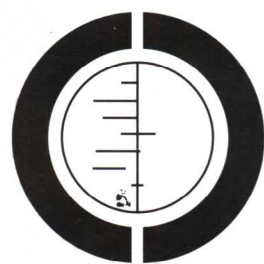


围产医学

最佳证据



首都医科大学附属北京妇产医院
中国循证医学中心

古梅 黄醒华 李坚 主 编
李幼平 张鸣明 总审校



浙江大学出版社

W E I C H A N Y I X U E Z U I J I A Z H E N G J U

围产医学最佳证据

首都医科大学附属北京妇产医院 古梅 黄醒华 李坚 主 编
中 国 循 证 医 学 中 心 李幼平 张鸣明 总审校

浙 江 大 学 出 版 社

图书在版编目(CIP)数据

围产医学最佳证据 / 古梅, 黄醒华, 李坚主编. — 杭州: 浙江大学出版社, 2004. 6
ISBN 7-308-03708-8

I. 围... II. ①古... ②黄... ③李... III. 围产期—产科学 IV. R714.7

中国版本图书馆 CIP 数据核字(2004)第 050855 号

出版发行 浙江大学出版社
(杭州浙大路 38 号 邮政编码 310027)
(E-mail zupress@mail.hz.zj.cn)
(网址 <http://www.zjupress.com>)

策划组稿 王锴
责任编辑 王锴
排版 浙江大学出版社电脑排版中心
印刷 浙江大学印刷厂
开本 787mm×1092mm 1/16
印张 11.75
字数 300 千字
版印次 2004 年 6 月第 1 版 2004 年 6 月第 1 次印刷
印数 0001—2000
书号 ISBN 7-308-03708-8/R·142
定价 35.00 元

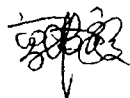
序

循证医学是 21 世纪临床医学发展的方向,普及和宣传循证医学的知识,介绍国际 Cochrane 协作网的系统评价,是临床医疗工作的需要,是科研工作的需要,也是卫生决策的需要。

既往的临床实践是以经验为指导的,但大多数临床实践未经过大样本设计良好的试验研究证实。而循证医学则是根据某一临床问题全面收集、选择设计合理、资料统计科学的大样本的随机对照试验(Randomized controlled trial, RCT)进行系统评价,对医疗保健的实践进行客观分析,同时也指出了在哪些问题上还需要进一步开展多中心的 RCT 研究,并按其研究质量参考价值进行分级,因此对临床医师的实践有指导意义。

由于 Cochrane 图书馆的系统评价是全球循证医学最全面、最可靠的临床研究证据,为使我国更多的妇产科医务工作者和妇幼保健人员能够及时阅读这些资料,首都医科大学附属北京妇产医院将 Cochrane 图书馆中妊娠和分娩组的系统评价资料 215 篇作为第二部分译成中文介绍给大家,用以指导临床实践和临床科研工作,避免无效劳动和不符合要求的科学研究,推动循证医学并正确利用和合理分析临床资料来总结临床经验与制定妇幼保健医疗卫生政策,规范医疗服务行为,提高医疗保健干预措施的效率;以人为本,运用有效、安全、廉价的医疗措施,提高医疗质量,满足患者需求,减少医疗纠纷,加快促进妇幼保健学科的发展。该书对妇幼保健工作是有指导和实用价值的。

中华医学会会长 曹泽毅



2004 年元月

前言

生命与健康是医学研究的永恒主题,妇女与儿童是医学研究的优选主题,而孕育和哺育新生命的围产医学,则是医学巨著中最重要的篇章。处于围产期的母、子是人类最娇弱的群体,对于他们,诊断、治疗的任何一点失误,都可能造成不可挽回的损失。

作为一名儿科医师,Iain Chalmers 爵士就是因为对亲身经历的 2 个围产期病例不幸夭折的深刻反思和责任驱使,转而研究临床问题,探寻最好的证据,于 20 世纪 80 年代末,与国际协作共同制作妊娠与分娩及新生儿的系统评价,用基于 7 个随机对照试验的系统评价结果,肯定了短程价廉的皮质类固醇治疗先兆早产有效,能降低早产儿死亡率。该结果一经推广,减少了欧洲早产儿死亡率的 30%~50%。受此结果的启发和鼓舞,Iain Chalmers 爵士及同事于 1992 年创建了英国 Cochrane 中心,次年创建国际 Cochrane 协作网,用 10 年时间,致力于组织全球医务工作者共同生产和传播高质量临床研究证据,以提高医疗服务质量。在他的领导下,妇产科和儿科领域成为协作网中生产证据最多的评价小组之一,在世界上享有盛誉。

1997 年 8 月,卫生部批准在原华西医科大学附属第一医院建立中国循证医学中心;1999 年 3 月,该中心成功注册成为国际 Cochrane 协作网在亚洲惟一的 Chinese Cochrane Center (ChiCC)。作为中国循证医学的拓荒者和组织者,我们率先引进循证医学,创建新兴学科方向,开展循证研究,促进循证实践,力图以多渠道、全方位的传播方式,孵育循证医学文化,唤醒公众意识,传播循证医学知识,为卫生决策者、临床医生、病人、公众及其他卫生保健用户提供循证决策的可靠科学依据,推动循证医学在中国的发展。

