

·述评·

大数据时代的药品安全主动监测： 对照选择的挑战与机遇

王胜锋 詹思延

100191 北京大学公共卫生学院流行病与卫生统计学系

通信作者：詹思延，Email:siyan-zhan@bjmu.edu.cn

DOI:10.3760/cma.j.issn.0254-6450.2016.07.001

【摘要】 电子医疗数据已成为大数据时代开展药品安全主动监测的重要资源。基于此确认药品与不良事件是否存在关联，要回归传统的流行病学研究设计，选取恰当的对照进行对比。本文主要阐述不同对照选取的原理、适用情形，介绍、评价并比较各种对照选取的思路与参数，引入对照选择批量化实现的进展性成果，以期为我国利用电子医疗数据开展上市后药品安全性监测提供方法学参考。

【关键词】 主动监测；电子医疗数据；对照选择

基金项目：国家自然科学基金(81502884, 81473067)

Active surveillance of adverse drug reaction in the era of big data: challenge and opportunity for control selection Wang Shengfeng, Zhan Siyan

Department of Epidemiology, School of Public Health, Peking University, Beijing 100191, China

Corresponding author: Zhan Siyan, Email: siyan-zhan@bjmu.edu.cn

【Abstract】 Electronic healthcare databases have become an important source for active surveillance of drug safety in the era of big data. The traditional epidemiology research designs are needed to confirm the association between drug use and adverse events based on these datasets, and the selection of the comparative control is essential to each design. This article aims to explain the principle and application of each type of control selection, introduce the methods and parameters for method comparison, and describe the latest achievements in the batch processing of control selection, which would provide important methodological reference for the use of electronic healthcare databases to conduct post-marketing drug safety surveillance in China.

【Key words】 Active surveillance; Electronic medical records; Control selection

Fund programs: National Natural Science Foundation of China (81502884, 81473067)

2015年，我国“十三五”规划提出“实施国家大数据战略，推进数据资源开放共享”，国务院亦发布《促进大数据发展行动纲要》，旨在大力促进中国数据技术的发展。医学作为大数据最活跃的科研领域之一，目前全球积累了海量的生物医学数据，如何对其进行合理高效的利用已经成为一个关键问题^[1]。与此同时，药品安全作为各国持续关注的重要公共卫生问题，仅基于不良反应自发报告这种被动监测模式已经不能满足风险管理的需求，而以流行病学调查为主的主动监测又因高昂的经费和时间投入，实际开展亦不多见^[2]。以电子病历(EMR)、医保数据库(Claim database)为代表的一组电子医疗数据，真实记录了病例的症状、体征、实验室检查、诊断和处方、费用等信息，能较好地反映临床实践，也为及时高效开展药品安全主动监测提供了新的思路^[3]。

一方面，借助人工智能、机器学习、神经网络等多种数据挖掘技术，可以充分发挥大数据发现信号(Signal detection)的优势，为获取各种未知信号提供可能^[4]；另一方面，通过快速识别某一药品的大批服用患者及未服药的对照患者，可以确认药品与不良事件之间的关联，即强化和确认信号(Signal refinement)，而这需要回归传统的流行病学设计。对比是流行病学研究方法的核心所在，对照的优劣直接决定了主动监测的质量，如何选择可比的对照成为设计过程中关键且耗时的环节^[5]。依赖于大数据的对照获取尽管相对快速广泛，但便捷性、可及性提高的同时，不同选择之间孰优孰劣、如何取舍的新问题自然产生，并成为目前药物流行病学领域方法学研究关注的热点。

一、对照选择的常用方法

海量电子医疗数据尽管容量大、种类多、产生和更新速度快,但在回答某一特定的药品安全问题时,仍经常面临对照个体无从选择的困境。如某些数据库中对于所有患者都使用的基础药物,不存在未用药对照的记录^[6];少有数据库包含生活习惯、社会经济地位、遗传因素等常见的混杂因素,无法调整对应的混杂偏倚^[7]。某些情况下,由于指征偏倚、疾病严重程度等难控制、难测量混杂因素的存在,欲直接寻求可比的对照个体更是无计可施^[8]。此外,大数据可以降低的是信息收集成本,但随着可及数据量的增大,数据的购买或共享成本、整合和分析成本毋庸置疑相应增加,加之对于数据隐私与伦理安全的要求愈来愈高^[9],如何权衡成本效益并明确最优样本量也值得研究者思考。

经典的观察性流行病学研究都是进行个体间比较,或者比较暴露组与非暴露组的结局事件率(队列研究),或者比较病例组与对照组的暴露率(病例-对照研究)。鉴于进行个体间比较选取对照的局限,学者开始将目光转向一组以自身作对照的设计方法,尝试将个体间比较转变为个体内比较,通过充分利用电子医疗数据库中用药时间、顺序等信息,回答某一药品安全性问题^[10]。此时不但无需再寻找可比的对照个体,对样本量和变量个数的需求也有所降低,同时可自动控制遗传等不随时间改变的因素,尤其部分难调整或难测量的混杂因素^[11]。

药品安全主动监测领域的研究设计基本原理见图1。纳入病例(特指发生所关注的不良事件),同时选择非病例作对照者为病例-对照研究,单纯只纳入病例者,可细分为自身对照病例系列

