·系统综述/Meta分析·

药物流行病学研究中具有中介作用的时依性混杂及其控制的系统综述

赵厚宇[1] 曾雪扬[1] 刘凤琪[1] 陈思源[1] 詹思延[1,2]

[1]北京大学公共卫生学院流行病与卫生统计学系/中国卫生发展研究中心 100191;[2]北京大学人工智能研究院智慧公众健康研究中心 100871

通信作者:詹思延,Email:siyan-zhan@bjmu.edu.cn

【摘要】 目的 了解药物流行病学研究中时依性混杂控制方法的应用情况。**方法** 系统检索英文数据库PubMed和Embase,中文数据库中国知网和万方数据知识服务平台,纳入发表于2020年6月15日以前且涉及具有中介作用的时依性混杂问题的药物流行病学研究,对其基本特征、药物暴露和结局、时依性混杂及其控制方法等问题进行分析。**结果** 共纳入298篇文献。涉及时依性混杂的药物流行病学研究在近年来明显增长,HIV/AIDS研究领域对该问题关注最多,占全部纳入研究的35.6%(106/298),且其中92个研究涉及抗反转录病毒药物。最常见的研究结局为死亡,而最常关注的时依性混杂因素为实验室检查指标(179,60.1%)、合并症(136,45.6%)和合并用药(108,36.2%)。边际结构模型(MSM)和逆治疗概率加权(IPTW)是最常用的控制时依性混杂的分析方法(244,81.9%)。与合理控制时依性混杂的分析相比,传统方法调整基线混杂引起偏倚的中位数为18.2%(*IQR*:7.4%~40.8%)。此外,分别有28.9%和64.8%的研究对因果推断的阳性假设和无未观测混杂假设进行了检验或讨论。**结论** 目前大多数慢性病药物治疗领域对时依性混杂关注仍然不足。实验室检查、合并症、合并用药等在常规医疗中容易获得的指标是最普遍考虑的时依性混杂。在控制方法上,MSM和IPTW等相对较简单、结果易理解的方法最常用。

【关键词】 药物流行病学; 时依性混杂; 因果模型; G-方法

基金项目:国家自然科学基金(81973146)

Methods for controlling time-varying confounding in pharmaco-epidemiological studies: a systematic reveiw

Zhao Houyu[1], Zeng Xueyang[1], Liu Fengqi[1], Chen Siyuan[1], Zhan Siyan[1,2]

[1] Department of Epidemiology and Biostatistics/China Center for Health Development Studies, School of Public Health Peking University, Beijing 100191, China; [2] Center for Intelligent Public Health, Institute for Artificial Intelligence, Peking University, Beijing 100871, China

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

【Abstract】 Objective To systematically review the application of methods for controlling time-varying confounding in pharmaco-epidemiological studies. **Methods** PubMed, Embase, CNKI, and Wanfang were searched for pharmaco-epidemiological studies involving time-varying confounding on June 15th, 2020. The basic characteristics, drug exposure and outcome, time-varying confounders and the application of methods to control these confounders were analyzed. **Results** A total of 298 articles were included. An increasing trend was observed in numbers of studies dealing with time-varying confounding in pharmaco-epidemiological studies in recent years. A total of 106 (35.6%) studies involved the safety or effectiveness of medication use in HIV/AIDS patients and 92 of them involved antiretroviral drugs. The most common outcome was mortality, while the

DOI:10.3760/cma.j.cn112338-20201016-01240

收稿日期 2020-10-16 **本文编辑** 李银鸽

引用本文:赵厚宇, 曾雪扬, 刘凤琪, 等. 药物流行病学研究中具有中介作用的时依性混杂及其控制的系统综述[J]. 中华流行病学杂志, 2021, 42(12): 2179-2187. DOI: 10.3760/cma.j.cn112338-20201016-01240.

most commonly concerned time-dependent confounders were laboratory examination results (179, 60.1%), comorbidities (136, 45.6%), and co-used medications (108, 36.2%). Marginal structure model (MSM) and inverse probability of treatment weighting (IPTW) were the most commonly used methods to control time-varying confounding factors (244, 81.9%). Compared with the results after properly controlling time-varying confounding, traditional methods adjusting only baseline confounders resulted in substantial bias (median 18.2%, interquartile range, 7.4%-40.8%). As for basic assumptions needed for causal methods controlling time-varying confounding, 28.9% and 64.8% of the included studies examined or discussed the assumptions of positivity and no unmeasured confounders, respectively. **Conclusions** At present, most of the fields of drug therapy for chronic diseases still pay insufficient attention to time-varying confoundings. Information collected in routine medical practice, such as laboratory tests, comorbidities, and co-used drugs, was the most commonly concerned time-varying confounder. MSM and IPTW were the most commonly applied methods for dealing with time-varying confounding.