为此,中国循证医学中心与 Cochrane 协作网更新软件公司(Update Sonware)经过近 2 年的谈判,最终于 2000 年 6 月达成协议,中国 Cochrane 中心作为国际 Cochrane 协作网的正式成员,被授权为翻译、出版 Cochrane 协作网产品及其他 Cochrane 信息的惟一合法中译机构。此后,我们翻译出版了由更新软件公司定期提供的《最新临床证据——Cochrane 系统评价摘要精萃》英文版两期,并定期选译 Cochrane 图书馆系统评价,摘要在《中国循证医学杂志》连载。2003 年 10 月,我们又与 Cochrane 协作网 Wiley and Sons 责任有限公司及 BMJ 出版发行集团谈判签约,将继续翻译出版 Cochrane 协作网产品,开拓新的领域——临床证据(Clinical evidence)的翻译出版。我们将陆续为各专业的医务工作者推出符合中国国情,采用当前最佳证据的科普读物和专业循证手册,让最好的证据真正能为政府采纳、医生接受、公众喜欢。

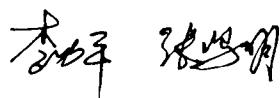
《围产医学最佳证据》是我们 2004 年推出的第一本书,主编古梅曾于 2002 年 4 月到成都参加中心主办的第二届亚太地区循证医学研讨会。她以临床医学专家的敏锐,意识到使用循证证据对提高医疗质量的作用和价值,并在首都医科大学附属北京妇产医院领导和同事

的支持下,很快完成了本书。她的努力得到了中国循证医学中心的肯定和支持,并出面为她解决了外方版权问题。

本书简明扼要地介绍了循证医学的基本概念和理论、实践方法和步骤,重点介绍了围产医学的临床证据,从临床实际出发,有机编排了近5年围产医学的200余个最佳临床证据,是一本针对性强,时效性和实用性俱佳的循证证据手册,能帮助临床医生在最短时间内找到当前该领域最新的临床研究进展,促进围产领域的循证实践。

如果说农业科学家的文章写在大地上,那么临床医学家的文章就应该写在临床上,健康、生命、质量、成本效益是检验我们成绩的标准。正如循证医学的伟大先驱 David Sakette 教授所言:“目前最重大的挑战之一,就是为第一线的临床医生提供高质量证据,并帮助他们采纳证据和实施循证医学。”我们将为此不断努力奋斗,做好服务。我们更期待有更多有志于此的中国同仁、国外同仁加入我们的行列,共同为之努力!

中国循证医学中心
Chinese Cochrane Center



2004年元月于成都

目 录

序

前言

第一部分 循证医学与围产医学

一、循证医学的提出	(1)
二、循证医学的研究方法——临床应用研究的方向	(2)
1. 临床流行病学(clinical epidemiology)	(2)
1.1 临床科学研究方案的设计原则	(3)
1.2 测量	(4)
1.3 评价	(4)
1.4 常用的临床科研设计方案	(4)
2. 循证医学的研究方法	(5)
2.1 系统评价(systematic review, SR)	(5)
2.2 荟萃分析(Meta-analysis)	(7)
2.3 系统评价与叙述性综述的区别与联系	(9)
三、循证医学和传统医学的区别	(10)
四、围产医学在循证医学中的发展	(11)
五、Cochrane 图书馆	(12)
1. 通过因特网检索 Cochrane 图书馆(Cochrane Library, CL)	(12)
1.1 中国循证医学/Cochrane 中心网站	(12)
1.2 Update Software 公司的网站	(12)
1.3 Cochrane 图书馆	(12)
2. 世界卫生组织(WHO)生殖健康图书馆电子版	(17)
六、把握好证据,实践循证医学	(20)

第二部分 围产医学最佳证据 (Cochrane 妊娠和分娩组)

一、产前保健模式	(23)
二、产前教育项目	(26)
三、孕期营养补充	(29)
四、孕期干预措施	(37)
五、妊娠期合并症	(47)
六、妊高征	(65)
七、多胎妊娠	(74)

八、胎儿生长受限.....	(78)
九、胎儿窘迫.....	(83)
十、臀先露.....	(86)
十一、早产.....	(90)
十二、胎膜早破.....	(98)
十三、产前诊断	(103)
十四、产前检查	(109)
十五、胎儿监测	(110)
十六、促宫颈成熟和引产	(116)
十七、剖宫产	(136)
十八、分娩期处理	(143)
十九、分娩镇痛方法	(160)
二十、产后保健和母乳喂养	(163)
二十一、社会支持	(175)