(Self-controlled case series, SCCS)、病例交叉研究(Case-crossover study, CCO)、病例-时间-对照研究

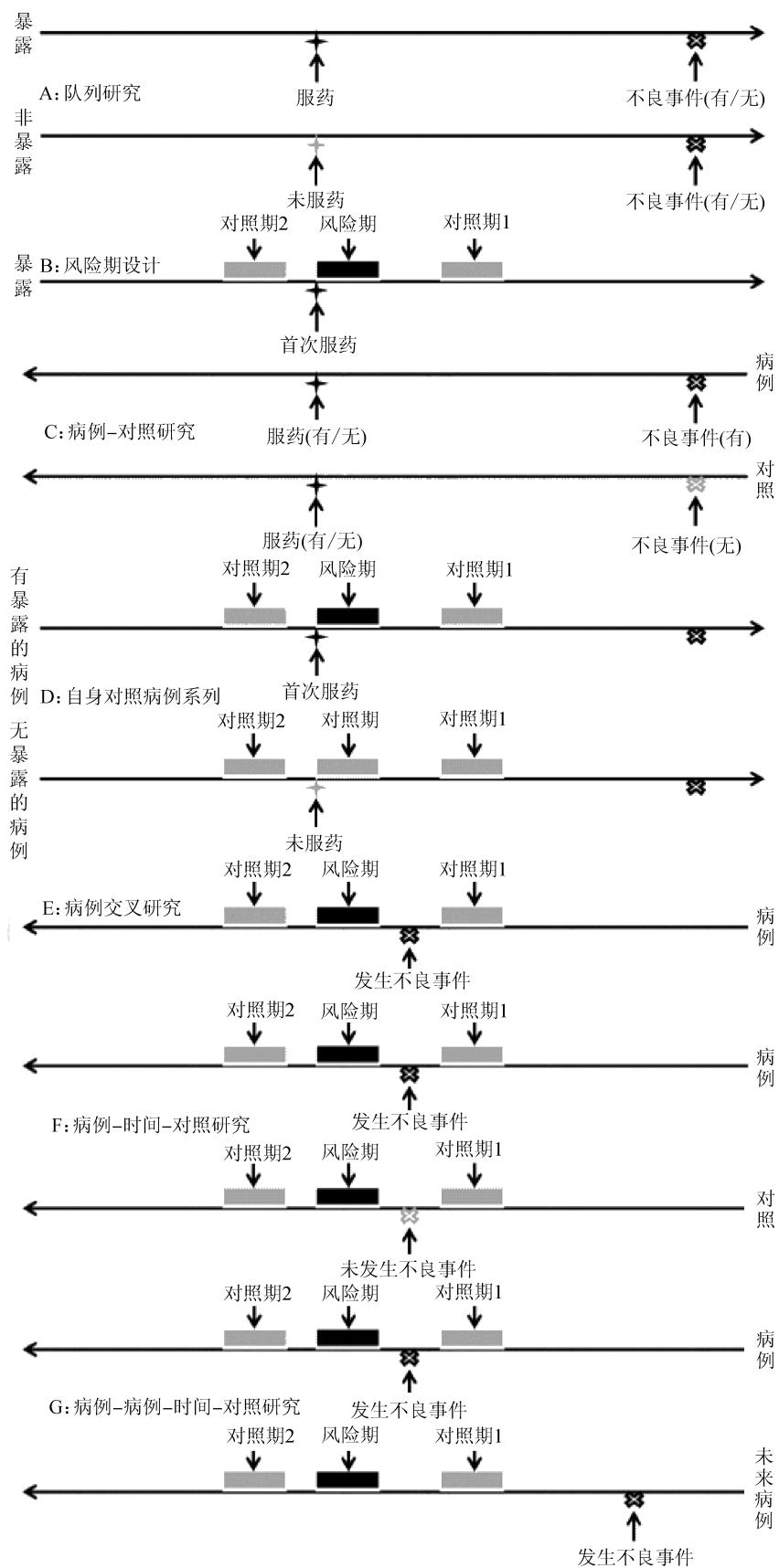


图1 药品安全主动监测的各种对照选择基本原理

(Case-time-control, CTC)、病例-病例-时间-对照研究(Case-case-time-control, CCTC) 4种研究设计(表1);纳入暴露者(特指使用所关注的药品),同时选择非暴露者作对照时为队列研究,单纯只纳入暴露者时,则为风险期设计(Risk-interval design, RI)。其实,SCCS、CCO、CTC、CCTC和RI并非刚刚提出的设计思路,国内也有学者分别介绍过各自的原理^[12-15]。遗憾的是,由于这组方法在适用范围、资料需求、统计分析等方面均相对独特(如对多时点信息的需求较高),导致之前应用受限,但海量电子医疗数据的挖掘使用为其发展和推广带来了契机。

多种可选方法的存在,满足了不同特征电子医疗数据的使用需求,很快被国外学者所接受和认可,新的研究成果不断发表于世界权威期刊(如*Lancet*^[16]、*Circulation*^[17]、*BMJ*^[18]等)。同时,被多家国际权威机构作为使用电子医疗数据的推荐策略,如美国的小规模安全警戒计划方法学组(Mini-sentinel)、观察医疗结果合作项目(The Observational Medical Outcomes

Partnership, OMOP)和欧盟的药品不良反应工作组(EU-ADR)等^[19-21]。但可选方法增多的同时,也带来了新的困惑,即当数据同时满足多种方法时,多个不同的结果如何解读、如何抉择,按照美国Mini-sentinel方法学组的总结,这种情况至少占到实际工作情境组合的1/3^[22-23]。

