选择性报告临床结局造成结果偏倚的担忧。
- ◎在注册平台上对研究初步临床结局可变性的预先描述并不会抑制在文章出版时报道其他结局。

引发的思考:

- ◎应考虑制定 对临床试验结果测量说明和报告的 规则。
- ◎期刊和一些检索工具 可以促进 系统综述纳入注册平台上的研究信息,比如可以通过采用更标准化的形式来列出注册号。

(2)发表随机临床试验初步/暂时性结果引起的偏倚

—美国Dartmouth学院临床实践研究所

结果显示:

- ◎只有一半(57%)的注册临床试验报道了最终结果
- ◎最终结果 与 初步结果 得到的结论很可能不同
- ◎在不同级别的杂志中,报道临床试验的 初步/最终结果 的比例差不多
- ◎如此多的对临床试验最终结果未报道的情况很可能会导致出版偏倚,因为真正最后的疗效还不清 楚。

期刊可做的工作:

- ◎限制/减少对暂时性初步结果的报道(尤其在未预设的情况下),以减少出版偏倚
- ◎当报道临床试验是初步成果还是最终结果时,须标明清楚并对此评价。

比如: 在标题末注明 "****: Unplanned/planned interim analysis"

***: Final results"

(3)以评价卵巢癌生物标记物的临床研究为例对Spin的识别和分类

—新西兰公共卫生研究所

- ◎在这个领域中, Spin普遍存在于摘要、结果、结论部分。
- ◎Spin的各种类型很多样(从多至少):
 - 1.结论说某标志物有什么用途,但这个用途与研究目的并不符,而且文章也证明不了这个用途;2.对结果的错误呈现,比如P>0.05却说有效果;3.对研究人群中生物标志物的运用推广到所有人群,比如研究的健康人群某指标有筛查作用,就同理推广到有症状人群中也有筛查作用;4.研究目的和所得结论不相符;5.摘要中的结论和正文结果部分呈现的不符;6.摘要中结果数据与正文结果数据不符;7.得到研究中并未调查的其他利好结论,比如并未比较医疗费却说能减少经济花费。
- ◎常导致Spin的原因有:
 - 1.未预先清楚阐述规范化的研究假设; 2.未计算最小样本量; 3.未预先阐明阳性阈值; 4.未写明统计值; 5.在摘要和正文结果部分,选择性报道一些有显著性的结果。
- ◎这种对结果膨胀地或选择性的报道将导致研究过程中巨大的浪费。需要一些策略来抑制这些 夸张的报告,以增强研究的质量和可信度。

(3)以评价卵巢癌生物标记物的临床研究为例对Spin的识别和分类

—新西兰公共卫生研究所

如何减少Spin,建议可做的工作有:

- ◎研究开展前:注册Protocol
 - ◎研究完成后,运用reporting guidelines
- ◎加强同行评议
- ◎培训科研工作者
- MiRoR Project : Methods in Research on Research

(4)综述已发表生物医学文献中存在的Spin

—悉尼大学药学院

- ◎Spin在发表文献中普遍存在,种类多样
- ◎列出Checklist有效
- ◎培训科研工作者
- ◎编辑和审稿人应熟悉各领域各种类的Spin,知道哪些因素会导致?哪些会抑制?
- ◎ 下一步还须研究: Spin对读者做出决策有什么影响?

3.完整性和学术不当 引起的偏倚

(1)荟萃分析在从原始纳入文献中移除撤销论文后的效应

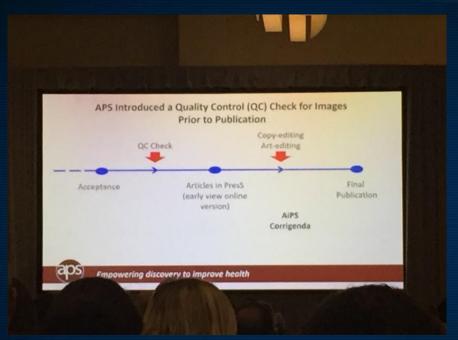
—斯坦福大学数据研究中心

- ◎撤稿对Meta分析造成的影响,总的来说很小,但在个别情况下也可能不同
- ◎对纳入文献量越少的Meta分析中,移除撤销稿件造成的影响越大

3.完整性和学术不当 引起的偏倚

(2)评价在生理学杂志的研究型论文出版过程中加入数字图片控制系统后的成果

一美国生理学学会



◎在稿件接收后处理流程中增加一个图片控制系统来审阅图片帮助作者改图,能有效提升图片质量并减少 发表后出错更正

◎在引入图片控制系统后,系统提示修改的问题总数量逐年下降。提示,再次给杂志投稿的作者作图也越来越符合杂志图片完整性规范,起到了很好的教育作用

3.完整性和学术不当 引起的偏倚

(3) 各种研究报告完整性不足类型

—Springer Nature研究完整性调查组

◎调查了300个生物医学杂志 投稿稿件+发表稿件:

伦理声明/患者知情(35%) > 剽窃/重复发表(23%) > 数据问题(18%) > 作者问题(11%) > 同

行评议问题(10%) > 利益冲突问题(4%)

◎仅投稿稿件:49% 伦理声明/患者知情

◎已发表稿件:41%数据问题

思考:

- ◎出版前高比例的伦理知情问题,早期就可检测出,应培训研究者重视
- ◎出版后高比例的数据问题,可能在同行评议时审稿人并没有总是去检查数据,反而只有在出版后数据问题才会曝光。

4.数据共享(期刊、大学共享数据库、临床试验注册中心三个角度)

(1)期刊数据共享政策的早期实施成果

一美国耶鲁大学医学院&BMJ

◎研究了三大要求数据共享的刊物(PLOS , the BMJ , Annals)中对数据共享的执行情况:

54%有共享计划,29%文中写了共享计划

48%在患者知情同意书中未告知原始数据将共享

◎在期刊要求共享后,作者会进行数据共享。但共享意愿程度取决于要求获取数据的类型和目的。

用于荟萃分析,96.7%愿意;重复原始研究结果,73.3%;

用于二次结果分析:如果数据需求者未计划相似的数据分析计划,77.8%愿意;有相似数据分析计划,仅16.7%愿意;

- ◎作者愿意花费准备共享数据的时间 18 (3-125) person-hours
- ◎障碍存在,仍然面临着很多挑战:准备不足;当数据获取者与研究者有利益冲突时,研究者往往不愿意 共享数据

4.数据共享 (期刊、大学共享数据库、临床试验注册中心三个角度)

(2)Yale University Open Data Access (YODA) Project 进行数据共享早期实践
—美国耶鲁大学医学院

- ◎实践证明了 研究者的数据共享需求
- ◎随着资助机构及研究者对数据共享行为以及对共享数据的运用的增加,很有必要进一步完善共享实践。

4.数据共享(期刊、大学共享数据库、临床试验注册中心三个角度)

(3)Clinical Trial上计划共享数据的声明

—美国NIH ClinicalTrials.gov

◎68%回应了有关计划共享的询问:其中13%计划分享,26%还未决定。

◎思考:在共享原始数据和其他相关文件成为临床研究者常规实践一部分之前,还必须要大量的文化思想 上和科学措施上的改变。



1.报告质量(评估报告规范与临床试验注册初步成果)

(1) '影像学期刊中 期刊水平和研究水平'与'原创研究结果与荟萃分析估算值贴合度'的关联分析

一加拿大Ottawa大学医学院

◎主要运用STARD规范标准评价影像学诊断性研究报告的质量

◎结果:原创研究结果与荟萃分析估算值贴合度 与高分vs.低分杂志 无差异

◎思考:是否符合写作规范该有谁来查证:编辑?审稿人?作者?

1.报告质量(评估报告规范与临床试验注册初步成果)

(2) 在 '高影响力期刊中刊发的临床试验报道'与 'ClinicalTrial注册平台中' 的差异

一加拿大McMaster大学循证医学中心

研究了综合类医学IF最高的5个杂志中发表的200个RCT

结果

- ◎在研究设计(3%)、研究类型(3%)、干预类型(5%)、研究组设置(12%)、初步研究成果(15%) 这些方面,两者差异比较小
- ◎但在研究开始时间(43%)、研究资助方(54%)、次级终点测量(58%)、数据检测委员会(70%)、初步完成日期(70%)这些方面,差异巨大
- ◎高分杂志中刊发的文章在最终发表时与试验注册时就存在很大差异,但这种现象不一定适用于所有期刊or 低分期刊

建议:

- ◎即便是高影响力期刊,也应认真检查论文与注册时的比较情况
- ◎研究者认为该由Editor和Reviewer来承担责任检查差异
- ◎建议临床试验注册研究发表论文应与注册地址做一个链接。

