

主编 杨克虎

HOW TO WRITE A SYSTEMATIC REVIEW

系统评价

指导手册



人民卫生出版社

系统评价 指导手册

主 编 杨克虎

编 委 (按姓氏笔画排序)

马 彬(兰州大学)

王梦书(四川大学)

艾昌林(四川大学)

田金徽(兰州大学)

刘雅莉(兰州大学)

李 晓(四川大学)

李 琳(四川大学)

……书(成都医学院)

虎(兰州大学)

龙(兰州大学)

柏建岭(南京医科大学)

戴俊程(南京医科大学)

人民卫生出版社

图书在版编目 (CIP) 数据

系统评价指导手册/杨克虎主编. —北京:
人民卫生出版社, 2010. 1

ISBN 978-7-117-12439-3

I. 系… II. 杨… III. 临床医学—系统评价—手册
IV. R4-62

中国版本图书馆 CIP 数据核字 (2009) 第 218915 号

门户网: www.pmph.com	出版物查询、网上书店
卫人网: www.ipmph.com	护士、医师、药师、中医师、卫生资格考试培训

系统评价指导手册

主 编: 杨克虎

出版发行: 人民卫生出版社(中继线 010-67616688)

地 址: 北京市丰台区方庄芳群园 3 区 3 号楼

邮 编: 100078

E - mail: pmph@pmph.com

购书热线: 010-67605754 010-65264830

印 刷: 北京人卫印刷厂

经 销: 新华书店

开 本: 850×1168 1/32 印张: 11.125

字 数: 279 千字

版 次: 2010 年 1 月第 1 版 2010 年 1 月第 1 版第 1 次印刷

标准书号: ISBN 978-7-117-12439-3/R · 12440

定 价: 26.00 元

版权所有, 侵权必究, 打击盗版举报电话: 010-87613394

(凡属印装质量问题请与本社销售部联系退换)



序

系统评价作为一种临床证据研究方法,具有一些与临床试验和循证医学实践不同的特殊性。一是国际性。其所以被称为“系统评价”,就是要求“全”,必须纳入和分析评估全球在某一领域的临床研究,如果只纳入某一语种文献,可能因发表偏倚和选择性偏倚而影响证据的准确性和强度。因此,要求研究者具有很好的英文读、写能力。另外,系统评价证据是最高级别证据,其交流传播都是国际性的。因此,系统评价研究可称为国际性的公众事件,研究者必须具有高度的社会责任感。二是研究者知识的全面性。不仅要求研究者熟悉本专业的基本理论和临床知识,还要求通晓临床科研设计知识,熟知临床试验实施过程和诸如统计学、文献检索等相关知识和技能。制作系统评价,要求研究者花大量时间搜集全球相关研究,通过鉴定临床试验的偏倚风险来评估证据的质量,这样获得的证据才不致误导证据使用者。

Cochrane 系统评价研究者均需经过严格的培训。中国 Cochrane 中心成立 12 年来,先后有近 900 人接受过为期三个月的专业培训,约 700 人接受过为期三天的短期培训。然而,在这些接受过培训的群体中,并不是所有的人都能够生产高质量 Cochrane 系统评价。这种情况从一个侧面说明了 Cochrane 系统评价研究的难度以及对质量和对研究者素质的严格要求。所以,我主张系统评价研究者必须直接阅读英文原版的指南,如

Systematic Reviews in Health Care: Meta-analysis in context, 尤其是《Cochrane 系统评价手册》, 以及《Meta 分析系统评价报告规范》(PRISMA)。通过阅读这些原版专著, 可迅速和大大提高研究者对英文文献的阅读、理解和写作能力, 这是对系统评价研究者素质的起码要求。然而, 一开始就直接阅读英文原版指南对中国研究者来说毕竟有一定困难, 这也可能是造成有的研究者将写英文系统评价视为畏途的原因之一。

看到杨克虎教授组织编写的《系统评价指导手册》不禁使人眼前一亮。本书内容翔实, 除了对系统评价/Meta 分析的基本原理做了很好的介绍和阐释外, 还跟进国际最新的方法学进展, 提供了大量最新的方法学信息资源, 除了给有志学习制作系统评价的读者以方法学指导外, 还可使读者循着这些资源进行深入的学习。这本书可帮助有志学习系统评价的读者加深理解系统评价/Meta 分析的原理和方法, 并与英文指南互相印证, 更快地掌握系统评价方法。