二、对照选择的评价与比较

对照的选择是否恰当,需要评价。这种评价工作目前主要由欧洲地区、美国的方法学工作组引领,他们多围绕同一电子医疗数据库,选取一定数量的已知药品-不良事件对,比较不同设计的综合信号发现能力^[19,21]。例如,OMOP工作组曾结合53对药品-不良事件对(其中9个阳性对),比较队列研究、CC、SCCS、CS的信号发现效果,并未发现哪种方法更为优越^[24]。随后,当将药品-不良事件对增加至399对(其中165对阳性),发现SCCS的信号发现能力可能最好^[25]。目前基本形成的共识是,不同方法在信号发现层面的能力差异很大程度上取决于药品-不良

表1 可用于药品安全主动监测的各种对照选择基本信息

对照选择	研究设计(提出时间)	基本思路	适用条件	统计分析												
暴露 & 非暴露 Cohort study	队列研究(19世纪) Cohort study	<table border="1"> <tr> <td></td><td>发生结局</td><td>未发生结局</td></tr> <tr> <td>暴露组</td><td>A</td><td>B</td></tr> <tr> <td>对照组</td><td>C</td><td>D</td></tr> </table> <p>Risk ratios, e.g. relative rate (RR): $\frac{A/(A+B)}{C/(C+D)} = \frac{\text{Incidence}_{\text{exposed}}}{\text{Incidence}_{\text{unexposed}}}$</p>		发生结局	未发生结局	暴露组	A	B	对照组	C	D	不适用于罕见结局	Cox回归			
	发生结局	未发生结局														
暴露组	A	B														
对照组	C	D														
病例 & 对照 Case-control study, CCS	病例-对照研究(20世纪) Case-control study, CCS	<table border="1"> <tr> <td></td><td>病例组</td><td>对照组</td></tr> <tr> <td>暴露</td><td>A</td><td>B</td></tr> <tr> <td>非暴露</td><td>C</td><td>D</td></tr> </table> <p>Odds ratios (OR_{Case-control}): $\frac{A/C}{B/D} = \frac{AD}{BC}$</p>		病例组	对照组	暴露	A	B	非暴露	C	D	不适用于罕见暴露	logistic回归 (匹配时,采用条件logistic回归)			
	病例组	对照组														
暴露	A	B														
非暴露	C	D														
仅病例 Self-controlled case series, SCCS	自身对照病例系列(1995年) Self-controlled case series, SCCS	<p>Risk ratios, e.g. relative rate (RR): $\frac{\text{Incidence}_{\text{risk period}}}{\text{Incidence}_{\text{control period}}}$</p>	暴露:瞬时效应 结局:急性事件 结局不影响暴露	条件Poisson回归												
	病例交叉研究(1991年) Case-crossover study, CCO	<table border="1"> <tr> <td></td><td colspan="2">对照期</td></tr> <tr> <td></td><td>暴露</td><td>非暴露</td></tr> <tr> <td>风险期</td><td>A</td><td>B</td></tr> <tr> <td>非暴露</td><td>C</td><td>D</td></tr> </table> <p>Odds ratios (OR_{Case-crossover}): $\frac{A/C}{B/D} = \frac{AD}{BC}$</p>		对照期			暴露	非暴露	风险期	A	B	非暴露	C	D	暴露:瞬时效应 结局:急性事件 混杂:无随时间变化的 混杂	条件logistic回归
	对照期															
	暴露	非暴露														
风险期	A	B														
非暴露	C	D														
	病例-时间-对照研究(1995年) Case-time-control, CTC	在CCO基础上,选择未发生所关注结局事件 的个体作对照	暴露:瞬时或慢性效应 结局:急性事件	条件logistic回归												
	病例-病例-时间-对照研究 (2011年) Case-case-time-control, CCTC	<p>在CCO基础上,选择未来发生所关注结局事 件的个体作对照</p> <p>Ratio of Odds = $\frac{OR_{\text{Case-crossover}}}{OR_{\text{Case-control}}}$</p>	暴露:瞬时或慢性效应 结局:急性事件	条件logistic回归												
仅暴露 Risk-interval design, RI	风险期设计(2001年) Risk-interval design, RI	<p>Risk ratios, e.g. relative rate (RR): $\frac{\text{Incidence}_{\text{risk period}}}{\text{Incidence}_{\text{control period}}}$</p>	暴露:瞬时效应 结局:急性事件	条件Poisson回归												

事件对与电子医疗数据库的特征^[26-27]。然而针对某一具体的药品-不良事件对,哪种方法的信号强化、验证能力更强,结果更为稳定,仍需进一步确认,为此OMOP工作组建议积极开展该方面研究,以便总结经验形成参考。本文亦侧重介绍该部分内容。

1. 理论基础及比较思路:不同的研究设计都有各自的最佳适用情形,已有学者对此进行梳理和总结^[10,28]。如队列研究适用于非罕见结局,病例-对照研究适用于非罕见暴露;SCCS、CCO与RI均适用于暴露效果为瞬时效应,结局产生为急性事件;CTC与CCTC均要求结局为急性事件,但暴露效果可急可缓。各方法的要求或限制本质上很大程度与所需求资料的可及性有关,如CCO关注发生不良反应的个体,比较的是其在结局发生前风险期与对照期药物暴露频率的差异^[13]。传统的CCO通过开展专题调查,由患者回忆在不同时点的各种药品暴露情况,可行性和精确性均较差。但如今利用数据抽提、转换和链接技术可以快速从已有电子医疗数据中得到所需信息(有时仍需辅助调查),既简单方便,又可控制某些混杂。随着电子医疗数据在药品安全领域的应用不断增加,越来越多的学者意识到个体内比较方法的优势,但在具体使用时,除了特别推荐某一种研究设计的典型情境(如关注儿童期接种疫苗的长期不良反应时推荐SCCS),更多要面对的是各种方法均适用时(如同一研究者对同一资料采用了几种方法,或者不同研究者对不同资料采用不同的方法)所得到的众多结果如何解读和比较^[29]。