AQ:两者不符的情况应该由谁来承担核查责任?争执和推脱:研究者 🥌 编辑/审稿人

1.报告质量(评估报告规范与临床试验注册初步成果)

(3)评价基于临床实践的系统综述的方法学和报告质量 —美国Oklahoma州立大学健康中心

用PRISMA(系统性综述的推荐报告规范)和AMSTAR(一种评价系统综述的测量工具)这两种工具来评价系统综述。

结果显示:

- ◎在方法学和报告质量上还是不够的
- ◎比如PRISMA条目15&22,AMSTAR条目10&11尤其需要注意是否缺失以便减少偏倚和利益冲突

2.科学文献的质量

(1)关于治疗干预措施的系统综述中统计方法运用的问题 —澳大利亚Monash公共卫生学院

结果显示:◎系统综述中统计方法存在普遍问题

建议:

- ◎应增强综述方法学专家和编辑的合作
- ◎经费资助者应要求作者进行系统综述时必须有统计学专家的把关
- ◎综述相关指南比如PRISMA须进一步更新扩展,来协助提供更多统计学指导
- ◎在投稿过程中加入专门检测统计学的工具
- ◎建议用Checklist帮助检查论文

2.科学文献的质量

(2)放射学杂志中报道规范和清单是否有用?

—德国某医院放射科

- ◎STARD和STROBE用的多(与专业有关)
- ◎大部分作者和半数审稿人认为报告清单有用。
- ◎文章撰写时运用清单对稿件最终刊发后的效果是有关联的。(虽然规范里都说不作为是否评价研究质量的工具)

(3)牛津大学出版社JNCI杂志运用报道规范后,作者/编辑决策/审稿 人评级之间的关系

◎阴性结果:是否用报告规范 与 稿件最终是否采用以及审稿评价 之间无关联性。

(4)用ARRIVE评价动物体内研究依从性

一英国爱丁堡大学

◎确定了编辑政策要求提交ARRIVE清单能动物实验撰写时采用规范的依从性。未来还须对ARRIVE规范进一步补充。

AQ:到底运用报告规范与稿件最终决策结果是否有关联,目前尚不一致,有的说有,有的认为没关联。

3.临床试验注册

(1)临床试验注册和最终报告的关联性——加拿大女子医学院

- ◎50%临床试验数据并未发表,有很多在2年后发表
- ◎即便发表了,收集的数据也只报道了有限的部分。尤其是次级终点指标,安全性分析细节;注册中的结果与报道结果很可能不一致(21%)
- ◎患者和医生通常会基于潜在可获取临床数据的一小部分来制定治疗决策
- ◎需要方法来提高研究出版和报告水平
- ◎需要编辑,资金资助者,伦理委员会,管理者都来要求临床试验必须注册,且必须公开所有的Protocol

3.临床试验注册

(1)临床试验注册和最终报告的关联性——加拿大女子医学院

- ◎50%临床试验数据并未发表,有很多在2年后发表
- ◎即便发表了,收集的数据也只报道了有限的部分。尤其是次级终点指标,安全性分析细节;注册中的结果与报道结果很可能不一致(21%)
- ◎患者和医生通常会基于潜在可获取临床数据的一小部分来制定治疗决策
- ◎需要方法来提高研究出版和报告水平
- ◎需要编辑,资金资助者,伦理委员会,管理者都来要求临床试验必须注册,且必须公开所有的Protocol



1.同行评议方式的创新

(1)生理学研究所杂志评价作者对双盲审稿的需求

◎ 初步结果显示,作者对双盲审稿还是很有需求的。但作者投稿时文稿中隐去作者信息的习惯还未养成。

(2)公开审稿在各国及各学科之间的应用:分析Publon中展示的审稿和 编辑政策

◎在Publon平台中,公开审稿正在以各种不同的形式发生,但仍只属于小部分。

期刊中有3.5%允许审稿人签名公开

在各国家分布也不同,像在阿拉伯有14%,其他很多国家为0

学科也有不同,工程类的0.8%,人类科学有2.9%

◎数据仅出自Publon,不足以推广到其他场景

1.同行评议方式的创新

(3)比较不同审稿模式下,邀请同行评议接受率

—BioMed Center

结果:

◎调查498期刊,总的接受率是49%,双盲接受率最高,公开审稿接受率最低

双盲 40个刊实施,接受率60%;

单盲 388个刊实施,接受率53%;

公开审稿 70个刊实施,接受率42%

◎对研究领域来说,

临床医学,生命科学:接受双盲审稿的最多,接受公开审稿的最少

在不同医学领域,接受公开审稿率也显著不同:临床医学最多41%,生命科学最少31%

建议:

- ◎如果期刊实施单盲或公开审稿,邀请同行评议的数量要适当增加
- ◎公开审稿模式更适用于临床研究
- ◎ 审稿模式并不是评价期刊的指标。因为期刊固有模式已很多年,那么采取何种形式审稿的差异似乎并不 影响期刊质量。

1.同行评议方式的创新

(4) 《急诊医学Blog》中 一种新的公开审稿模式: Inline format —美国University of Southern California 急诊科

- ◎Inline Peer Review并不延迟阅读时间
- ◎读者阅读文献时,点击查看Inline的审稿意见的点击率高于 查看放在文后的审稿意见
- ◎增强读者/审稿人/作者互动
- ◎增强论文的科学性
- ◎Inline Format是否与增强研究的透明性,及增加文章下载量相关,还须下一步研究

Why do you peer review?

To ensure quality

To surrowe observe works.

As N=1 as editor, I mad other mends!

QUALITY of JOVENNI + WHAT U

Want to gave asserthing lock Good to my ego!

offigation as a scentist

To learn!

I makes me fal weight for 8 Cillul books probling Evillmen as well. To become A better wester-

Why do you appreciate peer reviewers?

They advance science!

They are volunteers! Commitment to quality

Because I can't possibly know everything'

VOLUMATERIA WOOD
Help impose contributions submissions

Science dus not duelped a beter approchange to the PR to annue assent quity, digite all the district. Becaus they are forced to read an open of the people's,

What is the future of peer review?

Persist

0PEN 101

MORE DIFFICULT TO FIND REVIEWEES

Mere useful reviews

Open 1

pen! :-)

Abolishment of Poer raise

OXFORD UNIVERSITY PRESS 公开审稿 呼声很高

2.编辑政策和同行评议过程的创新

(1)改变编辑政策可减少期刊发表偏倚 —NPG

(3)同行评议过程中使用ORCID,减少对审稿人的错误识别—ORCID

(4)患者参与 —the BMJ





RESEARCH NEWS

A third of clinical trials publishing interim findings fail to report final results, review finds

Susan Mayor

London

More than one third of clinical trials publishing interim findings fail to report final results and one in five change their conclusions in final publications, a review of randomised trials has found.¹

Researchers searched PubMed for all randomised trials published between 1 January 2006 and 31 December 2015 reporting interim results, excluding pilot studies, protocols, and cancer trials reporting final primary results but interim secondary outcomes. They then searched for final publications and compared the findings, or contacted the authors if no final publication could be found.

Results, reported in JAMA, showed that final results were published for only 61% of the 171 interim publications identified, after allowing sufficient time to have elapsed beyond study completion.

"Frequent non-publication of final results means true treatment effects often remain unknown," said the researchers, led by Lisa Schwartz from the Dartmouth Institute for Health Policy and Clinical Practice, New Hampshire, US.

Final results differed from interim findings in 21% (15) of the 73 matched pairs of publications reporting the same efficacy or safety outcome.

Six of these studies changed from reporting beneficial interim

and funders should commit to making final results accessible by linking interim publications to final reports.

A second study, also published in JAMA, showed that more than one in five (22.9%, 95% confidence interval [CI] 20.3% to 25.6%) of a random sample of 1002 articles from 269 journals included a positive conflict of interest disclosure.² Nearly one in six (13.6%, 95% CI 11.5% to 15.9%) had no disclosure statement.

There was no indication that disclosures were associated with bias. But the highest prevalence of positive disclosure occurred in articles most likely to contain subjective opinions, such as commentaries, editorials, or narrative reviews.

Studies where authors had positive disclosures were also more likely to have been published in journals with higher impact factors and to have received more media attention, potentially gaining greater interest from the research community and the public.

"This may distort the perception of how common conflicts of interest are and amplify the effect of their results," said the study authors, led by Quinn Grundy from the University of Sydney, Australia.

They recommended that all research journals should enforce requirements that manuscripts include both disclosures and



谢谢!