

观察性研究中针对未测量混杂干扰的敏感性分析方法

王丹华¹ 尤东方^{1,2} 黄丽红³ 赵杨^{1,4,5,6}

¹南京医科大学公共卫生学院生物统计学系 211166; ²南京医科大学现代毒理学教育部重点实验室 211166; ³复旦大学附属中山医院生物统计室, 上海 200032; ⁴江苏省恶性肿瘤生物标志物与防治重点实验室, 南京 211166; ⁵肿瘤个体化医学协同创新中心, 南京 211166; ⁶南京医科大学生物医学大数据重点实验室 211166

通信作者: 赵杨, Email: zhaoyang@njmu.edu.cn

【摘要】 目的 介绍敏感性分析方法, 并对不同方法进行探讨和比较。方法 通过模拟试验和实例比较混杂函数敏感性分析法和边界因子敏感性分析方法在观察性研究中校正未测量混杂因素准确性的差异。结果 模拟试验与实际例子研究结果均显示, 当暴露(X)与结局(Y)之间存在未测量混杂情况下, 混杂函数法和边界因子相比, 在分析未测量混杂因素的效应至少达到多大强度才能导致观测效应值大小和方向彻底改变的问题上, 混杂函数和边界因子分析结果相似。但混杂函数法在完全解释观测效应值时所需的混杂效应强度小于边界因子做出同样解释所需的混杂效应值。边界因子分析中设置两个参数, 而混杂函数中只有一个参数, 混杂函数法在分析计算过程中较边界因子法简便灵敏。结论 对于真实世界观察性研究数据, 分析暴露(X)与结局(Y)之间的因果效应时, 敏感性分析过程必不可少, 从计算过程和结果解释上, 混杂函数敏感性分析方法是一个值得推荐的方法。

【关键词】 观察性研究; 因果推断; 未测量混杂因素; 敏感性分析

基金项目: 国家自然科学基金(81872709, 81903407); 江苏省高等学校自然科学研究重大项目(18KJA110004); 江苏省青蓝工程学科带头人; 南京医科大学中青年学术带头人项目; 江苏省预防医学优势学科; 江苏省社会发展项目(BE2017749)

DOI: 10.3760/cma.j.issn.0254-6450.2019.11.023

Sensitivity analysis method for unmeasured confounding interference in observational study

Wang Danhua¹, You Dongfang^{1,2}, Huang Lihong³, Zhao Yang^{1,4,5,6}

¹Department of Biostatistics, School of Public Health, Nanjing Medical University, Nanjing 211166, China; ²Key Laboratory of Modern Toxicology, Ministry of Education, Nanjing Medical University, Nanjing 211166, China; ³Department of Biostatistics, Zhongshan Hospital, Fudan University, Shanghai 200032, China; ⁴Jiangsu Provincial Key Laboratory of Malignant Tumor Biomarkers and Prevention, Nanjing 211166, China; ⁵Collaborative Innovation Center for Individual Medicine in Cancer, Nanjing 211166, China; ⁶Key Laboratory of Biomedical Big Data, Nanjing Medical University, Nanjing 211166, China
Corresponding author: Zhao Yang, Email: zhaoyang@njmu.edu.cn

【Abstract】 Objective To introduce the methods for sensitivity analysis, discuss and compare the advantages and disadvantages of different methods. **Methods** The difference between confounding function method and bounding factor method in accuracy of identifying unmeasured confounding factors in observational studies through simulation trials and actual clinical data was compared. **Results** The results of simulation trials and actual clinical data showed that when there was unmeasured confounding between exposure (X) and outcome (Y), the results of confounding function and the bounding factor analysis were similar in terms of the effect of unmeasured confounding factor to lead to the complete change of the magnitude and direction of the observed effect value. However, the confounding function method needed smaller confounding effect to fully interpret the observed effect value than the bounding factor needed. In addition, the bounding factor method needed to analyze two confounding parameters, while only one parameter was needed in the confounding function method. The confounding function method was simpler and more sensitive than the bounding factor method. **Conclusion** For real-world observational data, the sensitivity analysis process is essential in analyzing the causal effects between exposure (X) and outcome (Y). In terms of

the calculation process and result interpretation the sensitivity analysis method of confounding function is worth to recommend.