第一部分 循证医学与围产医学

古梅 李坚

循证医学(evidence based medicine, EBM),意为“遵循科学依据的医学”。它起源于 20 世纪中叶,是伴随着全球医学界探讨如何解决医疗费用过度膨胀,寻求有限资源的充分利用而产生的。实践证明,它是一种适应社会、科技及经济发展的新模式,对医疗实践、新药研制、政府医疗卫生主管部门法规制度制订等有重要的指导和参考价值。例如:丹麦根据 Cochrane 系统评价(Cochrane systematic review, CSR)结果,取消了对孕妇进行常规超声波检查的规定,节约了大量人力、物力和财力;澳大利亚利用 CSR 结果,制定了晚期乳癌的治疗指南;英国利用 CSR 结果,帮助制定了防治骨质疏松的国家指南等。循证医学的发展将提高医疗保健干预措施的效率,合理、高效地利用和分配有限的卫生资源,提高卫生管理部门决策的科学性。

一、循证医学的提出

近半个世纪以来,随着人类社会的进步和人们生活水平的提高,一些曾经严重威胁人类生命的疾病,如传染性疾病、营养缺乏性疾病已基本得到有效控制,人类的疾病谱发生了改变,而像癌症、心脑血管疾病和各种慢性疾病则成了威胁人类健康的主要疾病。由于人体的复杂性和独特性,这些多因素疾病不可能在短时间内获得有效的治疗方法。为了使一些药物尽快用于临床,曾有一些临床观察研究缺乏严格的科研设计,导致其研究结论带有一定的偏倚(bias),例如,在 20 世纪五、六十年代,依据来自己公布的 6 个平行或历史对照方法的试验提示,使用乙烯雌酚,对治疗可能会流产的孕妇有效果。但同一时期进行的 5 个随机对照试验中,却没有一个试验证明乙烯雌酚治疗有效,有的甚至说是有害的。然而,尽管有这些不同甚至是矛盾的结果报告,乙烯雌酚却照样被广泛使用,直到 1969 年有临床资料表明,经过乙烯雌酚治疗的妇女,其出生的后代会发生腺癌及其他严重的副作用时,才被制止。可见,在探索人类疾病和健康的问题时,临床试验结果的重要程度(importance)及其适用性(appraisal)问题,已越来越凸显出来。

1956 年,美国进行了第一次大规模、多中心、真正意义上的随机对照临床试验(randomized controlled trail, RCT),研究新生儿最佳的氧气使用量。这种方法将随机误差和偏倚控制到最低限度,因此受到重视而被接受。1972 年,英国著名流行病学家和内科医生 Archi Cochrane 出版了专著《疗效与效益:健康服务中的随机反映》(Effectiveness and Efficiency Random Reflections on Health Services),他在书中明确指出:“由于资源终将有限,因此应该使用已被恰当证明有明显效果的医疗保健措施。”他还特别强调:“应用随机对照试验(RCT)证据之所以重要,是因为它比其他任何证据来源更为可靠。”并首次提出了在医疗服

务中如何才能做到既有疗效又有效益,即卫生保健效果的问题。1979年,Cochrane 首先提出并号召临床工作者系统地回顾科研资料,将所有的随机对照试验资料收集起来进行系统评价(systematic review, SR),并随着新的试验的出现随时更新这些资料以保持其先进性,为临床实践提供可靠依据。之后,随着大量临床研究成果的出现,又产生了如何将这些研究成果转化为临床实践,避免不必要的浪费和重复研究的问题。例如,就一项短程价廉的激素类药物氢化可的松对可能会发生治疗疗效的随机对照试验系统评价结果而言,自1972年第一个试验结果报道到1991年的6项试验结果相继报道,但7项试验结果却不一致。也就是说,该疗法是否利大于弊,根据单个临床试验结果难以确定。然而,通过系统评价分析,可以明确肯定:氢化可的松的确可降低新生儿死于早产并发症的危险,使早产儿死亡率下降30%~50%。遗憾的是,由于多数产科医师并未认识到该项治疗措施的效果,致使许多的早产儿可能因其母亲未接受这种治疗而死亡,同时还耗费了许多不必要的治疗费用。

20世纪80年代初期,国际著名的内科学家,加拿大的麦克马斯特大学(McMaster University)David L. Sackett 教授及其同事在长期的临床流行病学实践的基础上,正式提出一套医学文献评价原则,指导临床医生利用文献研究结果来进行临床实践,并将此称为循证医学(EBM),认为,医务人员应该“谨慎地(conscientious)、明确地(explicit)、明智地(judicious)应用当前最佳证据就如何对患者进行医疗做出决策。”也就是说,在临床医疗实践中,医务人员应尽量以客观的科学依据结果为证据制订病人的诊治方案,结合临床经验对病人进行处理,选择的治疗方法应尽可能接近安全、有效、经济。

1992年,在英国牛津正式成立了以已故 Archie Cochrane 博士姓氏命名的英国 Cochrane 中心;1993年英国牛津正式成立国际 Cochrane 协作网,其目的是制作、保存、传播和更新医学各领域的系统评价(SR),为循证医学实践提供最佳证据。我国于1999年3月在华西医科大学正式成立了中国循证医学中心,并加入了国际 Cochrane 协作网,以推动循证医学在中国的发展。