制作系统评价的要求如此之高, 反过来, 其制作过程也可成为一种提高研究者自身素质和技能的很好途径。在系统评价研究中学习临床研究, 通过制作系统评价, 可促进研究者熟悉所研究专业领域的全面知识和发展情况, 系统学习和掌握临床科研知识和循证医学批判性思维及技能, 促进临床研究质量的提高, 从而造就一批高质量的临床研究人才。Cochrane 协作网通过生产系统评价产生了一批国际知名的方法学家, 本书的编辑出版也说明了我国的年轻方法学家队伍正在迅速崛起, 这尤其令人欣慰。

中国临床试验注册中心

中国循证医学中心

Chinese Cochrane Centre

INCLIN 地区资源与培训中心

四川大学华西医院



前 言

随着循证医学在中国的引入和传播,系统评价这一综合研究、积累知识的方法引起我国越来越多科研工作者和临床医师的关注,尤其是1999年中国Cochrane中心成立后,系统评价在我国得以快速发展。仅以Cochrane系统评价为例,截至2009年10月,中国内地及香港地区在Cochrane图书馆累计发表系统评价全文和计划书275篇。我国在Cochrane协作网所有发表系统评价注册题目的国家中位列第二,发表计划书的国家中位列第五,发表全文的国家中位列第七。目前,PubMed数据库收录系统评价超过2万余篇,CBM数据库中系统评价的文献量也已达4000余篇。

尽管我国系统评价数量持续增长,质量也在不断提高,但目前国内尚无一本完整、系统、实用的全面介绍系统评价写作的图书。《Cochrane系统评价手册》因其权威性和专业性受到系统评价撰写者的广泛欢迎,但由于其为英文版,且只针对如何撰写Cochrane系统评价,故出版一本适合我国科研和临床人员阅读、使用的中文指导手册势在必行。出于此目的,由兰州大学循证医学中心牵头,联合四川大学、南京医科大学等单位组成了编委会,历时一年,精心撰写、仔细推敲、反复修改、不断完善,终于奉献给大家一本通俗易懂、简明实用的指导书。

本书有以下几个特点:第一,为便于读者阅读和理解,全书统一用“系统评价”这一术语,而非“系统综述”、“系统评估”,以

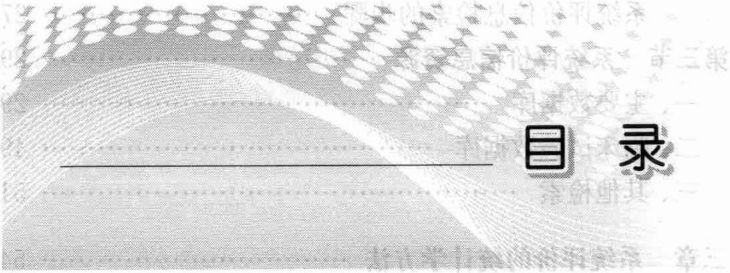
体现系统评价的研究性质与传统综述的区别;第二,较为详尽地阐述了系统评价的起源、发展及特点,澄清了对系统评价的常见误解和误区;第三,详细介绍了系统评价相关软件知识,并通过实际举例指导读者掌握和应用这些软件;第四,在介绍各类系统评价撰写时,尽量选取有代表性的实例进行讲解,以方便读者的理解掌握和灵活运用;第五,专门介绍了如何阅读和利用系统评价的内容,以体现学以致用宗旨;第六,全面归纳和总结了系统评价相关名词术语,便于读者使用时查阅;第七,列出了 SCI 发表系统评价的主要杂志名称,方便读者投稿和发表;第八,成稿后多次征求不同层次使用者的意见和建议,反复修改,不断完善,从而更加贴近实际,贴近读者;第九,每章后均列出推荐阅读材料及详尽的参考文献,便于读者进一步深入学习和研究。本书可供系统评价制作者参考,同时,可作为高等学校本科生、研究生循证医学课的辅助教材。

作为国内第一本全面介绍系统评价写作的书,尽管编者经过长期酝酿和认真准备,但由于我们对系统评价的认识和理解的局限性,以及系统评价方法学研究的不断进步,书中必定存在不足、缺陷甚至错误,读者在阅读时一定要用批判的眼光和辩证的态度,有选择地采纳和吸收。我们热切欢迎各位同道及广大读者对书中的观点和文字提出宝贵意见(通信邮箱:Lzu-xtpj@163.com),以便再版时能够及时更新、改进,止于至善。