【Key words】 Observational study; Causal inference; Unmeasured confounding factor; Sensitivity analysis

Fund programs: National Natural Science Foundation of China (81872709, 81903407); Key Project of University Natural Scientific Research of Jiangsu Province (18KJA110004); Qing-lan Project of Jiangsu Province; Excellent Young Faculty Program of Nanjing Medical University; Priority Academic Program Development of Jiangsu Higher Education Institution; Social Development Program of Jiangsu (BE2017749)

DOI:10.3760/cma.j.issn.0254-6450.2019.11.023

临床研究中随机对照试验(randomized controlled trial, RCT)采用随机化和盲法等原则,尽量避免试验过程中可能出现的偏倚,被认为是评价干预措施的金标准。但RCT研究受诸多因素限制,属于理想条件下疾病亚人群干预措施的效果^[1]。近年来,真实世界研究(real world study)在真实临床环境下评估治疗措施对患者健康的影响,为RCT研究结果提供有力的补充,得到了研究者的关注^[2]。观察性研究是真实世界研究中一种常见的手段,成为医学研究设计的一个重要类型。

然而,观察性研究中受试对象处理分配往往不是随机的,会存在某些重要未测量混杂(unmeasured confounder)在不同处理组间分布不均导致效应估计受到影响,从而无法准确评价干预措施对结局的作用^[3]。Juurink等^[4]研究显示服用SSRIs类药物的患者自杀率较高可能是由于患者自身抑郁症导致自杀而并不是由于药物本身的危害所致,该研究中因果推断结论的准确性取决于研究者能否准确校正未测量混杂因素。在真实世界研究中疾病严重程度、身体功能或认知功能等因素在研究过程中很难或无法测量获得,从而对研究结果造成偏倚^[5]。

针对这类问题,Rosenbaum和Rubin^[6]提出依据敏感性分析思路评估此类未测量混杂因素对结果效应值造成的影响。通过敏感性分析,研究者试图回答“未测量混杂是否能全部或部分解释研究中观察到的暴露因素的效应”。Vanderweele和Arah^[7]在潜在结果(potential outcome)框架基础上推导出敏感性分析的一般通用公式,可用于多种数据类型。2016年,Ding和Vanderweele^[8]提出边界因子(bounding factor)法,用其衡量因混杂因素导致暴露与疾病之间效应值偏倚大小。2017年,Kasza等^[9]提出“混杂函数(confounding function)”敏感性分析方法。无论是边界因子法或混杂函数法,与其他方法相比,特点是不对未测量混杂做任何具体假设,且其分析目的不仅仅是针对未测量混杂因素而是考虑到所有混杂因素可能对结果造成的影响。

本文主要介绍混杂函数和边界因子敏感性分析法的特点和性质,在观察性研究中校正未测量混杂因素对因果效应值稳健性的影响,并通过一个真实案例展示2种方法各自的特点。

原理与方法

1. 混杂函数法:因果推断结论准确性依赖于多个假设的成立,可交换性(exchangeability)假设是指暴露总体的潜在结果 Y^1 和 Y^0 的分布与非暴露总体的潜在结果的分布相同^[10]。混杂函数法是以因果推断的潜在结果模型为基础构建的函数^[11]。根据可交换性假设可知,当2个总体中不存在混杂因素时, $P(Y^1 | X=1) = P(Y^1 | X=0)$ 和 $P(Y^0 | X=1) = P(Y^0 | X=0)$ 等式成立。但观察性研究中,由于存在未测量混杂因素使可交换性假设很难成立,因此,定义混杂函数为:

$$c(x) = \frac{P(Y^x | X=1)}{P(Y^x | X=0)}, x \in \{0, 1\}$$

当研究中不存在未测量混杂时,可交换性假设成立, $c(1) = c(0) = 1$ 。在应用混杂函数法时,研究者应当结合专业知识,考察在专业知识认可的 $c(0)$ 和 $c(1)$ 下,根据研究数据所估计的效应是否在针对未测量混杂进行调整后依然存在。

