

构建中药上市后安全性评价证据体的思考

廖 星^{1,2} 谢雁鸣¹

(1 中国中医科学院中医临床基础医学研究所, 北京, 100700; 2 中国中医科学院博士后流动站, 北京, 100700)

摘要 中药上市后安全性评价关乎用药患者在临床应用中的生命安全, 关系着国家医疗卫生政策的导向, 以及药厂对药品开发规划设计。上市后药品的安全监测无疑是所有上市后药品的重要内容, 而监测主要关注药品在上市前未曾发现的那些潜在的不良反应/事件。既往, 上市后药物监测习惯以被动监测模式或者是四期的临床试验或者是大型的流行病学调查, 这些统被称之为安全性监测或药物警戒研究。现今, 研究者们发现这些方法或方式远远不能满足当下对于上市后药品安全性评价的证据需求。为此, 为了能够弥补这些证据积累的不足, 目前亟需发展新的证据研究模式, 比如转向大样本大范围的主动监测, 采用队列观察性研究方法, 基于真实世界医疗电子数据库等等。本文基于上市后中药的特点, 从循证医学证据理念出发, 提出构建中药上市后安全性评价证据体的设想来对未来中药上市后安全性评价研究提供参考。

关键词 安全性评价; 证据体; 上市后中药

Building up Evidence Body for Post-marketing Chinese Medicines Research

Liao Xing^{1,2}, Xie Yanming¹

(1 Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing 100700, China;

2 Postdoctoral research station of China Academy of Chinese Medical Sciences, Beijing 100700, China)

Abstract Safety surveillance is the most important aspect for post-marketing research of Chinese medicine. However, passive monitoring systems or phase four clinical trials or large sample epidemiology studies were historically the common ways of surveillance of post-marketing Chinese medicine. However those studies can no longer meet the expectations, and there are many more evidence sources for post-marketing Chinese medicine safety evaluation, eg. active surveillance from observational cohort study and study on hospital information system. These data and study can provide adequate source data for post-marketing Chinese medicine research. The most prominent problem is how to integrate different evidences from multiple sources. This article suggests we should build an evidence-based system to evaluate post-marketing Chinese medicines safety.

Key Words Safety evaluation; Evidence body; Post-marketing Chinese Medicines

中图分类号: R28 文献标识码: A doi:10.3969/j.issn.1673-7202.2014.09.008

随着人们对药品上市后所发生的全面效应认识逐渐深入和医学研究不断进展, 药品上市后安全性再评价已由基于临床试验的上市后安全性再评价发展到包括基于药物流行病学的药物监测、给药方案、临床药理学(群体药代动力学)及药物经济学等内容的临床再评价, 甚至更包括药剂学、药品稳定性等药品本身品质的上市后再评价。中药上市后安全性再评价研究内容主要包括对中药上市后临床应用中不良反应/事件自发报告信息的收集, 主动监测中药上市后新的或严重的不良反应, 个体和群体的中药上市后不良反应/事件的评价和分析, 建立用药安全信息反馈的机制, 建立获益-风险控制体系, 包括制定中药风险最小化计划等的研究。中药上市后近几年的发展, 重视以循证医学理念和方法为基础展开, 循证医学评价已经深入到不

同领域。国内外有关上市后中药安全性证据分级和分类研究尚缺如, 研究者们发现系统地收集和评价有关上市后中药安全性证据有重要意义。借鉴循证医学有效性证据分类、分级的成功经验, 探索对上市后中药安全性证据评价、分类和分级标准研究的相关瓶颈问题, 并形成基本框架, 为以后该领域证据体系的研究提供方向和技术支持, 同时构建上市后中药安全性评价证据体。随着 EBM(Evidence Based Medicine, EBM)未来五个发展方向的变化: EBM 信息资源, 临床循证指南, 循证实践, 探索最佳途径, 确保临床决策符合患者价值观和喜好, 在卫生政策领域引入循证原则。证据评价发展方向也会随之有所改变。正如牛津大学循证医学中心在研发证据评价体系时提出的一个核心问题: “在所有可获得不同治疗的措施中, 哪一种措施能够提供

基金项目: 国家自然科学基金青年基金资助项目(编号: 81202776); 中国中医科学院: 客座研究员联合创新研究项目(编号: ZZ070817); “重大新药创制”科技重大专项(编号: 2009ZX09502-030)“中药上市后再评价关键技术研究”