最后,我们要特别感谢中国循证医学中心/中国 Cochrane 中心李幼平、吴泰相、刘关键、李静、张鸣明、陈静、王莉、卫茂玲,四川大学华西临床医学院刘鸣、董碧蓉等各位老师长期以来对我们在学习、研究、实践循证医学过程中给予的真诚指导和无私帮助,感谢兰州大学循证医学中心张鹏、贾文琴、移康、郑波波等同学在成稿过程中提出的宝贵建议和对全书进行的校对工作。

杨克虎

2009 年 12 月



目 录

第一章 绪论	1
第一节 历史与现状	2
一、系统评价的历史	2
二、系统评价的现状	4
第二节 概念与特点	5
一、系统评价的概念	5
二、系统评价的特点	6
第三节 系统评价的方法与步骤	9
一、研究综合的方法与步骤	9
二、系统评价的方法与步骤	10
三、系统评价的报告	11
第四节 未来与发展	15
一、消除对系统评价的偏见	15
二、控制系统评价的偏倚	18
三、系统评价的未来与发展	19
第二章 系统评价信息检索	24
第一节 概述	24
第二节 系统评价信息检索基础	25
一、信息检索技术	25

二、系统评价信息检索的步骤	27
第三节 系统评价信息资源	29
一、主要数据库	29
二、临床试验数据库	49
三、其他检索	51
第三章 系统评价的统计学方法	54
第一节 效应量的确定	54
一、二分类资料的效应量及其 95%可信区间	54
二、连续型资料的效应量及其 95%可信区间	55
三、其他效应量	56
第二节 Meta 分析中的异质性	57
一、Meta 分析中异质性的类型及产生原因	57
二、异质性检验方法及其步骤	57
三、异质性资料的处理方法	58
第三节 合并效应量的统计推断	59
一、固定效应模型	60
二、随机效应模型	67
第四节 发表性偏倚	71
一、发表性偏倚的产生来源	71
二、发表性偏倚的识别方法	72
第四章 系统评价常用软件	76
第一节 Review Manager	76
一、数据的建立	77
二、数据的分析	83
三、分析结果的输出	88
第二节 Stata 在 Meta 分析中的应用	89
一、前言	89

二、Stata 界面功能简介	90
三、Meta 分析中常用的命令	92
四、实例分析	95
第三节 文献管理软件	102
一、为什么需要文献管理软件	102
二、什么是文献管理软件	103
三、如何把文献从数据库导出	105
四、如何把文献导入文献管理软件	116
五、如何查重	124
六、如何筛查文献	128
七、小结	132
第五章 干预性试验的系统评价	134
第一节 概述	134
一、基本概念	134
二、常用临床研究设计	135
第二节 提出问题,制订研究计划	139
一、提出临床问题	140
二、制订研究计划	142
第三节 资料检索	145
第四节 筛选文献和提取资料	149
一、筛选文献	149
二、提取资料	150
第五节 纳入研究的方法学质量评价	152
一、随机对照试验中偏倚的来源	152
二、各种偏倚及其预防措施	153
三、盲法和隐蔽分组的区别	156
四、方法学质量评价标准	156

第六节 分析资料 and 解释结果	158
一、分析资料	158
二、解释结果	161
三、讨论和结论	163
第七节 应用举例	164
第六章 诊断性试验的系统评价	179
第一节 概述	179
一、诊断性试验的研究设计	180
二、诊断性试验的评价指标	181
第二节 提出问题, 制订研究计划	184
一、题目	184
二、研究背景	185
三、研究目的	186
四、纳入和排除标准	186
第三节 检索资料	187
第四节 文献筛选和资料提取	189
一、文献筛选	189
二、资料提取	190
第五节 纳入研究的方法学质量评价	191
第六节 分析资料 and 解释结果	196
一、分析资料	196
二、解释结果	198
三、讨论和结论	203
第七节 应用举例	205
第七章 其他研究的系统评价	214
第一节 病因研究的系统评价	214
一、概述	214