2. 边界因子法:Ding和Vanderweele^[8]在2016年提出一种边界因子方法。以RR为例,令 U 表示未测量混杂因素(可以为任意尺度), RR_{XY}^{obs} 表示暴露因素 X 与结局变量 Y 之间观测到的效应值, RR_{XY}^{true} 表示校正已知和未测量混杂后的暴露因素 X 与结局变量 Y 之间真实效应大小, RR_{XU} 是暴露因素 X 与未测量混杂 U 之间效应值, RR_{UY} 是未测量混杂因素 U 与结局变量 Y 之间效应值,有

$$RR_{XY}^{true} \geq RR_{XY}^{obs} / \frac{RR_{XU} \times RR_{UY}}{RR_{XU} + RR_{UY} - 1}$$

$\frac{RR_{XU} \times RR_{UY}}{RR_{XU} + RR_{UY} - 1}$ 为联合边界因子(joint bounding factor)。上述不等式表明,即便未测量混杂存在时,真实的RR值不会比等式右边小。

3. 模拟试验:笔者通过模拟试验比较混杂函数法和边界因子法在因果效应估计中敏感性分析值敏感性校正方面的表现。试验步骤:

(1)生成自变量:随机产生服从二项分布 $B(1\ 000, 0.5)$ 的处理因素 X 和已知的混杂变量 C_1 和 C_2 , 模拟产生的处理因素 X 与混杂因素 C_1 和 C_2 之间的正向相关大小分别为 0.6 和 0.5。另外产生 10 个噪音变量 $A_1 \sim A_{10}$ [均服从 $U(0.05, 0.7)$]。

(2)生成结局变量:根据模型 $\text{logit}(Y) = -2 + \ln(4)X + \ln(5) \times C_1 + \ln(3) \times C_2$ 模拟产生二分类结局变量 Y , 试验样本总例数 1 000 例。

(3)结果评价指标:用上述 2 种校正方法分别对效应估计值进行敏感性分析,比较各自混杂参数取值不同时所解释的因果效应大小。

(4)统计学软件:本研究统计分析通过 SAS 9.4 软件以及 R 软件“boot”语言包计算效应值及 95%CI, 使用“bindata”语言包产生模拟数据。

4. 实例分析:估计危重病例进行右心导管插入术(right heart catheterization, RHC)治疗对结局影响,数据来源于 SUPPORT 研究^[12]。该研究主要考察 ICU 第 1 个 24 h 内 RHC 的使用与预后、留院时间、医疗费用的关系。从主要结局变量中选择 30 d 的生存结局变量作为本研究的结局变量。

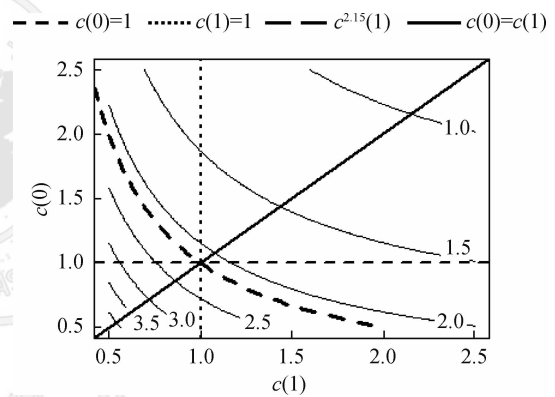
尽管原研究中,研究者考虑了多个潜在混杂因素并利用倾向性评分进行了匹配。由于受试者是患有 COPD、急性呼吸衰竭等 9 种重症病例,是否接受 RHC 并非随机分组所决定,医生对病情的认识及经验决定起了非常重要的作用,因此不能排除研究中还存在其他可能的未知混杂因素。故在原分析的基础上,采用敏感性分析思路考察 RHC 增加死亡风险这一结论的稳健性。