作者简介: 廖星, 博士, 在站博士后, 助理研究员, 研究方向为循证中医药临床研究, Tel: (010) 64014411-3302, E-mail: okfrom2008@hotmail.com

通信作者: 谢雁鸣, 研究员, 博士生导师, 研究方向为中医临床评价, Tel: (010) 64014411-3302, E-mail: datamining5288@163.com

最佳的利/弊平衡?”^[1]。基于这样一个“理想化”的问题,如何获得关于这个问题的证据,将是下一个阶段EBM研究者朝向奋斗的一个目标。而有关“不同治疗措施的利弊权衡”正是当下医学领域研究热点“比较效益研究(Comparative Effectiveness Research, CER)”关注的核心所在^[2]。

1 证据评价需要适应新的发展需求

随着循证医学的迅速发展以及不断渗入到医学研究和临床实践领域,根据研究质量,对各种不同类型的研究进行分级评判形成不同级别的证据已然被广泛应用于医学领域,而且也使得广大医学工作者从中受益。作为金标准的RCT无可厚非地用于评价干预措施的疗效,但是对于疗效以外的评价,随机对照试验有很大的局限性,譬如无法进行药物不良反应的研究。根据传统的证据分级体系,临床医生可以很快从最高级别的证据,综合多个RCT的系统综述获得有关某种干预措施的疗效信息,但是对于非疗效相关的临床问题,比如干预措施的危害,则无法从这种证据评价体系中获得最佳证据,往往要转向观察性研究设计。

虽然主流证据评价体系在世界医学研究领域里起到了很大的影响,但是不可否认的一个事实是最强的推荐有可能来自于低质量的证据,最高质量的证据也不意味着可以预示最强的推荐。证据的分级和评价应该根据研究问题以及研究设计类型的质量不同而不同。临床实际中的问题,虽然以治疗/干预疗效评价为核心问题,但是还有许多其他方面的问题,比如诊断、治疗(干预)、预防、病因,因此证据的评价不仅仅只是关注干预措施的疗效,而应该拓展其他问题证据的评价^[3-4]。在治疗领域里诞生的金字塔证据分级,将RCT视为最高证据级别的划分,就不能用于预防和病因学研究领域,后两者更多时候以队列(观察性)研究为最高证据级别。那么如何区分不同问题来评价不同的证据呢?

结合医学真实世界的实际情况,应该依据不同的研究问题选择不同类型的研究证据,比如RCT用于回答治疗或干预措施疗效的最佳证据,病因学、伤害等问题多倾向于队列或病例对照研究,诊断学和疾病筛查等问题则适合用横断面设计,以及疾病自然史和预后等问题则适合运用队列设计。而对于医学证据缺乏的领域则往往可以依靠历史回顾性研究,甚至是日常经验。可以针对不同的研究问题,选择不同类型的证据以分类也是医学证据研究领域一种可行的做法。

无论是证据的分级系统还是推荐级别的分级,都应该简洁、明了,推荐强度应该和证据质量的高低判断

一致^[5-6]。而且证据分级本为医学研究者的一种主观科研行为,目的是为了获得可观的、可行的、适用的证据,为临床实践服务,因此这个过程或者说体系是在不断发展和变更的。所以即使是来自牛津大学循证医学中心如此权威机构发布的证据分级体系,自1998年诞生以来,也一直在接受挑战和争议,同时也随着临床实际应用过程中问题发现不断增多,如最新2011版的证据分级Ⅱ(Levels of Evidence 2),就是最好的说明^[7]。该更新的证据分级最鲜明的特点在于,证据的分级基于七个分类问题来划分,把临床问题分成治疗、预防、病因、诊断、预后、危害、经济学七个方面,更具有针对性和适用性。

2 安全性证据评价

有关“安全性”的英文表述词有:safety、risk、harm、adverse event、adverse effect、ADE/R。概而言之,就是医疗实践中产生了不好的结果,即“伤害”。在EBM领域里,不难发现极少有专门的RCT或RCT的系统评价专门用来评价安全性问题。原因如下:RCT无法有针对性地报告不良事件或不良反应,或者是干预疗程不够长无法发现ADE/ADR,或样本量太少,不能发现ADE/ADR有差异性研究结果。有关如何评价干预措施的“伤害”或安全性(harms/safety)一直以来也是让医学研究者特别是EBM研究者所困扰的一个研究主题。所遇到的评价困难也诸多,如一些罕见的难以发现的“伤害”并不能通过RCT来获得,反而是一些低级别的研究才是其发现的唯一证据^[8]。而一些特殊的身心伤害更是需要一些特别的其他研究设计才能获得证据^[9]。再比如吸烟和肺癌的因果关系是通过队列研究发现的,而RCT对于此类问题的研究就无法通过伦理审核。RCT是一种理想的控制混杂因素的实验设计,但是由于其经费昂贵,费时,并不适用于评价日常的临床实践。再次,好的观察性研究结果同样可以获得较之一个差质量的RCT的高级别证据。