【**Key words**】 Time-varying confounding; Causal inference; Pharmaco-epidemiology; G-methods

Fund program: National Natural Science Foundation of China (81973146)

基于观察性数据的真实世界研究弥补了随机对照试验(randomized controlled trial,RCT)样本量和观察时间有限,研究人群有严格的入选标准,不能完全反映药品在真实医疗实践中的效果等方面的不足[1-3]。因此,真实世界研究是药品安全性和有效性证据的重要来源,对RCT证据具有重要的补充和验证作用[4-5]。在常规医疗实践中,随着患者病情、用药依从性等因素的改变,药物治疗的中断、加药和换药等随时间变化的情况十分常见[6-7]。同时,在药物流行病学研究中混杂因素常常受既往暴露水平影响而在暴露和结局关联中起到中介作用[8-9]。例如,在比较两种降糖药物与心血管疾病(CVD)风险关联的研究中,降糖药的使用会降低患者的糖化血红蛋白(HbA1c)水平,而HbA1c反映了血糖控制情况,又会影响降糖药种类和剂量的选择,同时HbA1c也被认为与CVD发生风险相关,因而HbA1c在特定时间成为治疗和结局之间关联的中介和混杂因素[8]。类似于HbA1c这样的混杂因素在药物流行病学研究中普遍存在,其满足:①是研究结局的时间依赖性危险因素,同时该因素对后期的治疗或暴露有预测作用;②过去的药物暴露或治疗史影响随后的混杂因素水平。满足条件①的混杂因素被称为时依性混杂(time-varying confounders)[9-10]。对于这类时依性混杂,如果采用传统的分析方法,如条件回归、分层、匹配、限制等,将面临两个严重问题[11]:①消除了先期暴露通过当期混杂中介作用对结局产生的影响,使暴露与结局之间的关联偏向无效假设;②碰撞分层偏倚(collider stratification bias)[8,10]。针对这一问题,Robins及其合作者提出了一组可以控制有中介作用的时依性混杂的方法,即G-方法(generalized methods, G-methods)[8,12-13],包括G-计算公式(G-computation formula)、逆治疗概率加权的边际结构模型(inverse probability of treatment weighting marginal structural model, IPTW-MSM),以及结构嵌套模型和G-估计(G-estimation of structural nested model, G-SNM)。虽然目前已经有数个研究[13-16]对时依性混杂的处理问题进行了综述,但尚未见到对药物流行病学研究中的时依性混杂及其控制方法的系统分析和报告。Clare等[14]的综述关注因果模型的应用而忽视了药物流行病学研究,其对主题的分类中药物相关的研究仅25篇(4.6%)。Suarez等[15]对2000–2009年使用MSM的文献进行了综述,同样没有关注药物流行病学研究。Yang等[16]综述了药物流行病学研究中时依性混杂的控制问题,但该研究仅关注MSM且仅纳入了2012年发表的文献。Pazzagli等[13]对G-方法等用于药物流行病学研究的基本原理进行了介绍,但未系统和定量地分析这些方法的应用现状。鉴于此,本研究旨在对截至目前的涉及时依性混杂的药物流行病学研究的基本特征、发展趋势、方法学质量等问题进行更加全面地系统分析和总结。

研究方法

1. 文献检索:于2020年6月15日检索了PubMed、Embase、中国知网和万方数据知识服务平台4个中、英文数据库。由于本研究所关注的问题与Clare等[14]的研究相似,即具有中介作用的时依性混杂及其控制,我们采用了类似的检索策略。检

索词涉及时依性混杂和G-方法等主要分析方法，例如“time-varying confounder”“time-dependent confounder”“marginal structural model”等。详细检索策略及其他与本研究相关的补充内容见https://www.researchgate.net/publication/356174434_Suppl_112338-20201016-01240。

2. 文献纳入和排除：本研究中纳入满足以下所有条件的文献：①研究问题为某种或某类药物的安全性或有效性分析；②控制了潜在的具有中介作用的时依性混杂；③纯应用研究，或者以药物有效性或安全性评价作为实例阐述方法学基本原理的方法学研究。

排除满足以下任一条件的文献：①纯方法学研究，或者只含有模拟数据的研究；②不涉及药物的安全性或有效性，如药代动力学研究等；③不涉及对比组的研究和非原始研究，包括综述、评论、指南、共识、病例报告等；④会议摘要等研究方法介绍不全的文献；⑤不涉及具有中介作用的时依性混杂问题；⑥除中文和英文之外的其他语言文献；⑦无法获得全文的文献。

文献经过去重、阅读题目和摘要、阅读全文等3个步骤完成纳入排除过程，后2个步骤由研究者分工独立完成并经双人核查。

3. 文献信息定义和提取：提取的信息包括：文献基本特征，包括作者、发表年份、期刊、研究设计类型、数据来源等信息；暴露和结局定义，包括研究药物、对照、结局事件，以及这些变量的定义形式；数据分析，包括分析策略、评估的治疗策略、时依性混杂变量、统计学分析方法、研究的主要结果等；研究方法学质量评价，主要针对因果模型的2个基本假设，即阳性假设（positivity）和无未观察混杂（unmeasured confounding）提取了相关信息。此外，由于既往研究表明，IPTW/MSM方法是最常用的时依性混杂控制方法[14,16]，因此对这类文献，我们还提取了IPTW的类型、估计方法、估计后处理和分布描述等信息。

本研究中，有的文献对某一类信息的报告可能涉及多个类别。对于结局，有的文献同时报告了多个结局的结果，例如死亡和CVD，对这类文献我们最多提取了2个结局。同时，我们提取了全部时依协变量信息，因此同一个研究可能涉及多个类别的时依性混杂变量。此外，部分文献可能使用了多个时依性混杂的控制方法，我们提取了全部这类信息。因此除特殊说明外，本文报告的百分比均以最终纳入的文献数（298篇）作为分母计算。

本研究评估了忽视时依性混杂可能对效应估计造成的影响，并通过如下两个公式评估未控制混杂的“粗分析”结果、仅控制基线混杂的结果与控制时依性混杂结果的差异：

$$Bias_c=|(Eff_{粗}-Eff_{时依})/Eff_{时依}|\times100;$$

$$Bias_b=|(Eff_{基线}-Eff_{时依})/Eff_{时依}|\times100$$

式中$Bias_c$和$Bias_b$分别表示粗分析和仅控制基线混杂时结果的偏倚程度，$Eff_{时依}$、$Eff_{粗}$和$Eff_{基线}$分别为控制时依性混杂、粗分析和仅控制基线混杂时的效应估计结果。