结 果

1. 模拟试验结果:

(1)混杂函数法:在非随机对照研究中,处理组与对照组由于某些特征分布不同,导致研究对象是否接受处理因素的概率受到其他混杂变量的影响。倾向性评分(propensity score)是反映观察到的混杂变量在两组间均衡性的一个近似函数^[13]。利用倾向性评分估计的权重对每个观察对象加权产生一个虚拟的标准人群,在该人群中,两组混杂因素趋于一致^[14]。逆处理概率加权法(inverse probability of treatment weighting)是以所有观察对象(处理组与对照组合并的人群)为“标准人群”进行调整^[15]。

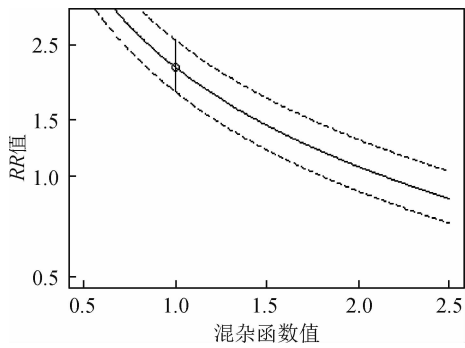
本模拟试验中,经倾向性评分调整已知混杂后求出暴露因素 X 与结局变量 Y 之间观测的效应估计值 $RR=2.15(95\%CI: 1.82 \sim 2.60)$, 说明模拟试验中暴露因素对结局变量是危险因素。混杂参数 $c(0)$ 和 $c(1)$ 不同取值情况下真实效应值 RR 等高曲线图见图 1。水平虚线表示 $c(0)=1$; 垂直虚线表示 $c(1)=1$; 实线表示 $c(0)=c(1)$; 从 $c(0)=c(1)$ 对应的实线可以得到,随着混杂参数值逐渐增大,真实效应值逐渐减小。当此模拟试验中不存在未测量混杂,混杂参数 $c(0)=c(1)=1$ 时的真实效应值为 2.15。假设暴露与结局事件间的真实效应值为 1, 根据公式可知混杂参数至少达到 $c(0)=c(1)=\frac{P(Y=1|X=1)}{P(Y=1|X=0)}=2.15$, 才能认为研究中由于没有充分调整某些未测量混杂变量,使得原本并没有任何显著性效应的处理因素被错误的解释为对结局产生有害效应。



注:当 $c(0)=c(1)=1$ 时,未校正未测量混杂的效应值 $RR=2.15$; $c^{2.15}(1)$ 标注的线表示当效应值为 2.15 时的 $c(0)$ 和 $c(1)$ 的各自取值情况

图 1 模拟试验数据结果效应值敏感性分析等高曲线图

由于模拟中设置了混杂 C_1 和 C_2 且在分析中将混杂因素 C_2 作为未测量处理,此模拟研究中的可交换性假设显然不成立。由于在模拟中, X 与混杂存在正向关系,使得在反事实情况下,若所有人都不接受处理,接受处理的患者比未接受处理的患者更有可能发生结局事件, $c(0) > 1$; 若所有人都接受处理,与事实相反的未接受处理的患者同样比实际接受处理组发生结局风险低, $c(1) > 1$ 。根据图 1 可知,在 $c(0)$ 和 $c(1)$ 均 > 1 的范围内,当真实效应值 $RR=1$ 时,其对应的混杂参数值为 2.15。由于在模拟中假设混杂因素 C_2 与结局 Y 之间的效应值为 $\ln(3) \approx 1.10$, 所以在充分的倾向性评分校正所有可能对结局产生影响的协变量基础上,不可能存在如此大小能改变结果的重要未测量混杂变量。图 2 为 $c(0)=$