回答这一问题,理想状态是存在某一组数据适用于各种不同设计,在此基础上比较各自针对同一问题的结果^[30]。这无疑对数据源提出了极高要求,因为不同的研究设计利用数据的不同信息层面,即便处于大数据时代,可同时满足上述7种对照选择的数据也难以寻觅。为此学者提出采用Monte Carlo等数据模拟技术,通过构建一个人群队列,然后在此基础上衍生其他6种设计的思路^[31]。队列人群中,一般暴露与非暴露的比例设定为1:1,并按照年龄、性别进行匹配,以确保统计效率最高。模拟中,研究者可以自主设定药品与不良事件的真实关联,直接比较不同设计所得到的关联强度与真实值的区别,并可充分探索不同参数设置对结果的影响。针对疫苗安全,McClure等^[30]与Glanz等^[31]开展了两项研究分别考虑不同的影响参数,均提示队列研究的效果相对较好,这与队列研究在观察性流行病学研究中的地位相一致,具有重要的实际意义。

针对一般药品与不良事件的关联也可以考虑把队列研究的结果作为金标准与其他方法进行比较。

2. 参数设定与评价指标:目前以自身作对照的几种方法间效果比较,数量有限且结果存有差异,究其原因如OMOP工作组所提醒,各种方法的效果直接受其参数设定影响,而每种方法又都涉及多种参数设定^[8]。以SCCS为例,Musonda^[32]提出暴露率、结局事件率、结局人数、发生结局个体的暴露比例、暴露与结局相对危险度、风险期长短、对照期长短及个数、随时间变化的混杂因素分布及效应等参数。此外由于该研究仅关注SCCS,尚未将随时间固定的混杂因素这一参数考虑在内。已有研究中,Viboud等^[33]考虑了风险期长短、暴露率、暴露时长分布、暴露与结局相对危险度4个参数,Glanz等^[31]考虑了结局事件率、结局人数、随时间变化的混杂因素、随时间固定的混杂因素、暴露与结局相对危险度5个参数,McClure等^[30]考虑了结局事件率、结局人数、暴露与结局相对危险度3个参数,而Li等^[34]考虑样本量、风险期长短、暴露与结局相对危险度和随时间变化的混杂因素4个参数。参数选择的不尽相同反映出学者对于该问题的认识尚未统一,直接影响着研究结果的可比性,但尚无研究综合考虑上述参数。值得一提的是,CCTC由哈佛大学Wang等^[35]于2011年提出,将未来病例作为当前病例的对照,用以避免暴露的时间趋势所带来的混杂。该方法将导致样本量至少损失一半,影响统计效率,到底与其他方法相比,在处理随时间变化的混杂因素上优势如何,仍无实例加以说明^[7]。

任何数据模拟,都建立在一定的前提假设基础上。以Douglas等^[36]利用英国临床研究数据集(Clinical Practice Research Datalink)探索奥利司他(Orlistat)与急性肝损伤的关联为例,即包括3条假设:①复发结局相对独立;②结局的出现不影响之后药物的使用;③整个观察期间不存在截尾数据。前提假设很多情况下偏于理想,但为了可行却常不得不如此,以求简化模拟情境。前提设定之后,需要结合所选参数,明确研究所涉及的基本变量,并按照已知规律逐一拟合所需数据。药品的暴露剂量、频率和时间一般基于已有数据库的真实数据信息采用Bootstrap抽样等方法拟合,而发生不良事件个体的分布拟合常基于暴露人时、事件发生率和关联强度共同组成的函数^[31],其中关联强度的大小可自由设定。以自身作对照的设计需要考虑对照期的暴露和结局分布(多指此次暴露前),此处拟合常基于一般

人群中的暴露率、暴露持续时间和事件发生率,而风险期长短、对照期长短及个数则需要结合具体的专业知识^[37]。此外,年龄常被选作随时间变化的混杂因素,主要由于一般研究均需要控制年龄,且年龄较易获取,其影响也比较容易发现。年龄常采用Beta分布[α, β]拟合,与不良反应的关联可结合实际,拟合弱对称、弱单调递增、强对称、强单调递增等多种情境^[32]。

实际 Monte Carlo 模拟过程中,由于参数较多,常采用逐步模拟的思路,每一步针对一个特定参数,同时为减少运算次数,后续参数设置会根据前面模拟结果适当精简。一般来讲,常要求每种情境组合模拟 10 000 个独立样本,以确保 95% 可信区间的标准误不超过 0.2%^[32]。相比模拟过程的繁琐,效果评价部分比较简单,常用指标包括关联强度指标(*RR* 或 *OR*)的点值及 95% 可信区间、真实值低于或高于 95% 可信区间的比例、参数 β 与真实值(或队列研究结果)的偏差和均方误差(Mean Squared Error, *MSE*)等。