二、循证医学的研究方法——临床应用研究的方向

循证医学是以解决临床问题、认识和预防疾病、提高临床诊断的准确性、应用有效的治疗措施、改善患者预后并提高人们的生存质量为目的的。具体地说,循证医学是指对临床研究的文献,运用流行病学的原则和方法以及相关的质量评价标准,进行认真分析与评价,以获得最真实可靠且有临床应用价值的研究成果或称之为最新最佳证据(current best evidence)的科学。

循证医学是临床应用研究的方向,临床流行病学的基础理论和临床研究的方法学,是实践循证医学的学科基础。

1. 临床流行病学(clinical epidemiology)

临床流行病学是指在临床医学领域内引入流行病学和卫生统计学的方法,从患者个体的诊治扩大到群体特性的研究,以探讨疾病的病因、发病机理、临床表现、诊治、预防及预后等临床规律,并进行严格的设计、测量和评价的临床基础科学。

临床流行病学这一概念,是1938年美国流行病学家 John Paul 正式提出的,但当时并

未引起临床学家的重视。直到 20 世纪 70 年代后期, David L. Sackett 等人在临床研究和实践中, 创造性地将流行病学及卫生统计学的原理和方法, 有机地与临床医学相结合, 随后又将卫生经济学及社会医学融合于临床医学, 从而丰富和发展了临床研究的方法学, 才使临床流行病学逐渐完善。日益发展和完善的临床流行病学, 以其先进的临床科研方法学强调临床科研设计(design)、测量(measurement)和评价(evaluation, DME)的科学性, 排除各种偏倚、干扰因素的影响, 确保研究成果的真实性(validity)和研究结论的可靠性(reliability), 从而进一步推动了临床科学研究, 深化了对疾病的发生、发展和转归整体规律的宏观认识, 提高了对疾病的诊断和治疗水平。

1.1 临床科学研究方案的设计原则

1.1.1 根据不同性质的临床研究课题, 选择合理的设计方案, 在注重科学性和可行性的前提下, 进行抉择。

依不同性质的研究课题抉择不同的研究方案*

研究的性质	备选方案	论证强度	可行性
病因/危险因素 (causation / risk factor)	— 随机对照试验(randomized controlled trail, RCT)	++++	----
	— 队列研究(cohort study)	+++	+++
	— 病例对照研究(case-control study)	+	+++
	— 描述性研究(descriptive study)*	±	++++
防治性研究 (prevention / treatment)	— 随机对照试验(RCT)	++++	++
	— 交叉试验(cross-over study)	++	++
	— 前后对照试验(before-after study)	+	++
	— 病例对照试验(case-control study)	+	+++
	— 描述性研究(descriptive study)	±	++++
预后研究 (prognosis)	— 队列研究(cohort study)	++	++
	— 病例对照研究(case-control study)	+	+++
	— 描述性研究(descriptive study)	±	++++

* 含横断面调查(cross-sectional study)

1.1.2 研究对象一定要符合临床公认的诊断标准, 按课题设计的具体要求确定对象的纳入标准及排除标准, 确保研究对象的可靠性。

1.1.3 样本量的估算要适度。

根据研究设计的假设条件, 计算合适的样本需要量, 以防样本量不足造成假阴性错误结论, 同时又可避免样本量太大而造成资源和时间的浪费。

1.1.4 试验措施中的创新性或存在的某种特殊性, 要有科学依据, 否则试验没有价值。

1.1.5 临床试验观察的期间要根据试验终点的设计指标而定。

如终点指标为痊愈、死亡、有效、无效等, 那么大多数试验对象达到终点需要的时间即为

试验观察期间,且结果的确定要以生物学及临床试验数据为依据。

1.1.6 由于临床研究的复杂性,其研究过程中的各个环节不可避免地存在着各种偏倚因素的干扰,因此要有识别和防止偏倚因素干扰的对策,确保研究结果的真实可靠。

1.1.7 正确应用统计分析方法,不同质的研究资料需采用不同的统计学方法进行比较和评价。

1.2 测量

1.2.1 试验的措施一定要有反应性和可度量性。

因为试验性的致病因素或治疗药物,其本身存有致病或治疗作用,所以这种作用要能被临床和实验室等检查方法以及指标量度客观地反映出来。

1.2.2 测量的方法要求有良好的敏感性和特异性。

即当致病或治疗效应发生后,采用的测量方法要有良好的敏感性予以发现,并有良好的特异性予以确定,从而防止漏诊和误诊的发生。

1.2.3 测量指标的判断标准和临床意义要明确。

如实验室的生化指标,血压测量等定量指标,病人主观征候的感觉状况如疼痛减轻等定性指标,这些指标的数据及程度必须有临床意义及确定公认的判断标准如有效、无效或恶化等予以量度。死亡、痊愈或病残属于最终效应的硬指标。