二、基本步骤	215
三、应用举例	217
第二节 预后研究的系统评价	221
一、概述	221
二、基本步骤	222
三、应用举例	228
第三节 观察性试验的系统评价	232
一、概述	232
二、基本步骤	233
第四节 动物实验的系统评价	237
一、概述	237
二、基本步骤	239
三、应用举例	242
第五节 群随机的系统评价	246
一、群随机试验简介	246
二、群随机的系统评价	247
三、群随机系统评价的问题与展望	255
四、应用举例	255
第八章 如何阅读系统评价	265
第一节 为何阅读系统评价	265
第二节 如何阅读系统评价	266
一、文献阅读方法的沿革	266
二、阅读系统评价时读者应该提出的 10 个问题	268
三、从阅读到利用	278
第九章 Cochrane 协作网与 Cochrane 系统评价	287
第一节 Cochrane 协作网	287
一、起源	287

二、目的、任务和原则	289
三、组织结构	290
四、挑战	291
第二节 Cochrane 系统评价	293
一、概念与影响	293
二、现状与问题	295
三、注册与撰写	296
四、发表与更新	299
附录一 Cochrane 中心一览表	306
附录二 Cochrane 协作网系统评价小组一览表	310
附录三 Cochrane 协作网研究领域/网络一览表	318
附录四 Cochrane 协作网方法学组一览表	321
附录五 SCI 收录系统评价的研究领域、来源国家及 机构排名	324
附录六 SCI 收录系统评价的主要期刊列表 (以刊名字顺为序)	325
附录七 系统评价常见术语	329

第一章

绪 论

如果像通常所认为的,科学除了辛苦的堆积事实之外什么也不是,那么它将很快在自身重负下停滞、崩塌。一种新思想的提出或一种规律的发现大大减轻了记忆的负担,并且通过引入统一的规则,使记忆以一种可获得的形式保存变得容易。因此,新事物的接受和旧事物的理解、同化两个过程在工作中联系紧密。因为都很重要,我们可以分别讨论其相对的重要性。然而,应该明确一点,最应该得到但恐怕并不总是得到的是那些既有发现又有解释,不仅介绍新的事实,而且指明了其与旧事实联系的工作^[1]。

上述一段文字是 1885 年剑桥大学物理学 Rayleigh 教授在英国科学促进协会大会上的讲话,他洞悉了科学研究的本质并提出了可以更好地传承和创造知识的伟大理念。然而此后在长达一个世纪中,很少有科学组织、研究人员和其他团体重视和实践他的远见卓识。科学应该是逐渐积累而成的,但科学家们只有很少的时候在科学地积累证据。他们很大程度只关心自己当前的研究,绝少意识到需要将有关某一科学问题此前所有的研究进行系统全面地归纳总结^[2]。

卫生服务的质量关系到每个人的健康,卫生保健决策应该基于当前可得的最佳证据,然而教科书上的知识往往过时,最新文献中报道的研究又存在片面和局限,想要获取最佳的证据,就

需要恰当地重复研究和对这些研究的全面总结。系统评价作为一种科学、合理、高效、实用的研究综合方法,在短短 20 年间迅速被卫生保健及其相关领域所接受。本章将简要介绍这种方法的沿革、概念、原理、特点、步骤及发展趋势。

第一节 历史与现状

.....

一、系统评价的历史

(一) 从研究综合到系统评价

虽然认识到需要综合研究证据的历史由来已久,但直到最近 30 年,才逐渐形成这种研究的明确方法。20 世纪 80 年代,社会学家们注意到,在进行研究的综合时,由于缺乏规范的过程,在纳入研究、分析资料及得出结论方面主要靠研究者的主观臆断,而非客观透明的方法,因而针对同一问题得出的结论大相径庭。他们开始意识到,正如原始研究需要严格的方法学指导,研究的综述质量也需要严格的方法学来保证,因为综述本身就是一种科学研究^[3]。1983 年,Light 等出版的《综述研究评价年鉴》一书,包含了 15 篇解决研究综合方法学问题及步骤的文稿,还有 20 篇阐述如何在实践中应用方法学的独立文章^[4]。1984 年,Light 等出版的《总结:综述研究的科学》一书系统全面地阐述了综述研究这门科学的方法和原理,启发了医学研究人员对医学综述质量的关注^[5]。