目前针对如何评价并获得评价安全性证据,美国研究者已开始着手这方面的研究。他们认为^[10]:那些RCT往往是只有基于预先假设才能获得最有效的因果推断研究结果,而一定研究时间段里的用药并不能真正揭示实际用药人群的安全性问题。因此有关安全性结局评价的证据往往级别不高,即使是产生于一些RCT的结果评价也多是因“机遇”而产生。如何去获得高证据级别的有关安全性结局评价的证据?可以针对报告了安全性结局的RCT进行系统评价,同时也可以针对其他非RCT报告的研究设计进行系统评价。针对有些报告对比了安全性结局的研究,值得注意的

一点是,不应仅仅关注统计学上的差异性,还应该关注在同一个可信区间里的临床差异性,这样也可以获得有关安全性结局评价的有用信息,而不致遗漏。其中对于当下热门的有关干预措施安全性评价的注册登记研究如何评价提出了具体的评价框架^[11]:结果是否真实?各种偏倚是否导致干预措施看起来更安全还是使其看起来可能会造成更大的伤害,而实际这些伤害却不是由于本研究中的干预措施导致的?国内有学者也提出由于上市后药品 ADR(尤其严重 ADR)罕见,RCT 常不足以有效发现药品安全事件(样本含量较小、观察时间较短、研究对象严格挑选等)。因此,安全性评价证据分级标准与有效性评价存在差异,其最佳证据往往来自队列研究或病例对照研究,为此,还拟定了一个仅供参考的有关药物安全性评价证据级别,并提到该评价标准的层次也只是相对的,当前的,并非永久不变^[12]。

3 中医药研究领域安全性证据的评价

过去 30 多年里,随着 EBM 的应用和推广,中医药临床有效性研究已经从系统评价和随机对照试验中获得了重要的研究证据。然而,有关上市后中药安全性研究的证据研究却处于缓慢发展阶段。

此外,对于目前中医药研究领域里大量涌现的系统评价,也有研究者报道其并没有给中医药临床实践带来太多有利影响,反而引发了一连串对于中医药开展随机对照试验的反思^[13]。系统评价处于证据体系的高级别,但是多针对于干预措施的有效性,很少有单独针对安全性结局评价的系统评价,这或许是因为在 RCT 等干预性研究中有关安全性的信息比较少报告^[14-15],更或许是 RCT 并不是用来评价“伤害效应”的最佳设计^[16]。早在 2004 年英国约克大学研究者就提出,在进行系统评价的时候,不仅仅对评价干预措施的疗效结局时应该标准化,而对于如何去评价干预措施的“有害”结局时也应该系统标准化^[17]。研究者在进行评价“Harm”的系统评价时,应该首先确定好研究问题,比如评价长期效应,药物交互作用,某种对患者重要的但不严重的伤害,不用仅仅局限于某种特殊的不良事件/反应。然而,如果研究的目的在于获得之前可疑的不良事件/反应,那么对于原始的监测数据的分析往往更加优于一个系统评价^[18]。换言之,有关 ADR 的证据,抛开 EBM 传统的设计类型,往往来自于其他渠道的数据,如国家 SRS 数据也可以成为证据源。因此一味追求金标准、高级别证据产生的模式,是否适宜于中医药临床研究值得深思,同时就上市后中药安全性研究是否合适使用随机对照试验也是目前方法学专

家争论的热点问题之一。虽然目前医学研究领域的标准已趋于成熟,并逐步统一,但在中医药研究领域特别是上市后中医药安全性证据评价方面上存在很大的挑战。因此,研究制定符合循证医学思想,满足上市后中医药研究领域和实践需要的高质量证据分类分级标准和推荐意见强度有着重要意义。