1.3 评价

临床研究结果究竟是否真实可靠,有多大的临床价值,是否对疾病的诊治和预防适用,患者是否可以承担其费用等,都是应考虑的重要的评价原则。

1.3.1 对临床意义的评价

要用严格评价(critical appraisal)的标准和方法,分析和评价临床研究内容和结果的真实性、可靠性及临床意义。临床意义所应用的效果指标是指事件发生率(event rate,包括痊愈率、有效率、残废率、病死率、药物不良反应率、发病率和患病率),绝对、相对危险降低率或增高率等的评价。

1.3.2 对研究结果的统计学分析和评价

如果研究结果有临床意义的话,必须用正确的统计学方法对这一结果进行显著性检验,以评价临床差异的真实程度,即肯定结果的真阳性、真阴性的几率以及检验效能的水平。当某种研究结果既有临床意义又有统计学显著差异时,即能作出肯定性的结论;如仅有临床显著意义而无统计学显著差异,也不能因此而否定临床意义,此时应计算Ⅱ型错误和检验效能的水平。

1.3.3 研究结果的卫生经济学评价

临床医学研究的结果应对其社会效益及经济效益进行评价,应用卫生经济学的原则方法计算成本—效果(cost-effectiveness),成本—效益(cost-benefit)以及成本—效用(cost-utility),并作比较和评价,肯定那些成本低、效果佳的研究成果。

1.4 常用的临床科研设计方案

临床科研的类型可以分为两大类:试验性研究和观察性研究。

1.4.1 试验性研究

试验性研究可以人为地控制条件,随机分组,有目的地设置各种对照,直接探讨某个或

某些被研究因素与疾病或事件之间的联系。临床试验常用的方法是随机对照试验(RCT)、前后对照试验(before-after study)与交叉试验(cross-over study),前后对照试验与交叉试验是随机对照试验的两种特例。前后对照试验是从个体着眼,对同一个体使用的两种措施或药物进行前或后的随机分配处理。交叉试验是从研究组群着眼,分试验组与对照组,然后互相对换,最后统一评价其治疗措施的效果。

1.4.2 观察性研究

观察性研究则不能人为地控制试验条件,只能在自然情况下模拟试验性研究,尽量控制非研究因素,以得到真实性结果。临床常用的观察性研究方法有病例对照研究(case-control study)、队列研究(cohort study)、横断面研究(cross-sectional study)和描述性研究(descriptive study)。病例对照研究与队列研究都是设立对照组进行的比较性研究,又统称为分析性研究,其论证强度比横断面研究和描述性研究要高。

上述这些方法广泛地被应用于临床疾病的病因、诊断、治疗和预后的研究。病因、诊断和预后的研究主要使用观察性研究方法,而治疗措施(药物、手术、放疗、化疗、理疗康复和心理治疗等)效果的评价则常用临床试验研究方法,特别是随机对照试验。

传统的临床科研或经验总结,利用的往往是现存的资料,即使是前瞻性研究,由于缺乏严格设计,资料的真实可靠性也会有所缺陷。不解决这一问题的症结,就很难使临床科研产生质的飞跃。临床流行病学这门新兴的临床医学基础科学(basic sciences for clinical medicine),为临床医生从事医学科学研究和指导临床实践提供了有用的理论与方法,为证据的产生奠定了基础。因此,要实践循证医学,就必须掌握临床流行病学的基础理论和临床研究的方法,只有得到足够数量的高质量的随机对照试验才能据此进行高质量的系统评价。由此可见,临床流行病学是基础,循证医学是其发展及应用,反过来又促进随机对照试验等临床研究工作质量的提高。

2. 循证医学的研究方法

循证医学的研究方法是对临床资料进行二次分析评价。

随机对照试验花费大,耗时多,需要一定规模的机构来组织和协调,因此不能满足临床研究实践的需要,这是当前临床研究存在问题的一个方面。另一方面,目前大多数临床研究的试验样本量偏小,故结果不能令人信服,尤其是小样本试验的亚组分析结果,更会造成对临床实践的错误导向。所以,将所有涉及同一问题的各单个 RCT 结果联合起来进行系统评价或荟萃分析,就既能增大样本量,增强检验效能,又可以得出可靠的结论。

2.1 系统评价(systematic review, SR)

系统评价与其他科研一样,也是一种基本的科学研究活动,它是一种将多个临床研究按照预定的方法和标准,根据某一题目(课题)或临床问题,系统、全面地收集、选择、评价和分析相关的文献资料,从而得出综合结论的方法。系统评价的结论不是一成不变的,为保持证据的新颖性、可靠性和科学性,需对该系统评价进行定期更新。系统评价的具体操作步骤如下:

2.1.1 提出临床实践中所要解决的问题

一个好问题应包括四个关键因素:① 何种病人;② 采用何种干预措施;③ 运用何种指标来判断疗效;④ 使用何种研究设计方法。例如:产程中胎儿窘迫(病人)的手术治疗和保守