三、对照选择的批量化实现

目前,针对疫苗不良反应主动监测选择何种对照的研究相对成熟,SCCS 被高度重视和推荐,如美国儿童流感疫苗安全性监测白皮书已将其明确为首选方法^[38]。始建于 1990 年的美国疫苗安全性数据库链接系统(Vaccine Safety Datalink, VSD)也越来越多地采用基于 SCCS 的设计,并不断探讨其他自身病例对照设计的应用价值^[39]。与此同时,随着电子医疗数据在主动监测中应用不断增多,新的实际需求也不断提出。一方面,需要明确符合成本-效益的最小有效数据集。最小有效数据集体现在记录数与变量数,尤其当涉及到数据库购买、需链接其他数据库获取更多信息、或需进一步调查核实部分信息时,所需信息的多少将直接影响经济、时间成本。不同的对照选择方式对数据记录、变量的需求大不相同,确定对照选择方式就等同于明确了对数据集的要求^[6]。另一方面,实际监管工作中,药品-不良事件对组合繁多,面对的分析任务常常是批量化、多进程的,这无形中要求能够针对各个问题快速确定可行、最优的研究策略,以便做出及时反馈^[19]。最早学者尝试利用决策树分析,仅基于成本和不良事件率两个指标制定了一套研究方法推荐流程,但过于宽泛^[40]。2012 年,美国 Mini-sentinel 工作组确定研究目的、暴露特征、结局特征、暴露与结局的关联特征四组主要指标,细化为 11 个条目,经专家讨论根据

每个条目不同选项的组合,将药品安全性问题归纳为 64 种情境(表 2)^[41]。其中,个体内混杂主要针对指征偏倚(Indicator bias),如降糖药的使用是糖尿病的标志,而糖尿病是心血管事件的危险因素,因此简单比较个体自身用药前后的心血管事件发生率容易得出降糖药增加心血管事件风险的错误结论^[19]。

基于上述情境分类,工作组于 2013 年开发出一款基于 Excel 宏程序的快速分析工具 PROMPTS (Prospective Routine Observational Monitoring Program Tools)^[23],并随后以 SAS Macro 模块嵌入分布式数据的分析模板,极大方便了美国 FDA 利用电子医疗数据主动监测某一新上市药品或生物制品的潜在不良反应^[42]。以比较利伐沙班(rivaroxaban)与华法令(warfarin)的缺血性卒中、颅内出血和消化道出血风险为例,研究者只需要结合专业知识,明确用药为瞬时暴露,其作用风险立即出现,作用风险窗持续较长,不良反应为突发事件,个体内混杂因素不可忽略,即可确定最优选择为队列研究,自动开展后续分析^[43]。目前该工具仍有待完善之处:其一,仅考虑了队列研究和 SCCS 两种研究设计,并未纳入 RI、CCO 及其衍生类型等,不能有效处理结局事件影响暴露频率的情境,如华法令导致出血后,本身会降低华法令的暴露;其二,对两种方法均可的情境,并未给出进一步的具体推荐,而两者对于数据的需求差别很大,有必要明确,这也与前述相关研究不够充分有关,制约了对应环节的流程细化。

综上所述,欧美发达国家利用电子医疗数据开展主动监测所证实的明显优势,带动了业界对该领域的关注,也推动着包括我国在内的各国科研人员在此方向上开始不断的摸索与实践。我国仍处在一个初级起步阶段,多为学术界自发探索,利用电子医疗数据开展主动监测的研究本身数量有限。这一现象与我国大陆地区的电子医疗数据分类和记录形式与国外存在显著差别有关,因为标准化水平较低、结构类型较多,导致相当数量的国内学者仍积极致力于如何整合多种来源的数据^[44]。实际上,整项工作需要一系列配套基础,包括政策的导向和资金的支持,政府部门、医疗卫生机构、各学术机构的共同努力和通力合作^[45],欣喜的是已经看到良好的发展态势。《中共中央国务院关于深化医药卫生体制改革的意见》指出实现卫生信息化互通互联、信息共享,呼吁全国执行统一的数据标准和规范,已取得初步成效,为未来有效挖掘电子医疗数据、开展药品安全主动监测奠基铺路。同时,国内一些地区也开始出

表2 美国Mini-sentinel方法学组归纳的64类药品-不良事件情境组合

暴露持续性	暴露风险窗 出现时间	暴露风险窗 持续时间	混杂大小		关注结局事件	
			个体内	个体间	突发急性	潜伏慢性
			可忽略	需要关注		
瞬时(如疫苗接种,首次使用某种药物,包括间断药物使用)	立即	短	可忽略	可忽略	自身对照(或队列)	队列(或自身对照)
				需要关注	自身对照(或队列)	自身对照或队列
			需要关注	可忽略	队列	队列
		长		需要关注	队列(或自身对照)	队列
			可忽略	可忽略	队列(或自身对照)	队列
				需要关注	自身对照或队列	自身对照或队列
	延迟	短	可忽略	可忽略	队列(或自身对照)	队列
				需要关注	自身对照或队列	自身对照或队列
			需要关注	可忽略	队列	队列
		长		需要关注	队列	队列
			可忽略	可忽略	自身对照或队列	自身对照或队列
				需要关注	队列	队列
持续(如慢性药物使用,植入物的连续暴露等)	立即	短	可忽略	可忽略	队列(或自身对照)	队列
				需要关注	自身对照或队列	队列(或自身对照)
			需要关注	可忽略	队列	队列
		长		需要关注	队列	队列
			可忽略	可忽略	队列	队列
				需要关注	队列(或自身对照)	队列
	延迟	短	可忽略	可忽略	队列(或自身对照)	队列
				需要关注	队列(或自身对照)	队列
			需要关注	可忽略	队列	队列
		长		需要关注	队列	队列
			可忽略	可忽略	队列	队列
				需要关注	队列	队列