4 如何构建上市后中药安全性评价证据体

中国中医科学院中医临床基础医学研究所谢雁鸣研究员等首次在上市后中药临床安全性研究领域提出构建“安全性证据体”的理念^[19],首次针对该领域尚无证据分类分级理念的现状,借鉴循证医学有效性证据分类分级的成功经验,探索对该研究领域证据整合和分级。根据当前可得的证据,将前瞻性大样本的注册登记研究和来自国家不良反应中心自发呈报系统的 SRS 数据列为最高级证据、将来自 SR 或 RCT 中报告的 ADR/ADE 和医院 HIS 数据回顾性分析结果列为第二等证据、将来自医院临床实际的个案病例讨论报告和文献中个案报告以及其他研究类型报告的 ADR/ADE 列为第三等证据、将相关国家政府及相关机构颁布的文件列为第四等证据,将专家意见和共识列为第五等证据。有关上市后中药安全性证据评价体系的构建需要从点、线、面相结合的整体体系构建。从点的角度来看,五种不同来源的证据可以分级进行评价,依据质量强度来判断警戒信息强度。从线的角度来看,凡是具备五种证据源中的两种及其以上者则根据证据源之间结果是否一致获得警戒信息强度,进行评价。从面的角度来看,五种证据源可以形成一体化证据源,即如果不具备高质量的临床研究,但是多个源头的研究结果都具备异质性,则所形成的证据具有高度可信性,可以形成强的警戒信息。由于当前 EBM 领域有关药物安全性证据评价体系不能满足现实需求的现状,以及上市后中药安全性证据评价的特殊性,本研究建议在该领域的证据评价应该以“证据体”,以点、线、面相结合的形式,进行灵活多元的证据评价。从中药药物警戒研究角度出发,多角度获得发掘上市后中药安全性评价的证据。

总之,当今医学时代,循证医学的迅猛发展,证据分级的逐渐成熟,标志着一个以证据为基础的新医学时代的到来。临床证据的分类分级和临床推荐意见强度系统的逐渐成熟,但其内容复杂、应用局限、标准各异,对指导全球范围内各级医疗机构的循证实践并非万能。近年来,社会对上市后中药安全性问题越来越关注,引起了国家政府和中医药界的高度重视,对于其临床安全使用的证据支持呼声也越来越高。研究者

们越来越发现系统地收集和评价有关上市后中药安全性的研究证据有着时间的紧迫性和时代重要意义。因此,为了能够系统、完整地上市后中药相关的既有研究证据进行整合,运用循证医学的技术和方法进行全面研究,使应用者获得最佳临床使用证据,同时使得上市后中药的临床使用和研究信息在全世界范围内进行传播,非常必要。构建中药上市后安全性评价证据体研究是一个需要多方积极参与的长期动态过程,尤其要充分调动中药生产企业的积极主动性,使其充分认识和理解中药上市后安全性证据积累的重要性,资助有条件的临床医疗机构开展中药上市后的安全性再评价工作,获得最佳证据。同时,国家药品管理部门要加强管理和监督,从证据角度出发,制定和健全中药上市后再评价的政策、法规和技术标准将使决策更为明智和审慎。

参考文献

- [1] Jeremy Howick, Iain Chalmers, Paul Glasziou, Trish Greenhalgh, Carl Heneghan, Alessandro Liberati, Ivan Moschetti, Bob Phillips, and Hazel Thornton. "Explanation of the 2011 Oxford Centre for Evidence – Based Medicine(OCEBM) Levels of Evidence (Background Document)" [C]. Oxford Centre for Evidence – Based Medicine, 2011.
- [2] Sox HC, Greenfield S. Comparative effectiveness research: a report from the Institute of Medicine[J]. Ann Intern Med, 2009, 151(3): 203 – 205.
- [3] Sackett DL, Wennberg JE. Choosing the best research design for each question[J]. BMJ, 1997, 315: 1636.
- [4] Guyatt GH, Haynes RB, Jaeschke RZ, Cook DJ, Green L, Naylor CD, et al. Users' guides to the medical literature: XXV. Evidence – based medicine: principles for applying the users' guides to patient care[J]. JAMA, 2000, 284: 1290 – 1296.
- [5] Atkins D, Best D, Briss PA, et al. GRADE Working Group. Grading quality of evidence and strength of recommendations [J]. BMJ, 2004, 328 (7454): 1490.
- [6] Guyatt GH, Oxman AD, Vist GE et al; GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendation[J]. BMJ, 2008, 336(7651): 924 – 926.