4. 趋势性分析：使用主题词（“Drug-Related Side Effects and Adverse Reactions”[Mesh]）OR（“Comparative Effectiveness Research”[Mesh]）检索了PubMed收录的药物安全性或有效性研究的文献，并将其历年文献数与涉及时依性混杂的文献数进行比较。

本研究使用EpiData 3.1软件提取文献信息，使用SAS 9.4软件完成所有统计学分析。

结　果

1. 文献检索与纳入排除：从4个数据库中共检索到7 149篇文献，剔除重复文献2 416篇。通过阅读题目、摘要和全文，剔除综述、会议摘要、非药物安全性或有效性相关研究等文献，最终纳入298篇（图1）。

2. 纳入文献的基本特征：纳入文献中最早的一

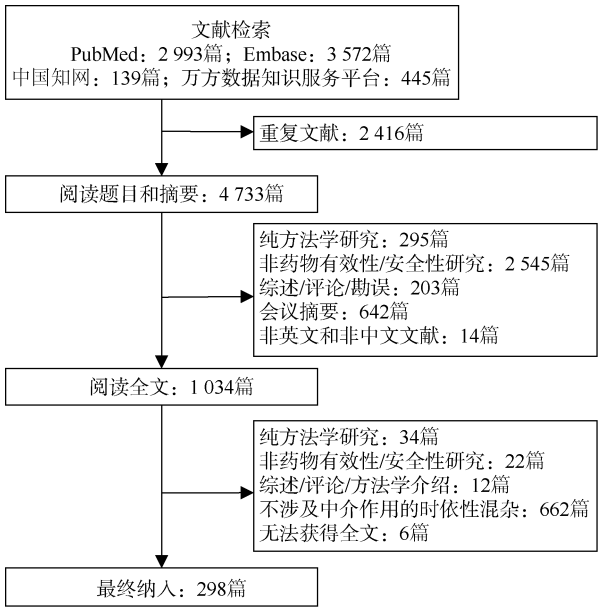


图1　文献纳入排除流程图

篇发表于1992年，介绍了结构嵌套模型和G-估计。相关文献在2007年后出现明显增加，2013年出现波动性下降后再次明显增加，2015年后每年文献维持在30篇左右。同时，涉及时依性混杂的药物流行病学研究占药物安全性或有效性研究的比例呈现明显的增长趋势，但在整个药物流行病学研究中占比仍然很小。

纳入文献中261个（87.6%）为队列研究，34个（11.4%）为RCT数据的二次分析，另有两个病例对照和一个病例队列研究（表1）；239篇（80.2%）文献为药物安全性或有效性的纯应用研究。涉及HIV/AIDS的研究最多，占全部研究的35.6%（106/298）；

表1 纳入文献的基本特征

研究特征	文献数	百分比（%）	研究特征	文献数	百分比（%）
研究设计			文献类型		
队列	261	87.6	纯应用研究	239	80.2
RCT	34	11.4	理论研究/方法学介绍	59	19.8
其他	3	1.0	数据来源[a]		
基础人群			Medicare	20	6.7
HIV/AIDS	106	35.6	VHA	16	5.4
肾病	32	10.7	MACS	15	5.0
糖尿病/代谢性疾病	24	8.1	IeDEA	11	3.7
神经和精神性疾病	23	7.7	HIV-CAUSAL Collaboration	9	3.0
风湿免疫性疾病	20	6.7	WIHS	9	3.0
心脑血管疾病	16	5.4	USRDS	8	2.7
危重症	11	3.7	MarketScan	7	2.3
其他循环系统疾病	9	3.0	CPRD	7	2.3
恶性肿瘤	9	3.0	Medicaid	6	2.0
感染性疾病	8	2.7	HMO	5	1.7
妊娠	8	2.7	MoBa	5	1.7
罕见病	6	2.0	SickKids	4	1.3
COPD	5	1.7	THIN	3	1.0
其他	21	7.1	结局类别		
药物类别			死亡	110	36.9
抗反转录病毒	92	30.9	心血管疾病	42	14.1
他汀类	18	6.0	基础疾病进展	42	14.1
其他心血管系统用药	17	5.7	基础疾病改善	27	9.1
ACEI/ARB	16	5.4	肝肾肺功能	22	7.4
抗风湿和免疫调节药	16	5.4	生化指标改善	20	6.7
抗抑郁药	15	5.0	感染	17	5.7
抗精神病药	14	4.7	AIDS	17	5.7
促红细胞生成素	13	4.4	妊娠、分娩和发育相关疾病	10	3.4
降糖药	12	4.0	精神性疾病	10	3.4
其他抗感染药	12	4.0	恶性肿瘤	10	3.4
镇痛麻醉药	12	4.0	不良反应	6	2.0
抗炎药	10	3.4	伤害和骨折	6	2.0
其他抗病毒药	10	3.4	治疗依从性	5	1.7
抗肿瘤药	8	2.7	其他	7	2.3
激素类避孕药	6	2.0			
其他神经系统用药	27	9.1			