1985~1986 年间美国 Mulrow^[6] 调查了 4 种发行量超过 50 000 册的医学期刊 JAMA、NEMJ、Ann Intern Med、Arch Intern Med 上发表的 50 篇综述,发现普遍质量低下,基本没有使用科学的方法去甄别、评价和综合信息,故也无法充分为读者提供有价值的信息。Mulrow 提出:医学综述应该致力于解决一个具体明确的问题;检索应有效率;应该制定详细的纳入排除

标准;评价方法和过程应标准化;结果的整合应客观全面;只有经过系统全面的收集、评价和整合信息,最后的结论才可信;评价者应指出当前综述的局限性并提出以后的改进建议。这一发现立刻引起医学界的关注,并为系统评价奠定了方法学基础。1988年,Oxman 和 Guyatt 开始发表系列文章指导读者如何阅读、评价医学综述质量^[7]。北美医学研究者们的这些动向正在呼唤一种全新研究综合方法的出现。与此同时,欧洲一些临床医师也在不断关注和探索如何更科学系统地整理和收集研究证据的方法。

1972年,英国医师和公共卫生专家 Archie Cochrane 的《效果与效率:卫生服务中的随想》一书问世^[8]。书中明确提出:由于资源终将有限,因此应该使用已被证明有明显效果的卫生保健措施,而随机对照试验是检验干预效果最好的方法。1979年,他进一步提出应该将医学领域里所有相关的随机对照试验收集起来综合分析,并随着新的临床试验的出现不断更新,以便得出更为可靠的结论^[9]。

英国产科医师 Iain Chalmers 深受 Cochrane 的影响,并将其设想付诸实践。1989年他与同行出版的《妊娠和分娩领域的有效治疗》中对短疗程、低价格类固醇药物治疗有早产倾向孕妇的 RCT 进行总结归纳,结果有力证明了这种药物可大大降低婴儿死于早产并发症的风险。该结果在欧洲的推广,减少了欧洲新生儿死亡率的 30%~50%^[10],现代意义上的 Cochrane 系统评价雏形在这本著作中已经形成。

(二) 从 Meta 分析到系统评价

19 世纪上半叶,为减少随机误差的影响,德国数学家 Karl Gauss 和法国数学家 Pierre-Simon Laplace 发明了一系列合并效应量的统计方法,形成了 Meta 分析的雏形。Meta 分析从理论走向实践最早是在天文观测中。当时的天文学家发现,在多个场合测量恒星的位置往往导致稍有不同的估计,需要一定的

方法来合并估计,从收集的结果中得到一个平均值。1861年,英国皇家天文学家 George Airy 出版了一本针对天文学家的“教材”,在书中阐述了这种定量合成过程所采用的方法^[2],即 Meta 分析。

1904年,Karl Pearson 发表在《英国医学杂志》上的一篇文章被认为是 Meta 分析早期在医学中的应用。时任伦敦大学学院生物测定学实验室主任的 Karl Pearson 应政府要求,评价一种伤寒疫苗有效性的证据。他收集了在英国不同地区服役士兵中进行的 11 个相关研究的免疫力和病死率数据,计算其中每一个的相关系数,并将这些相关系数合为两组,得出平均相关系数,使得结果更加客观全面^[11]。

20 世纪 30 年代,英国农业统计学家 Ronald Fisher 提出了合并 P 值的概念,在其领域发展并运用了类似于 Meta 分析的方法。1976 年,心理学家 Gene Glass 正式提出 Meta 分析的概念,并迅速在卫生保健领域得以传播和应用^[12]。此后,医学研究者将这种统计方法逐步应用到临床试验及卫生领域,并在应用的过程中考虑进一步减小系统误差,即向现在意义上的系统评价靠拢。

20 世纪 90 年代,为降低偏倚而提出的科学综述与为减小机遇而逐渐成熟的 Meta 分析在医学领域的最终结合,催生出了—种全新的生产证据的方法,1993 年 7 月,《英国医学杂志》与英国 Cochrane 中心的方法学家和编辑们在伦敦召开会议,将这种方法定名为“Systematic Reviews”,并大力推广和使用这一新术语,由此揭开了系统评价全新的一页。

二、系统评价的现状

从以上可以看出,系统评价并非横空出世的产物,而是研究综合和 Meta 分析发展的必然结果。现在系统评价的理念和方法已被广泛接受。2007 年,Moher 等^[13]对系统评价的现状做了一次全面调查,结果显示每年新发表英文系统评价约 2500 个,