治疗(干预措施)对母亲和围产儿发病率影响(判效指标)随机对照试验(研究设计)的系统评价。

2.1.2 收集资料,检索有关医学文献

例如,应用计算机检索各种国际、国内光盘或网络版数据库和 Cochrane 系统评价协作组专业资料库;人工查阅有关学术期刊、学术会议论文集;追踪查阅文献或综述的参考文献目录等。为防止“发表偏倚”的问题,还需从临床试验研究者或其他人员那里去获得未发表的有阴性结果倾向的资料,以确保所研究的结果真实可靠。

2.1.3 确定纳入和排除标准

根据提出的问题,确定纳入的标准和排除标准,并对临床试验进行筛选。例如,在产程中胎儿窘迫的手术和保守治疗的系统评价中,纳入标准应定为在产程中怀疑有胎儿宫内窘迫的孕妇,有手术产组(剖宫产或阴道助产)和保守治疗组的比较,真正随机分配的无混杂因素的随机对照试验可入选,非随机对照试验则被排除。

2.1.4 资料提取

按纳入标准确定所选试验后,需要对各试验进行资料的提取,可设计一个资料提取表格用于记录各组的病人数、结果、反映研究质量的指标及其他重要资料。值得注意的问题是:①在发表的文章中,常常缺乏所需要的数据,这时候应与作者联系以补充完善。②由于工作量大,为保证质量,应有两人单独进行资料提取,然后交叉核对,以防止在资料提取过程中出现错误。

2.1.5 对各试验的质量评估

试验质量的不同可以导致结果的不同,越严格的试验其结果越接近真实。以下四个方面的评估是最基本的:①选择性偏倚,即评估是否真正随机分配治疗组和对照组;②实施偏倚,即评估除所要研究的干预措施以外,两组其他处理措施是否一致(无混杂因素的);③排除偏倚,即评估两组受试者退出试验的情况是否有系统的差异,是否有过多的失访病例;④测量偏倚,评估最好由两名以上研究者独立进行盲法评估试验质量,通过讨论解决意见分歧,不足资料可以通过与研究者的本人联系予以补充以防止测量偏倚。

2.1.6 统计学处理

统计学处理是系统评价的重要步骤之一,包括四个方面:①检验各个试验结果是否基本一致(同质性)以证实合并这些试验的合理性,可用 χ^2 检验。如果不一致,应分析原因。如果试验结果不一致是因为某个试验失访病例过多,那么应排除这个试验,或者尽量去获得失访者的信息,或对失访人员采用恰当的统计学方法如意向分析法(intention to treat analysis)来处理,以减少对试验结果影响。②对各试验的统计量进行合并,得出合并后的统计量,如相对危险度(relative risk,RR)或比值比(odds ratio,OR),以及其 95%的可信区间(confidence interval,CI)。③对合并的统计量进行统计检验和统计推断,以确定某一治疗方法是否降低有害事件(死亡率或病死率)的发生或是否有治疗效果。④用 Meta-分析图(森林图 forest plot)来表示出单个试验结果和合并的结果,判断治疗方法是否有效。

2.1.7 敏感性分析(sensitivity analysis)

敏感性分析是指改变某些影响结果的重要因素,如纳入标准的确定、研究质量的差异、失访情况、统计方法(固定效应或随机效应模型)和效应量(effect size)(比值比或相对危险度)等的选择是否敏感,以观察同质性和合成结果是否发生变化,从而判断结果的稳定性和

强度。

2.1.8 发表性偏倚

影响发表性偏倚的主要因素有治疗效应量的统计学意义、样本量的大小、治疗方法有无创新、研究的问题是否重要、研究质量是否受资助,以及发表该报道的杂志的特色等方面。发表性偏倚可使 Meta-分析过分夸大治疗效应量或危险因素的关联强度,导致临床个体治疗与卫生决策的失误。所以在进行 Meta-分析前,必须评估发表性偏倚的可能影响,也可绘制漏斗图(funnel plot)来了解发表偏倚的情况。

2.1.9 评价者结论

评价者结论包括:① 说明根据所纳入试验的系统评价结果是否能够作出对某一疗法有效或无效的结论,是否可以在实践中推广。② 如果现有资料尚不足以作出结论时,那么有什么趋势?提出是否应该进一步临床试验的建议。例如:在人工破膜缩短产程的系统评价中,评价者的结论是常规早期人工破膜各有利弊。早期人工破膜可缩短产程,减少 5 分钟 Apgar 异常评分,但 Meta-分析未显示早期人工破水可减少剖宫产分娩的风险,而事实上它有增加剖宫产的趋势。一项大样本研究提示,早期人工破膜与胎儿窘迫而剖宫产分娩有关,这表明人工破膜应在产程进展异常时施行。

系统评价可以是定性的(qualitative systematic review,定性系统评价),也可以是定量的(quantitative systematic review,定量系统评价),即包含 Meta-分析过程。

2.2 荟萃分析(Meta-analysis)

Meta-分析由 Beecher 于 1955 年最先提出,英国心理学家 Glass 于 1976 年为它命名。Meta-分析的概念分广义和狭义两种:广义的概念是指当系统评价用定量合成的方法对资料进行统计学处理时,即称 Meta-分析;狭义的概念仅指 Meta-分析是一种定量合成的统计处理方法。目前在国外文献中,以广义的概念应用更为普遍。