注：IeDEA：International epidemiological Databases to Evaluate AIDS；MACS：Multicenter AIDS Cohort Study；VHA：Veterans Health Administration；USRDS：US Renal Data System；WIHS：Women′s Interagency HIV Study；CPRD：Clinical Practice Research Datalink；SickKids：Hospital for Sick Children；MoBa：Norwegian Mother and Child Cohort Study；HMO：HMO Research Network；THIN：The Health Improvement Network；ACEI：Angiotensin-converting enzyme inhibitor；ARB：Angiotensin Ⅱ receptor blocker；COPD：慢性阻塞性肺病；RCT：随机对照试验；[a]仅列出使用超过2次的数据源

其次为肾病，占 10.7%（32/298）。53 个（17.8%）研究使用了多个数据源，163 个（54.7%）研究使用了电子医疗数据库，其中使用最多的是 Medicare 和退伍军人健康管理数据（Veterans Health Administration，VHA），分别有 20 个（6.7%）和 16 个（5.4%）研究使用这些数据。

在涉及 HIV/AIDS 的研究中，关注抗反转录病毒治疗药物安全性或有效性的研究最多，占全部纳入研究的 30.9%（92/298）。此外，关注较多的药物还包括他汀类、心血管系统用药、ACEI/ARB、抗风湿和免疫调节类药物以及精神性疾病用药。在研究结局方面，110 个（36.9%）研究涉及死亡风险，是关注最多的结局类型，其次为心血管系统疾病，42 个（14.1%）研究关注研究药物的心血管安全性。在样本量方面，观察性研究样本量明显多于 RCT，规模最大的研究涉及超过 160 万的人群。

3. 暴露、对照和结局定义：所有纳入文献中，233 个（78.2%）研究将药物暴露定义为二分类变量，仅有 20 个（6.7%）研究考察了连续型的暴露（表 2）。在对照方面，50.3%（150/298）的研究将未使用暴露组药物的其他研究对象作为对照，而 55 个（18.5%）研究将不同于暴露组的其他药物使用者作为对照。93 个（31.2%）研究考察了同类或同种药物不同剂量、用药策略或用药变化之间的比较。在结局方面，53 个（17.8%）研究考察了药物暴露与复合终点之间的关联，其中 36 个（67.9%）研究的复合终点中包括了死亡。时间-事件和二分类变量是最常用的结局定义形式，分别有 54.4%（162/298）和 30.9%（92/298）的研究定义了这两类结局，13 个（4.4%）研究考察了多个类型的结局与药物暴露的关系。

4. 时依性混杂：在全部 298 个研究中，288 个（96.6%）报告了时依协变量。实验室检查指标、合并症、合并用药等信息是最常见的时依性混杂变量，分别有 179 个（60.1%）、136 个（45.6%）和 108 个（36.2%）研究考察了这些类型的时依性混杂变量（表 2）。而在分析中时依性混杂变量最常用的定义形式为连续变量和二分类变量，分别占 71.8%（214/298）和 64.4%（192/298）。在报告时依协变量的 288 个研究中，平均每个研究考虑了 6.4 个时依性混杂变量，33 个（11.1%）研究仅考虑了一个时依性混杂，56.0%（167/298）的研究中时依性混杂变量不超过 5 个。

5. 分析方法和结果：意向性分析（intention to treat）和实际治疗分析（as treated）是最常用的分析策略，分别有 137 个（46.0%）和 117 个（39.3%）研究使用了这两种分析策略（表 3）。仅 31 个（10.4%）研究考察了动态治疗策略（dynamic treatment strategy），

表 2　纳入研究的暴露、对照、结局定义和时依性混杂变量

研究特征	文献数	百分比（%）
暴露类型		
二分类变量	233	78.2
多分类变量	58	19.5
连续变量	20	6.7
对照类型		
非（研究药）用药者	150	50.3
其他对照	148	49.7
累积暴露	36	12.1
停药 *vs.* 维持用药	9	3.0
换药 *vs.* 维持用药	7	2.4
减药 *vs.* 维持用药	2	0.7
延迟 *vs.* 即时用药	8	2.7
不同用药策略	31	10.4
其他阳性对照	55	18.5
单一或复合结局		
单一事件终点	245	82.2
复合终点	53	17.8
含死亡	36	12.1
不含死亡	17	5.7
结局类型		
时间-事件型	162	54.4
二分类变量	92	30.9
连续变量	46	15.4
计数变量	11	3.7
时依性混杂类别		
实验室检查	179	60.1
合并症	136	45.6
合并用药	108	36.2
体格检查	62	20.8
卫生服务利用	46	15.4
综合评分	45	15.1
症状体征	43	14.4
人口学特征	33	11.1
非药物治疗	30	10.1
生活方式	30	10.1
时间相关性变量	26	8.7
其他	33	11.1
未报告	10	3.4
时依性混杂定义形式		
连续变量	214	71.8
二分类变量	192	64.4
多分类变量	93	31.2
未报告	10	3.4
时依性混杂数量		
1~	167	58.0
6~	71	23.8
>10	50	16.8
未报告	10	3.4

且其中 13 个(4.4%)研究考虑了随机治疗策略(random treatment strategy)。在统计学分析方面,IPTW/MSM 是最常用的分析方法,占 81.9%(244/298)。相比之下,G-公式、结构嵌套模型等其他 G-方法使用率不足 10%。因此,时依性混杂控制方法在药物流行病学研究中应用的趋势主要由 IPTW/MSM 方法决定。此外,在所有纳入研究中,最常报告的效应指标为 *HR* 值(156,52.3%)。