注:横坐标代表 $c(0)=c(1)$ 的不同混杂函数值,纵坐标代表真实效应值,竖直的短实线表示 $c(0)=c(1)=1$,实线表示校正不同混杂参数取值后对应的效应值 RR ,虚线表示校正未测量混杂因素后 RR 值的 95%CI,圆圈表示 $c(0)=c(1)=1$ 时的 RR 值

图2 模拟试验数据效应值 RR 的敏感性分析

$c(1)$ 时校正后的效应值 RR 及其根据 bootstrap 得到的 95%CI。

(2)边界因子法:已知模拟试验数据原始观测效应值为 2.15,假设 $RR_{xu}=RR_{ly}$,根据边界因子法可知,如果暴露与结局间真实效应值为 1,实际观测效应值 $RR_{xy}^{obs}=2.15$ 完全因混杂因素导致而与处理因素无关,此时敏感性参数 $RR_{xu}=RR_{ly} \geq 2.15 + \sqrt{2.15(2.15-1)}$,即 2 个参数至少满足 $RR_{xu}=RR_{ly} \geq 3.72$ 才能掩盖处理因素显著的有害效应,认为观测效应值 2.15 完全是因未测量混杂与处理因素和未测量混杂与结局之间关联错误导致。表 1 列出参数 RR_{xu} 和 RR_{ly} 不同取值所对应的校正后效应 RR 值及其 95%CI。行代表 RR_{xu} ,列代表 RR_{ly} 。根据表格内容可知,若 $RR_{xu} \neq RR_{ly}$,单元格 $(RR_{xu}, RR_{ly}) = (3.0, 3.5)$ 或 $(RR_{xu}, RR_{ly}) = (3.5, 3.0)$ 对应的校正效应值 $RR=1.13$ (95%CI: 0.95 ~ 1.36)。虽然效应值 $RR > 1$,但其 95%CI 下限 < 1 ,因此,只要 $\text{Min}(RR_{xu}, RR_{ly}) \geq 3$ 可以认为暴露与结局间的真实效应值完全是由混杂因素导致与暴露无关。同样,在该模拟试验中,在充分调整协变量的条件下,认为不可能存在该未测量混杂变量使其与暴露和结局的效应大小分别为 3.0(3.5)和 3.5(3.0)。在真实数据分析中,如果完全掩盖掉处理效应所需的未测量混杂强度太大而没有真实意义,那么认为处理与结局之间的因果关

系是真实存在的。

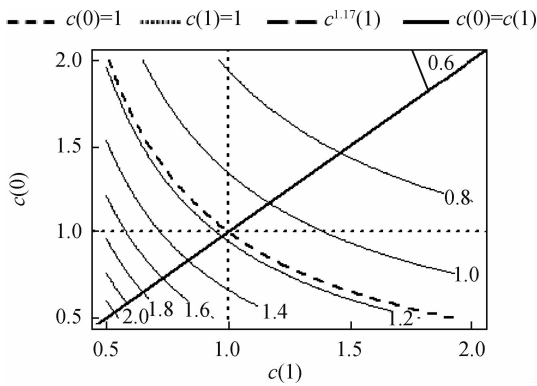
2. 实例分析结果:共有 5 735 例危重成年病例住进 ICU 接受治疗,2 184 例接受 RHC 治疗的病例中死亡 830 例,3 551 例没有接受 RHC 治疗病例中死亡 1 088 例,计算得粗 $RR=1.24$ 。根据前文介绍的倾向性评分法将数据中已知的协变量或混杂因素纳入到 logistic 回归模型中,得到校正后病例接受 RHC 治疗相对死亡风险 RR 值为 1.17(95%CI: 1.06 ~ 1.28)。考虑到可能存在某些未测量混杂对结果造成影响,以下介绍对效应值的稳健性进行敏感性分析的过程。