注:某些特征属性的确定相对主观(如立即或延迟出现),不同的研究情形可以酌情设定;表中仅有一种研究类型时,代表强烈推荐;表中有两种研究类型时,括号外的为推荐,括号内的为酌情考虑,若无括号区分则两种方法相当,需权衡混杂与错分的影响

现覆盖较完备、构建较统一的医疗保健信息系统,如宁波市鄞州区、福建省厦门市等,国内学术界与这些信息资源的结合正在探索中^[45]。以广东省药品不良反应监测中心为代表的监管机构也主动联手多家医院构建大规模区域化电子医疗数据库,摸索基于其中的信号挖掘与验证技术和经验。既往朱田田等^[46]多位学者在结合各种信号挖掘技术,及时快速从电子医疗数据库中挖掘更多药品不良反应信号领域做了很多工作。国内学者对于强化与验证信号的流行病学设计关注仍很少,就采用以自身作对照设计开展主动监测而言,目前仅台湾地区学者有屈指可数的几项研究^[47]。而从主动监测实际工作出发,信号发现永远仅是第一步,对所发现信号的因果确认才是药品安全更为关注的问题^[4]。将来采用电子医疗

数据开展主动监测将不可避免的要面对“如何选择对照”问题,围绕其中的思考和探讨将成为上市后安全性监测任重而道远的一部分,如何学习国际先进技术并充分结合国情建立可行高效的技术工具将是提升我国药品风险管理水平的关键链条之一。

利益冲突 无

参考文献

- [1] Trifonova OP, II'In VA, Kolker EV, et al. Big Data in Biology and Medicine: Based on material from a joint workshop with representatives of the international Data-Enabled Life Science Alliance, July 4, 2013, Moscow, Russia [J]. Acta Nat, 2013, 5 (3):13-16.
- [2] Wadman M. Experts call for active surveillance of drug safety [J]. Nature, 2007, 446(7134):358-359. DOI:10.1038/446358b.