Meta-分析的统计方法多种多样,应根据资料的性质选择某一统计分析方法:如果计算资料各试验间同质性时,可选择固定效应模型(fixed effect model),即 Peto 方法或 Mantel-Haenzel 方法;各试验间存在异质性时,选用随机效应模型(random effect model),或 DerSimonian-Laird(D-L 法)。Cochrane 系统评价有专门的软件(RevMan)进行上述统计学处理。

Meta-分析是系统评价的一个研究过程,它将多个独立的、可以合成的临床研究综合起来进行定量分析,比较和综合针对同一问题所取得的研究结果,比较和综合其结论是否有意义。在进行 Meta-分析时,应根据资料的类型及评价目的,选择效应量和统计分析方法。例如:① 对于分类变量,可选择比值比(OR)、相对危险度(RR)、危险度差值(risk difference, RD)和防止某一例事件需要治疗同类患者的人数(number needed to treat, NNT)等作为效应量,表示合成结果。② 对于连续变量,当结果测量采用相同度量衡单位时,应选择加权均数差值(weighted mean difference, WMD);当结果测量采用不同度量衡单位如疼痛评分在不同研究中采用不同的量表时,则应选择标化的均数差值(standardized mean difference, SMD)。③ 进行 Meta-分析合成结果时,可选择固定效应模型或随机效应模型,用 Meta-分析图(森林图)来表示其结果。Meta-分析图是以 OR 及其 95% 可信区间(CI)作图的一个简便方法,可用于描述每个研究的结果和其特征,展示各研究结果间的差异情况。图 2-2-1 是“早产前给予皮质类固醇治疗”系统评价的 Meta-分析的比值比(OR)图。

该图中,从左至右分别为比较(Comparison)或判效指标(Outcome)、比值比(OR)、加权

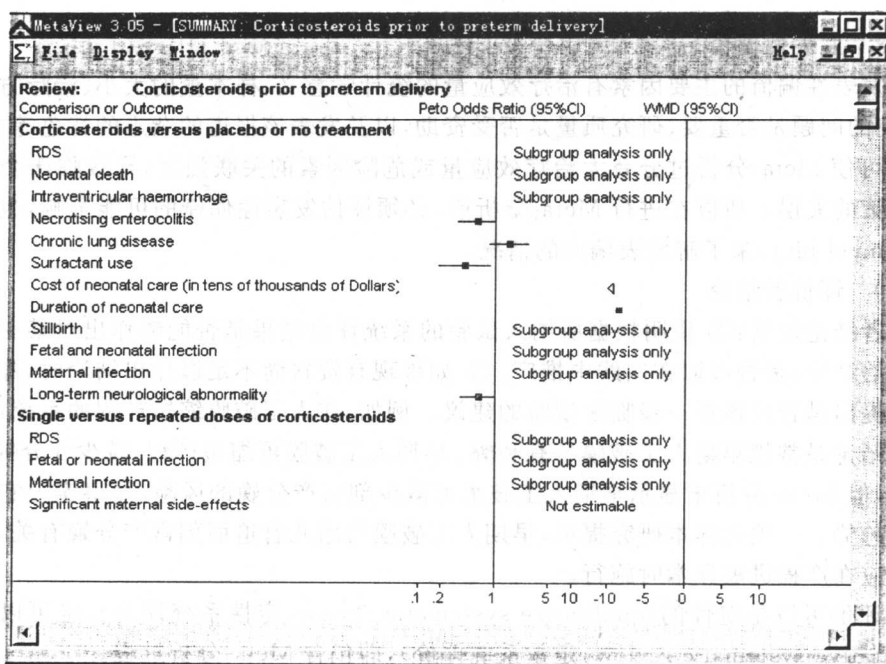


图 2-2-1 Meta-分析的比值比(OR)图

均数差值(WMD)。在对皮质类固醇与安慰剂和不治疗的比较, 判效指标中, 包括了对呼吸窘迫综合征、新生儿死亡、脑室周围出血、坏死性小肠结肠炎、慢性肺部疾病、表面活性物质的使用、新生儿护理成本、新生儿护理的延续时间、死产、胎儿和新生儿感染、母体感染、远期神经学方面异常改变以及单剂量与重复给药的比较。选中任何一项判效指标, 均可以看到其判效指标详细的 Meta-分析图, 例如新生儿死亡(图 2-2-2)