表 3　纳入文献控制时依性混杂的分析策略和统计学分析方法

研究特征	文献数	百分比(%)
分析策略		
ITT	137	46.0
AT	117	39.3
PP	33	11.1
无法判断/未报告	7	2.4
效应测量指标		
HR 值	156	52.3
RR 值	46	15.4
OR 值	42	14.1
RD 值	24	8.1
效应值差	40	13.4
斜率	13	4.4
其他	13	4.4
治疗策略的动态性		
静态治疗策略	267	89.6
动态治疗策略	31	10.4
治疗策略的随机性		
确定性治疗策略	285	95.6
随机治疗策略	13	4.4
统计学分析方法		
IPTW/MSM	244	81.9
G-公式	27	9.1
SNM	21	7.1
DR/TMLE	10	3.4
其他	20	6.7

注:ITT:Intention-to-treat;AT:As-treated;PP:Per-protocol;*HR*:Hazard ratio;*RR*:Risk/Rate ratio;*OR*:Odds ratio;*RD*:Risk/Rate difference;IPTW:Inverse probability of treatment weighting;MSM:Marginal structure model;SNM:Structural nested model;DR:Doubly robust method;TMLE:Targeted maximum likelihood estimation

共有 123 个(41.3%)研究同时报告了不调整混杂和使用因果模型调整时依性混杂的结果,136 个(45.6%)研究同时报告了调整基线混杂和使用因果模型调整时依性混杂的结果。与 MSM 等因果模型的结果相比,不调整混杂引起的偏倚程度中位数为 28.2%(*IQR*:9.6%~61.8%),而调整基线混杂使偏倚程度减小,其中位数为 18.2%(*IQR*:7.4%~40.8%)。两种分析与因果模型的结果相比,分别有 58.5%(72/123)和 43.4%(59/136)的研究偏倚超过 20%,15 个(11.0%)研究中传统方法调整基线混杂而获得的效应值方向与控制时依性混杂结果相反(表 4)。

表 4　粗分析与仅调整基线混杂的传统分析结果的偏倚程度

偏倚程度	时依性分析 *vs.* 单变量分析		时依性分析 *vs.* 基线混杂调整	
	文献数	百分比(%)	文献数	百分比(%)
≤20%	51	41.5	77	56.6
≤40%	26	21.1	24	17.7
≤60%	15	12.2	17	12.5
>60%	31	25.2	18	13.2

6. 边际结构模型与逆概率加权:在 244 个使用 IPTW/MSM 方法的研究中,1 个 MSM 分析中未报告参数估计方法。在剩余的 243 个研究中,204 个(84.0%)报告了 IPTW 估计方法,其中仅 8 个(3.3%)研究使用了机器学习方法估计权重(表 5)。在 193 个(79.4%)使用(广义)线性模型估计权重的研究中,189 个(97.9%)研究使用了 logistic 模型。81.5%(198/243)的研究估计了稳定的权重值,而 104 个(42.8%)研究对权重值进行了截尾或剔除极端权重的研究对象。以权重 99%(1%)分位数作为临界值截尾权重的方法最常见,而在以特定权重值截尾的研究中,以 10 为截尾水平最常见。

7. 因果模型基本假设:所有纳入的研究中,仅 86 个(28.9%)研究明确描述对阳性假设进行检验或讨论,95 个(31.9%)研究描述了 IPTW 的分布(表 5)。193 个(64.8%)研究讨论了未测量混杂的潜在影响,其中仅 21 个(10.9%)研究针对未测量混杂进行了敏感性分析。

讨　论

本研究是目前针对药物流行病学领域时依性混杂处理问题的最新和最全面的系统综述。近年来药物安全性和有效性研究对时依性混杂的关注明显上升,但是这些研究仍然主要集中于 HIV/AIDS 相关领域,慢性非传染性疾病用药的研究中对时依性混杂问题的关注仍然相对不足。同时,原理上相对简单但稳健性相对不足的 MSM 和 IPTW 方法[14]仍然是药物流行病学研究处理时依性混杂问题的主流方法。而现有药物流行病学研究中考

表 5 使用IPTW/MSM方法的研究中IPTW估计特征

IPTW估计特征	文献数	百分比(%)
IPTW估计方法		
(广义)线性模型	193	79.4
机器学习	8	3.3
其他	3	1.2
未报告	39	16.1
IPTW类型		
稳定的	198	81.5
不稳定的	45	18.5
IPTW分布[a]		
未报告	148	60.9
报告	95	39.1
均值	66	27.2
分位数	68	28.0
分布图	17	7.0
IPTW估计后处理[b]		
截尾	101	41.6
剔除	3	1.2
标准化	4	1.6
未做特殊处理	192	79.0
截尾/剔除权重的百分位[c]		
<99%	10	9.7
99%	52	50.5
>99%	16	15.5
截尾/剔除权重的权重值[c]		
<10	2	1.9
10	11	10.7
>10	12	11.7

注：一个使用MSM的研究未报告使用了IPTW估计，未纳入此部分分析；[a]54个研究报告了多个分布特征；[b]两个研究同时使用了标准化和截尾；[c]一个研究使用适应性截尾，未报告截尾水平；IPTW：Inverse probability of treatment weighting；MSM：Marginal structure model

察最多的时依性混杂因素主要为实验室检查指标、合并症和合并用药等医疗实践中常规收集的临床数据。

在大数据时代，能否利用医疗保险、电子病历等真实世界数据产生真实世界证据已经成为监管机构必须考虑的问题[17]。真实世界证据能否用于药品的有效性和安全性评价，关键取决于药物流行病学研究能否控制各种潜在偏倚以获得真实、无偏的关联估计[18-19]。我们的结果表明，当存在时依性混杂时，不控制任何混杂效应平均会造成68%的偏倚，而仅控制基线混杂的传统分析方法对结果的偏倚控制有限，仍有43%的研究偏倚超过20%，且有11%的研究可能完全颠倒了药物的真实疗效或安全性。这与此前仅针对MSM的综述结果一致[15-16]。这表明，在药物流行病学研究中，时依性混杂对结果的潜在影响不可忽视。然而，目前关注和控制时依性混杂的药物流行病学研究主要集中于某些特定的领域，例如HIV/AIDS、肾脏疾病和心血管疾病等，而其他领域的研究对这一问题的关注和相关方法的应用却十分有限。这可能与MSM最早且大量用于探索HIV/AIDS药物治疗的有效性和安全性问题有关。