- [3] Groves P, Kayyali B, Knott D, et al. The ‘big data’ revolution in healthcare [R]. McKinsey Quarterly, 2013.
- [4] Khoury MJ, Ioannidis JPA. Big data meets public health [J]. Science, 2014, 346 (6213) : 1054–1055. DOI: 10.1126/science.aaa2709.
- [5] Schneeweiss S, Stürmer T, Maclure M. Case-crossover and case-time-control designs as alternatives in pharmacoepidemiologic research [J]. Pharmacoepidemiol Drug Saf, 1997, 6 Suppl 3: S51–59. DOI: 10.1002/(SICI)1099–1557(199710)6:3+<S51:AID-PDS301>3.0.
- [6] Smeeth L, Donnan PT, Cook DG. The use of primary care databases: case-control and case-only designs [J]. Fam Pract, 2006, 23(5):597–604. DOI: 10.1093/fampra/cml025.
- [7] Wang SV, Gagne JJ, Glynn RJ, et al. Case-crossover studies of therapeutics: design approaches to addressing time-varying prognosis in elderly populations [J]. Epidemiology, 2013, 24(3): 375–378. DOI: 10.1097/EDE.0b013e31828ac9cb.
- [8] Bosco JL, Silliman RA, Thwin SS, et al. A most stubborn bias: no adjustment method fully resolves confounding by indication in observational studies [J]. J Clin Epidemiol, 2010, 63(1) : 64–74. DOI: 10.1016/j.jclinepi.2009.03.001.
- [9] Psaty BM, Larson EB. Investments in infrastructure for diverse research resources and the health of the public [J]. JAMA, 2013, 309(18):1895–1896. DOI: 10.1001/jama.2013.3445.
- [10] Schneeweiss S. A basic study design for expedited safety signal evaluation based on electronic healthcare data [J]. Pharmacoepidemiol Drug Saf, 2010, 19 (8) : 858–868. DOI: 10.1002/pds.1926.
- [11] Farrington CP. Control without separate controls: evaluation of vaccine safety using case-only methods [J]. Vaccine, 2004, 22 (15/16):2064–2070. DOI: 10.1016/j.vaccine.2004.01.017.
- [12] 潘金仁, 何寒青, 严睿, 等. 自身对照的病例系列研究方法在疫苗安全性评价中的应用 [J]. 中华流行病学杂志, 2013, 34(8) : 836–839. DOI: 10.3760/cma.j.issn.0254–6450.2013.08.019.
Pan JR, He HQ, Yan R, et al. Self-controlled case-series (SCCS) method as a tool for the evaluation on the safety of vaccine [J]. Chin J Epidemiol, 2013, 34 (8) : 836–839. DOI: 10.3760/cma.j.issn.0254–6450.2013.08.019.
- [13] 张政, 詹思延. 病例交叉设计 [J]. 中华流行病学杂志, 2001, 22 (4):304–306.
Zhang Z, Zhan SY. The case-crossover study [J]. Chin J Epidemiol, 2001, 22(4):304–306.
- [14] 王涛, 詹思延, 胡永华. 病例–时间–对照设计 [J]. 中华流行病学杂志, 2002, 23(2):142–144.
Wang T, Zhan SY, Hu YH. The case-time-control study [J]. Chin J Epidemiol, 2002, 23(2):142–144.
- [15] 王晶, 卓琳, 詹思延. 病例–病例–时间–对照设计 [J]. 中华流行病学杂志, 2014, 35 (12) : 1396–1399. DOI: 10.3760/cma.j.issn.0254–6450.2014.12.018.
Wang J, Zhuo L, Zhan SY. The case-case-time-control study design [J]. Chin J Epidemiol, 2014, 35 (12) : 1396–1399. DOI: 10.3760/cma.j.issn.0254–6450.2014.12.018.
- [16] Kwong JC, Vasa PP, Campitelli MA, et al. Risk of Guillain-Barre syndrome after seasonal influenza vaccination and influenza health-care encounters: a self-controlled study [J]. Lancet Infect Dis, 2013, 13 (9) : 769–776. DOI: 10.1016/S1473–3099 (13) 70104 –X.
- [17] Lin ST, Chen CC, Tsang HY, et al. Association between antipsychotic use and risk of acute myocardial infarction: a nationwide case-crossover study [J]. Circulation, 2014, 130(3) : 235–243. DOI: 10.1161/CIRCULATIONAHA.114.008779.
- [18] Cohet C, Haguinet F, Dos Santos G, et al. Effect of the adjuvanted (AS03) A/H1N1 2009 pandemic influenza vaccine on the risk of rejection in solid organ transplant recipients in England: a self-controlled case series [J]. BMJ Open, 2016, 6 (1):e009264. DOI: 10.1136/bmjopen–2015–009264.
- [19] Gagne JJ, Nelson JC, Fireman B, et al. Mini-Sentinel methods: Taxonomy for monitoring methods within a medical product safety surveillance system: Report of the Mini-Sentinel Taxonomy Project Work Group [R/OL]. (2012–01–31) [2016–02–22]. http://www.mini-sentinel.org/work_products/statistical_Methods/Mini-Sentinel_Methods_Taxonomy-Year-2-Report.pdf.
- [20] Patadia VK, Coloma P, Schuemie MJ, et al. Using real-world healthcare data for pharmacovigilance signal detection—the experience of the EU-ADR project [J]. Expert Rev Clin Pharmacol, 2015, 8 (1) : 95–102. DOI: 10.1586/17512433.2015.992878.
- [21] Li XC, Hui S, Ryan P, et al. Statistical visualization for assessing performance of methods for safety surveillance using electronic databases [J]. Pharmacoepidemiol Drug Saf, 2013, 22 (5) : 503–509. DOI: 10.1002/pds.3419.
- [22] Coloma PM, Avillach P, Salvo F, et al. A reference standard for evaluation of methods for drug safety signal detection using electronic healthcare record databases [J]. Drug Saf, 2013, 36 (1) : 13–23. DOI: 10.1007/s40264–012–0002–x.
- [23] Carnaham R, Gagne JJ, Nelson J, et al. Mini-Sentinel surveillance plan: Mini-Sentinel prospective routine observational monitoring program tools (PROMPT) : rivaroxaban surveillance [R/OL]. (2014–03–28) [2016–02–22]. http://www.mini-sentinel.org/work_products/Assessments/Mini-Sentinel_PROMPT_Rivaroxaban-Surveillance-Plan.pdf.
- [24] Ryan PB, Stang PE, Overhage JM, et al. A comparison of the empirical performance of methods for a risk identification system [J]. Drug Saf, 2013, 36 Suppl 1: S143–158. DOI: 10.1007/s40264–013–0108–9.
- [25] Ryan PB, Schuemie MJ, Welebob E, et al. Defining a reference set to support methodological research in drug safety [J]. Drug Saf, 2013, 36 Suppl 1: S33–47. DOI: 10.1007/s40264–013–0097–8.
- [26] Boyce RD, Ryan PB, Norén GN, et al. Bridging islands of information to establish an integrated knowledge base of drugs and health outcomes of interest [J]. Drug Saf, 2014, 37 (8) : 557–567. DOI: 10.1007/s40264–014–0189–0.
- [27] Koutkias VG, Jaoulent M. Computational approaches for