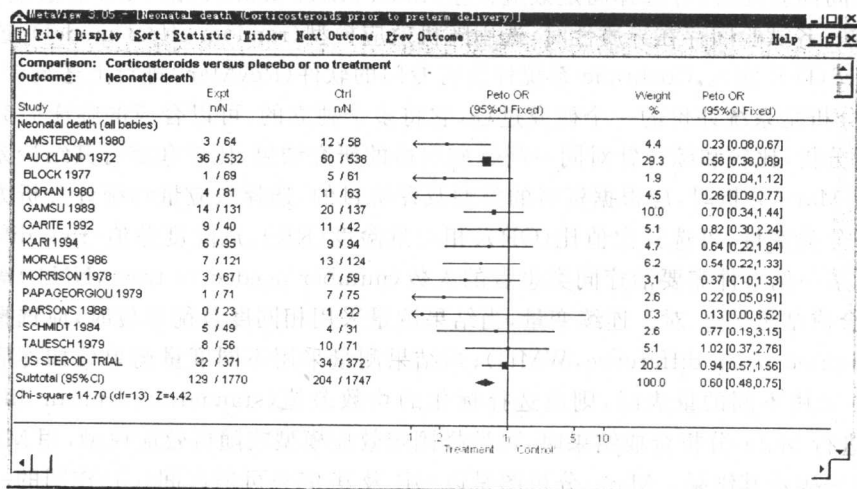


图 2-2-2 新生儿死亡的 Meta-分析图

在新生儿死亡的 Meta-分析图中, 从左到右依次为单个试验(Study)、试验组(Expt)、对

对照组(Ctrl)、比值比(OR)、权重(Weight)等项目。中间的短横线代表一个试验结果的可信区间(CI),位于横线中部的小方块代表比值比。可信区间是指比值比的真值可能存在的范围,它反映结果的精确性:范围越宽,横线越长,说明样本量较小,结论欠精确可靠;范围越窄,横线越短,说明样本量较大,结论较精确可靠。Cochrane 系统评价中使用的可信区间是 95%。中线(竖线)代表 $OR=1$;最下方的菱形符号代表所纳入的全部试验的综合结果;短横线/菱形符号与中线接触或相交,表示差异无统计学意义。对不利结局,短横线/菱形符号在中线左边表示有效,在右边表示无效;对有利结局则相反。权重(Weight)表示各单个试验结果在总体结果中所占的百分比,一般病例数越多,权重越大。

图 2-2-2 第一列显示单个试验的名称,第二、三列表示试验组和对照组。 n 表示新生儿死亡病例数, N 表示各组病例总数。以第一行 1980 年 AMSTERDAM 试验为例,在该图中可知:试验组共 64 例病人,新生儿死亡 3 例,对照组共 58 例病人,新生儿死亡 12 例,其比值比为 0.23,95%可信区间下限为 0.08,上限为 0.67;其短横线在中线左边表示治疗有效。权重为 4.4%,表示该试验结果在最后总体结果中权重为 4.4%。其他各项试验也依次可以类推。图中最下方的菱形符号代表合并以上各项试验得出的综合结果;位于中线的左侧,所采用的判效指标是新生儿死亡的不利结局,因此,上图所示 Meta-分析最后得出“早产前给予皮质类固醇治疗可有效降低新生儿死亡”的结论。

综上所述,通过对以往的临床试验进行 Meta-分析,可以发现某些问题已经有结论,避免再重复研究。并且,Meta-分析可有效地整合已有的信息,从而为合理决策提供依据,特别是当众多试验结果相互矛盾时,Meta-分析可以提供一个可靠的综合的答案。

2.3 系统评价与叙述性综述的区别与联系

系统评价是一种临床研究方法,是全面收集所有相关临床研究并逐个进行严格评价和分析,必要时进行定量合成的统计学处理,得出综合结论的过程。由于“系统”和“评价”是 Systematic review 的重要特点,与一般综述有着本质的不同,故中国循证医学/Cochrane 中心将其翻译为“系统评价”是为了避免混淆,并且比其他译名如系统综述更为准确。

系统评价和叙述性文献综述都是对临床研究文献的分析和总结,但系统评价不同于描述性综述。文献综述是一种叙述性综述,常常涉及某一问题的多个方面,如妊娠糖尿病的病理、病理生理、流行病学、诊断及预防、治疗、康复的措施,也可是仅涉及某一方面的问题如诊断、治疗等方面的综合分析。在获取原始数据或在综合发现方面,文献综述往往带着综述者的主观判断,因此综述文章的质量就无法得到恰当的评价;另一方面,综述者没有通过定量分析来综合数据,当发现相同的几个具有不同结果时,综述者往往会主观地加以选择和摒弃。而系统评价则是一种全新的文献综合评价方法,它是针对临床某一具体问题,例如糖尿病孕妇选择性分娩,“是积极处理还是期待处理?是剖宫产还是引产?”先全面、系统收集全世界所有能收集到的,包括发表的或未发表的随机对照试验,再按照临床流行病学严格评价文献的原则和方法,即评估单个试验在设计、实施和分析过程中防止或减少系统误差或偏倚和随机误差的程度,分析医学文献所报道的研究结果的真实性、研究过程中和文献里是否存在有关偏倚(bias)和混杂因素(confounder)的影响,以及可被接受的程度,筛选出符合质量标准的文献,然后进行质量评估的定性分析,并同时符合条件的研究论文进行定量的 Meta-分析。这样去粗取精,去伪存真,所得出的综合评价具有一定的深度,并可随着新的临床研究的出现得到及时更新。因此可以说,叙述性文献综述有助于了解某一疾病的全貌,而