本综述还发现，药物流行病学研究中考察最多的时依性混杂变量是实验室检查指标、合并症和合并用药，这些也是医疗实践中常规收集的患者信息，这与近年来基于电子医疗数据库的药物流行病学研究显著增加有关[19]。随着临床实践中常规收集和电子存储的医疗信息的激增，基于大型电子医疗数据的观察性研究逐渐成为药物流行病学研究的主流，特别是在北美和欧洲等地区[5,20]。电子健康记录等电子医疗数据包含常规医疗过程中获取的各类纵向信息[21]，相对于RCT和专门的队列研究，基于这种观察性医疗数据开展研究具有明显的优势：耗时短、成本低，可以反映药物在实际临床应用中的效果和安全性；规模大、时间长，可以回答药物的罕见不良事件风险等问题[5,20]。而且电子健康记录数据中药物处方等数据记录了用药起始和改变的时间，同时其数据量大而保证亚组分析具有较大的样本量，因而允许更精细地研究治疗时序对结局的影响[22]。然而电子医疗数据也有其自身局限性，例如无法获得非处方用药信息；某些数据库，特别是管理型数据库，缺乏特定的潜在混杂变量，例如关于疾病严重程度的信息经常缺乏，因此可能无法排除疾病严重程度等因素造成的混杂[20,23]。现有研究大量应用实验室指标、共患疾病等信息可能是在数据有限性和混杂控制的充分性之间的妥协，这些药物流行病学研究也表明控制这些潜在的时依性混杂具有可行性和必要性。

MSM和IPTW是药物流行病学研究中控制时依性混杂最常用的分析方法，这可能是因为相比于其他方法，IPTW/MSM更易于理解、结果易于解释且更容易通过常用分析软件实现[24]。但该方法稳健性较差，对模型设定错误较敏感[24-25]。当混杂因素对暴露有高度的预测性，或当暴露是连续的或水平较多时，估计结果容易出现较大的偏倚和方差[24]。近年来，一种被称为超级学习(super learning，SL)的集成学习方法被提出用于估计倾向性评分(PS)和逆概率权重，以避免依赖于单一的估

计方法而引起模型设定错误[26]。与传统的PS估计方法相比，基于SL的PS估计具有数据自适应性且不依赖数据分布假设，可以有效地减少偏倚，提高估计效率[26-27]。然而，在使用IPTW/MSM的文献中，仅有3%的研究使用了SL方法，这表明现有考虑时依性混杂的药物流行病学研究在分析方法的选择上较为单一，而SL等更稳健的方法应用于药物流行病学研究的效果还需要更多的实证研究加以验证和推广。此外，其他G-方法，如G-公式和结构嵌套模型在药物流行病学研究中应用较少。与MSM相比，结构嵌套模型的优势在于能够估计时依性混杂变量的效应修饰作用，而传统的MSM只能估计基线协变量的效应修饰作用。同时结构嵌套模型对未测量混杂等假设的违背也更加稳健[28]。而在动态治疗策略的比较研究中，参数G-公式通常能获得比IPTW/MSM更加精确的效应估计[4]。此外，近年来一些新的方法被提出和应用于药物流行病学研究，例如具有双稳健性的目标极大似然估计[29-30]、贝叶斯边际结构模型[3]、半参数嵌套累积生存时间模型(SNCSTMs)[24]、结构嵌套限制平均时间损失模型(SNRMTLMs)[31]等，然而由于这些方法理论较复杂，其在药物流行病学研究中的应用明显不如IPTW/MSM[14]。

时依性混杂的有效控制依赖于3个重要的基本假设，即一致性假设、阳性假设和无未观测混杂。一致性假设通常认为其成立而很少有研究对其进行检验和讨论[16]，然而违反阳性假设和无未观测混杂的情况却并不罕见。本研究纳入文献中仅有很少一部分研究明确描述了对阳性假设进行检验或讨论。IPTW方法通常对阳性假设的违背更为敏感，这可能造成估计结果的严重偏倚[32]。然而在使用IPTW控制时依性混杂的药物流行病学研究中，仅有不到40%的研究对权重分布进行了描述。稳定权重和截尾权重是避免违背阳性假设的常用方法[33]，但截尾权重并不是一种理想的模型检验过程，因为极端权重通常包含了大量混杂效应，而截尾权重通常会使结果偏向于非加权估计值[34]。因此，采用不同的截尾水平进行多组敏感性分析对于评估分析结果的可靠性十分重要[35]。Santacatterina等[33]提出了最优概率权重(optimal probability weights，OPW)，即与原始权重之间有特定方差且最小欧式距离的概率权重值，来代替权重截尾，其模拟和实证研究表明，OPW的分布比截尾权重更窄，并且在所有违反阳性假设的情况下，OPW的表现都优于截尾权重[33]。此外，通常用于药物流行病学研究的大型数据库由于并非为特定研究目的而建立，通常会缺乏一些关键的临床信息，因而可能遗漏一些重要的混杂变量[23]。本研究中约有60%的研究考察的时依性混杂变量≤5个，部分研究可能因时依性混杂控制不充分而面临偏倚风险[16,23]。因此，当这种偏倚风险存在时，针对未测量混杂进行敏感性分析[36-37]，或者利用混杂测量更充分的验证数据对结果进行校正[23]等方法有助于评估估计结果的稳健性。