- pharmacovigilance signal detection: toward integrated and semantically-enriched frameworks [J]. *Drug Saf*, 2015, 38 (3) : 219–232. DOI: 10.1007/s40264-015-0278-8.
- [28] Madigan D, Stang PE, Berlin JA, et al. A systematic statistical approach to evaluating evidence from observational studies [J]. *Annu Rev Stat Appl*, 2014, 1: 11–39. DOI: 10.1146/annurev-statistics-022513-115645.
- [29] Harpaz R, DuMouchel W, Shah NH, et al. Novel data-mining methodologies for adverse drug event discovery and analysis [J]. *Clin Pharmacol Ther*, 2012, 91 (6) : 1010–1021. DOI: 10.1038/clpt.2012.50.
- [30] McClure DL, Glanz JM, Xu S, et al. Comparison of epidemiologic methods for active surveillance of vaccine safety [J]. *Vaccine*, 2008, 26 (26) : 3341–3345. DOI: 10.1016/j.vaccine.2008.03.074.
- [31] Glanz JM, McClure DL, Xu S, et al. Four different study designs to evaluate vaccine safety were equally validated with contrasting limitations [J]. *J Clin Epidemiol*, 2006, 59 (8) : 808–818. DOI: 10.1016/j.jclinepi.2005.11.012.
- [32] Musonda P. The self-controlled case series method: performance and design in studies of vaccine safety [D]. London: The Open University, 2006.
- [33] Viboud C, Boelle PY, Kelly J, et al. Comparison of the statistical efficiency of case-crossover and case-control designs: application to severe cutaneous adverse reactions [J]. *J Clin Epidemiol*, 2001, 54 (12) : 1218–1227. DOI: 10.1016/S0895-4356(01)00404-8.
- [34] Li LL, Kulldorff M, Russek-Cohen E, et al. Quantifying the impact of time-varying baseline risk adjustment in the self-controlled risk interval design [J]. *Pharmacoepidemiol Drug Saf*, 2015, 24 (12) : 1304–1312. DOI: 10.1002/pds.3885.
- [35] Wang S, Linkletter C, Maclure M, et al. Future cases as present controls to adjust for exposure trend bias in case-only studies [J]. *Epidemiology*, 2011, 22 (4) : 568–574. DOI: 10.1097/EDE.0b013e31821d09cd.
- [36] Douglas IJ, Langham J, Bhaskaran K, et al. Orlistat and the risk of acute liver injury: self controlled case series study in UK Clinical Practice Research Datalink [J]. *BMJ*, 2013, 346: f1936. DOI: 10.1136/bmj.f1936.
- [37] Wei YD, Guo HX, The E, et al. Persistent lipid abnormalities in statin-treated coronary artery disease patients with and without diabetes in China [J]. *Int J Cardiol*, 2015, 182: 469–475. DOI: 10.1016/j.ijcard.2015.01.024.
- [38] Halsey NA, Talaat KR, Greenbaum A, et al. The safety of influenza vaccines in children: An Institute for Vaccine Safety white paper [J]. *Vaccine*, 2015, 33 Suppl 5: F1–67. DOI: 10.1016/j.vaccine.2015.10.080.
- [39] Platt R, Carnahan RM, Brown JS, et al. The U.S. Food and Drug Administration's Mini-Sentinel program: status and direction [J]. *Pharmacoepidemiol Drug Saf*, 2012, 21 Suppl 1: 1–8. DOI: 10.1002/pds.2343.
- [40] Figueiras A, Tato F, Takkouche B, et al. An algorithm for the design of epidemiologic studies applied to drug surveillance [J]. *Eur J Clin Pharmacol*, 1997, 51 (6) : 445–448. DOI: 10.1007/s002280050228.
- [41] Gagne JJ, Baker M, Bykov K, et al. Mini-Sentinel methods: Development of the mini-sentinel taxonomy prompt selection tool: year three report of the mini-sentinel taxonomy project workgroup [R/OL]. (2013–09–30) [2016–02–22]. http://www.mini-sentinel.org/work_products/Statistical_Methods/Mini-Sentinel_Methods_Taxonomy-Year-3-Report.pdf.
- [42] Gagne J, Wang S, Schneeweiss S. Mini-Sentinel prospective routine observational monitoring program tool: cohort matching [R/OL]. [2016–02–22]. http://www.mini-sentinel.org/work_products/Statistical_Methods/Mini-Sentinel_PROMPT_Cohort-Matching-Tool_Technical-Users-Guide.pdf.
- [43] Carnahan R, Gagne JJ, Nelson J, et al. Mini-Sentinel prospective routine observational monitoring program tools (PROMPT): rivaroxaban surveillance plan [R/OL]. [2016–02–22]. http://www.mini-sentinel.org/work_products/Assessments/Mini-Sentinel_PROMPT_Rivaroxaban-Surveillance-Plan.pdf.
- [44] 徐维, 王志勇, 陈薇薇, 等. 从中美纸质病历元数据比较看中国电子病历元数据 [J]. 中国全科医学, 2009, 12 (10) : 925–927. DOI: 10.3969/j.issn.1007-9572.2009.10.045.
- Xu W, Wang ZY, Chen WW, et al. Enlightenment on electronic medical record metadata by comparing traditional paper record metadata between China and America [J]. *Chin Gen Pract*, 2009, 12 (10) : 925–927. DOI: 10.3969/j.issn.1007-9572.2009.10.045.
- [45] 王丹. 药品不良反应主动监测及其发展趋势 [J]. 中国药物警戒, 2015, 12 (10) : 600–602.
- Wang D. Active surveillance of adverse drug reaction and its development trend [J]. *Chin J Pharmacovigil*, 2015, 12 (10) : 600–602.
- [46] 朱田田, 叶小飞, 郭晓晶, 等. 基于电子医疗数据库的药物流行病学安全性研究的特点及要求 [J]. 中国药物警戒, 2014, 11 (7) : 413–415. DOI: 10.3969/j.issn.1672-8629.2014.07.007.
- Zhu TT, Ye XF, Guo XJ, et al. Characteristics and requirements of pharmacoepidemiologic safety studies using electronic healthcare databases [J]. *Chin J Pharmacovigil*, 2014, 11 (7) : 413–415. DOI: 10.3969/j.issn.1672-8629.2014.07.007.
- [47] Chui CSL, Man KKC, Cheng CL, et al. An investigation of the potential association between retinal detachment and oral fluoroquinolones: a self-controlled case series study [J]. *J Antimicrob Chemother*, 2014, 69 (9) : 2563–2567. DOI: 10.1093/jac/dku145.

(收稿日期: 2016-03-01)

(本文编辑: 王岚)