- [7] OCEBM Levels of Evidence Working Group. "The Oxford 2011 Levels of Evidence". Oxford Centre for Evidence – Based Medicine. [http://www.cebm.net/index.aspx? o = 5653](http://www.cebm.net/index.aspx?o=5653).
- [8] Eikelboom JW, Mehta SR, Pogue J, Yusuf S; Safety outcomes in meta – analyses of phase 2 vs phase 3 randomized trials: Intracranial hemorrhage in trials of bolus thrombolytic therapy [J]. JAMA, 2001, 285(4): 444 – 450.
- [9] Scaf – Klomp W, Sanderman R, Wiel HB, Otter R, Heuvel WJ; Distressed or relieved? Psychological side effects of breast cancer screening in The Netherlands [J]. J Epidemiol Community Health, 1997, 51(6): 705 – 710.
- [10] Delfini Group, LLC, 2010. Delfini Appraisal Tool. Safety Issues & Considerations http://www.delfini.org/Delfini_Pearls_Basics_Safety.pdf.
- [11] Delfini Group, LLC, 2010. Delfini Appraisal Tool. Registry Evaluation Checklist; Safety of Interventions http://www.delfini.org/Delfini_Tool_RegistryAppraisalChecklist_SafetyInterventions.pdf.
- [12] 李幼平, 文进, 王莉. 药品风险管理: 概念、原则、研究方法与实践 [J]. 中国循证医学杂志, 2007, 7(12): 843 – 844.
- [13] 罗辉, 刘建平. 从系统综述看中国随机对照试验的质量 [J]. 中西医结合学报, 2011, 9(7): 697 – 701.
- [14] Ioannidis JPA, Lau J; Completeness of safety reporting in randomized trials: an evaluation of 7 medical areas [J]. JAMA, 2001, 285: 437 – 443.
- [15] Loke YK, Derry S; Reporting of adverse drug reactions in randomized controlled trials – a systematic survey [J]. BMC Clin Pharmacol, 2001, 1: 3.
- [16] Cuervo GL, Clarke M; Balancing benefits and harms in health care [J]. BMJ, 2003, 327: 65 – 66.
- [17] Heather M McIntosh, Nerys F Woolcott and Anne – Marie Bagnall. Assessing harmful effects in systematic Reviews [J]. BMC Medical Research Methodology, 2004, 19: 1 – 6.
- [18] Medawar C, Hershheimer A; A comparison of adverse drug reaction reports from professionals and users, relating to risk of dependence and suicidal behaviour with paroxetine [J]. Int J Risk Saf Med, 2004, 16: 5 – 19.
- [19] LIAO Xing, Nicola Robinson. Methodological Approaches to Developing and Establishing the Body of Evidence on Post – marketing Chinese Medicine Safety [J]. Chin J Integr Med, 2013, 19(7): 494 – 497.

(2014 – 08 – 06 收稿 责任编辑: 洪志强)

(上接第 1140 页)

- [23] 赵玉斌, 肖颖, 谢雁鸣. 清开灵注射液过敏患者血清免疫球蛋白及细胞因子变化规律的临床研究 [J]. 中国中药杂志, 2011, 36(10): 1382 – 1385.
- [24] 王连心, 唐浩, 谢雁鸣, 等. 巢式病例对照研究在 HIS 真实世界参麦注射液疑似过敏因素分析中的应用 [J]. 中国中药杂志, 2013, 38(18): 3019 – 3023.
- [25] 王连心, 程豪, 谢雁鸣, 等. 基于倾向性评分的不同疗程使用参麦注射液肝肾功能变化分析 [J]. 中国中药杂志, 2013, 38(18): 3053 – 3059.
- [26] 姜俊杰, 唐浩, 谢雁鸣, 等. 基于医院信息系统的参麦注射液不同使

用剂量对肾功能作用分析 [J]. 中国中药杂志, 2013, 38(18): 3060 – 3067.

- [27] 庄严, 谢邦铁, 谢雁鸣, 等. 中药注射剂不良反应信号自动预警算法的研究与实现 [J]. 中国中药杂志, 2012, 37(17): 2695 – 2697.
- [28] Hopstadius J, Niklas NG, Bate A, et al. Impact of stratification on adverse drug reaction surveillance [J]. drug safety, 2008, 31(11): 1035 – 1048.
- [29] 孙亚林. 基于自发呈报系统数据的药物不良反应信号检测数据挖掘研究 [D]. 上海: 第二军医大学, 2010.
- [30] 王连心, 向永洋, 谢雁鸣, 等. 参麦注射液自发呈报系统的药物警戒研究 [J]. 中国中药杂志, 2013, 38(18): 2987 – 2993.

(2014 – 08 – 06 收稿 责任编辑: 洪志强)