相对于既往研究，本研究对涉及时依性混杂的药物流行病学研究的人群、暴露、结局、分析方法等特征进行了更加全面地分析和总结。有利于领域内研究者全面充分地了解时依性混杂及其控制在药物流行学研究中的应用现状和发展趋势。本研究有一定的局限性。首先，我们只检索了两个英文数据库PubMed和Embase，未被两个数据库收录的药物流行病学研究可能被遗漏。其次，在检索中文数据库时，我们使用了英文概念的直译词，由于不同的研究者对时依性混杂及其控制方法理解上的不同，可能在发表文章中使用了不同的中文概念(例如，本文中的“时依性混杂”在其他文献中被称为“时变混杂”“依时混杂”等)，尽管我们尝试将同一个英文概念翻译为意思相近的不同中文词(例如上述例子的时依性混杂的不同翻译)，仍然有可能遗漏部分研究。此外，近年来发展的新方法可能不在我们的检索范围内，因而可能遗漏处理时依性混杂的其他方法的研究文献。

综上所述，近年来，关注和控制时依性混杂的药物流行病学研究呈现增长趋势，但慢性非传染性疾病长期用药的有效性和安全性研究对时依性混杂关注仍然不足。实验室检查、合并症、合并用药等在常规医疗中容易获得的指标也是最普遍关注的时依性混杂，这表明基于电子医疗数据库的药物流行病学研究控制时依性混杂可行且有必要。在控制方法上，MSM和IPTW等相对较简单、结果易理解但稳健性较差的方法最常用。未来需要更多的研究探索和促进更加稳健的分析方法在药物流行病学研究中的应用，以获得更加可靠的药物安全性和有效性证据。

利益冲突 所有作者均声明不存在利益冲突

参 考 文 献

[1] Agency for Healthcare Research and Quality, US. Methods guide for effectiveness and comparative effectiveness

reviews[EB/OL]. (2015-02-26) [2019-02-20]. https://effectivehealthcare. ahrq. gov/topics/cer-methods-guide/overview.
[2] Frakt AB. An observational study goes where randomized clinical trials have not[J]. JAMA, 2015, 313(11): 1091-1092. DOI:10.1001/jama.2015.0544.
[3] Liu K, Saarela O, Feldman BM, et al. Estimation of causal effects with repeatedly measured outcomes in a Bayesian framework[J]. Stat Methods Med Res, 2020, 29(9): 2507-2519. DOI:10.1177/0962280219900362.
[4] Young JG, Cain LE, Robins JM, et al. Comparative effectiveness of dynamic treatment regimes: an application of the parametric g-formula[J]. Stat Biosci, 2011, 3(1):119-143. DOI:10.1007/s12561-011-9040-7.
[5] Takahashi Y, Nishida Y, Asai S. Utilization of health care databases for pharmacoepidemiology[J]. Eur J Clin Pharmacol, 2012, 68(2): 123-129. DOI: 10.1007/s00228-011-1088-2.
[6] Faries D, Ascher-Svanum H, Belger M. Analysis of treatment effectiveness in longitudinal observational data [J]. J Biopharm Stat, 2007, 17(5): 809-826. DOI: 10.1080/10543400701513967.
[7] Latimer NR, Abrams KR, Lambert PC, et al. Adjusting for treatment switching in randomised controlled trials−a simulation study and a simplified two-stage method[J]. Stat Methods Med Res, 2017, 26(2): 724-751. DOI: 10.1177/0962280214557578.
[8] Daniel RM, Cousens SN, de Stavola BL, et al. Methods for dealing with time-dependent confounding[J]. Stat Med, 2013, 32(9):1584-1618. DOI:10.1002/sim.5686.
[9] Hernán MA, Cole SR, Margolick J, et al. Structural accelerated failure time models for survival analysis in studies with time-varying treatments[J]. Pharmacoepidemiol Drug Saf, 2005, 14(7):477-491. DOI:10.1002/pds.1064.
[10] Hernán MÁ, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men[J]. Epidemiology, 2000, 11(5): 561-570. DOI: 10.1097/00001648-200009000-00012.
[11] Hernán MA, Brumback BA, Robins JM. Estimating the causal effect of zidovudine on CD4 count with a marginal structural model for repeated measures[J]. Stat Med, 2002, 21(12):1689-1709. DOI:10.1002/sim.1144.
[12] Naimi AI, Cole SR, Kennedy EH. An introduction to G methods[J]. Int J Epidemiol, 2017, 46(2): 756-762. DOI: 10.1093/IJE/DYW323.
[13] Pazzagli L, Linder M, Zhang ML, et al. Methods for time-varying exposure related problems in pharmacoepidemiology: an overview[J]. Pharmacoepidemiol Drug Saf, 2018, 27(2):148-160. DOI:10.1002/pds.4372.
[14] Clare PJ, Dobbins TA, Mattick RP. Causal models adjusting for time-varying confounding−a systematic review of the literature[J]. Int J Epidemiol, 2018, 48(1): 254-265. DOI: 10.1093/ije/dyy218.
[15] Suarez D, Borràs R, Basagaña X. Differences between marginal structural models and conventional models in their exposure effect estimates: a systematic review[J]. Epidemiology, 2011, 22(4): 586-588. DOI: 10.1097/EDE.0b013e31821 d0507.
[16] Yang SB, Eaton CB, Lu J, et al. Application of marginal structural models in pharmacoepidemiologic studies: a systematic review[J]. Pharmacoepidemiol Drug Saf, 2014, 23(6):560-571. DOI:10.1002/pds.3569.
[17] Sherman RE, Anderson SA, Dal Pan GJ, et al. Real-world evidence−what is it and what can it tell us? [J]. N Engl J Med, 2016, 375(23): 2293-2297. DOI: 10.1056/NEJMsb1609216.
[18] US Food and Drug Administration. Framework for FDA's real-world evidence program[EB/OL]. (2018-12-07) [2019-02-18]. https://www. fda. gov/media/120060/download.
[19] Hall GC, Sauer B, Bourke A, et al. Guidelines for good database selection and use in pharmacoepidemiology research[J]. Pharmacoepidemiol Drug Saf, 2012, 21(1): 1-10. DOI:10.1002/pds.2229.
[20] Torre C, Martins A P. Overview of pharmacoepidemiological databases in the assessment of medicines under real-life conditions[M]//Lunet N. Epidemiology-Current Perspectives on Research and Practice. Rijeka: InTechOpen, 2012.
[21] Johnson KW, Glicksberg BS, Hodos RA, et al. Causal inference on electronic health records to assess blood pressure treatment targets: an application of the parametric g-formula[J]. Pac Symp Biocomput, 2018, 23: 180-191. DOI:10.1142/9789813235533_0017.
[22] Hu LY, Hogan JW, Mwangi AW, et al. Modeling the causal effect of treatment initiation time on survival: application to HIV/TB co-infection[J]. Biometrics, 2018, 74(2): 703-713. DOI:10.1111/biom.12780.
[23] Burne RM, Abrahamowicz M. Adjustment for time-dependent unmeasured confounders in marginal structural Cox models using validation sample data[J]. Stat Methods Med Res, 2019, 28(2): 357-371. DOI: 10.1177/0962280217726800.
[24] Seaman S, Dukes O, Keogh R, et al. Adjusting for time-varying confounders in survival analysis using structural nested cumulative survival time models[J]. Biometrics, 2020, 76(2): 472-483. DOI: 10.1111/biom.13158.
[25] Karim ME, Platt RW. Estimating inverse probability weights using super learner when weight-model specification is unknown in a marginal structural Cox model context[J]. Stat Med, 2017, 36(13):2032-2047. DOI: 10.1002/sim.7266.
[26] Neugebauer R, Schmittdiel JA, van der Laan MJ. A case study of the impact of data-adaptive versus model-based estimation of the propensity scores on causal inferences from three inverse probability weighting estimators[J]. Int J Biostat, 2016, 12(1): 131-155. DOI: 10.1515/ijb-2015-0028.
[27] Gruber S, Logan RW, Jarrín I, et al. Ensemble learning of inverse probability weights for marginal structural modeling in large observational datasets[J]. Stat Med, 2015, 34(1):106-117. DOI:10.1002/sim.6322.
[28] He JW, Stephens-Shields A, Joffe M. Structural nested mean models to estimate the effects of time-varying treatments on clustered outcomes[J]. Int J Biostat, 2015, 11(2):203-222. DOI:10.1515/IJB-2014-0055.
[29] Schnitzer ME, Sango J, Guerra SF, et al. Data-adaptive longitudinal model selection in causal inference with collaborative targeted minimum loss-based estimation[J]. Biometrics, 2020, 76(1): 145-157. DOI: 10.1111/biom.13135.
[30] Schomaker M, Luque-Fernandez MA, Leroy V, et al. Using longitudinal targeted maximum likelihood estimation in complex settings with dynamic interventions[J]. Stat Med, 2019, 38(24):4888-4911. DOI:10.1002/sim.8340.
[31] Hagiwara Y, Shinozaki T, Matsuyama Y. G-estimation of structural nested restricted mean time lost models to estimate effects of time-varying treatments on a failure time outcome[J]. Biometrics, 2019, 76(3): 799-810. DOI: 10.1111/biom.13200.
[32] Platt RW, Delaney JA, Suissa S. The positivity assumption and marginal structural models: the example of warfarin use and risk of bleeding[J]. Eur J Epidemiol, 2012, 27(2): 77-83. DOI:10.1007/s10654-011-9637-7.
[33] Santacatterina M, García-Pareja C, Bellocco R, et al. Optimal probability weights for estimating causal effects of time-varying treatments with marginal structural Cox models[J]. Stat Med, 2019, 38(10): 1891-1902. DOI: 10.1002/sim.8080.
[34] Cole SR, Hernán MA, Margolick JB, et al. Marginal structural models for estimating the effect of highly active antiretroviral therapy initiation on CD4 cell count[J]. Am J Epidemiol, 2005, 162(5): 471-478. DOI: 10.1093/aje/kwi216.
[35] Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models[J]. Am J Epidemiol, 2008, 168(6): 656-664. DOI: 10.1093/aje/kwn164.
[36] Yang S, Lok JJ. Sensitivity analysis for unmeasured confounding in coarse structural nested mean models[J]. Stat Sin, 2018, 28(4): 1703-1723. DOI: 10.5705/ss. 202016.0133.
[37] Brumback BA, Hernán MA, Haneuse SJPA, et al. Sensitivity analyses for unmeasured confounding assuming a marginal structural model for repeated measures[J]. Stat Med, 2004, 23(5):749-767. DOI:10.1002/sim.1657.