(1)混杂函数敏感性分析法:根据 Connors 等^[12]的结果可知,ICU 的患者接受 RHC 治疗比不接受 RHC 治疗的病例 30 d 的死亡风险增加了 17%,表明 RHC 手术治疗结果比不接受 RHC 手术治疗的结果要差。这个结果比较不符合临床经验,因为临床医生对于危重病例建议采取 RHC 手术治疗。因此,我们怀疑这个结果是否由于某些未测量到的混杂因素导致。

同模拟试验分析过程一样,图 3 表示混杂函数参数 $c(0)$ 和 $c(1)$ 不同取值对应的效应值 RR 等高曲线图。图 4 所示在校正不同混杂函数参数后的效应值 RR 及 bootstrap 产生的 95%CI。图 3 中 $c^{1.17}(1)$ 标注的线是指当调整未测量混杂后的真实效应应为 $RR_{xy}^{true}=1.17$ 时,对应 $c(1)$ 的不同取值。根据临床经验可知,实际接受 RHC 的病例比没有接受 RHC 的病例病情严重,因而,若在反事实的情形下,若所有人都不接受 RHC,且已调整的混杂不能解释所有的混杂,则反事实接受 RHC 治疗患者最终出现死亡结局的可能性更高,即 $c(0)$ 和 $c(1)$ 一般取值 > 1 。但是,若所有人都反事实地接受 RHC,原来未接受 RHC 的病例病情较轻,RHC 副反应导致死亡风险可能增加,但原来接受 RHC 的病例本身病情较重,通过实行 RHC 更好地了解心脏的有关参数,从而进行有针对性的治疗,反而能减少死亡率,故也可假定 $c(1) \leq 1$ 。图 4 中,随着混杂函数值的增大,效应值 RR 逐渐减少,理论上是合理的。因为未测量到的未知混杂越多,混杂函数校正的效应值 RR 就越小,将观测到

表 1 未测量混杂的校正估计值及其 95%CI

RR_{xu}	$RR_{ly}=1.5$	$RR_{ly}=2.0$	$RR_{ly}=2.5$	$RR_{ly}=3.0$	$RR_{ly}=3.5$	$RR_{ly}=4.0$
1.5	1.91(1.62 ~ 2.31)	1.79(1.52 ~ 2.17)	1.72(1.46 ~ 2.08)	1.67(1.42 ~ 2.02)	1.64(1.39 ~ 1.98)	1.61(1.37 ~ 1.95)
2.0	1.79(1.52 ~ 2.17)	1.61(1.37 ~ 1.95)	1.51(1.27 ~ 1.82)	1.43(1.21 ~ 1.73)	1.38(1.17 ~ 1.67)	1.34(1.14 ~ 1.63)
2.5	1.72(1.46 ~ 2.08)	1.51(1.27 ~ 1.82)	1.38(1.16 ~ 1.66)	1.29(1.09 ~ 1.56)	1.23(1.04 ~ 1.49)	1.18(1.00 ~ 1.43)
3.0	1.67(1.42 ~ 2.02)	1.43(1.21 ~ 1.73)	1.29(1.09 ~ 1.56)	1.19(1.01 ~ 1.44)	1.13(0.95 ~ 1.36)	1.08(0.91 ~ 1.30)
3.5	1.64(1.39 ~ 1.98)	1.38(1.17 ~ 1.67)	1.23(1.04 ~ 1.49)	1.13(0.95 ~ 1.36)	1.05(0.89 ~ 1.27)	1.00(0.85 ~ 1.21)
4.0	1.61(1.37 ~ 1.95)	1.34(1.14 ~ 1.63)	1.18(1.00 ~ 1.43)	1.08(0.91 ~ 1.30)	1.00(0.85 ~ 1.21)	0.94(0.80 ~ 1.14)



注:基于逆处理概率加权法校正已知混杂因素数据的 $c(0)$ 和 $c(1)$ 不同取值时 RR 变化情况; $c^{1.17}(1)$ 标注的线是指当真实效应估计值 $RR_{XY}^{true} = 1.17$ 时对应 $c(0)$ 和 $c(1)$ 各自不同取值变化

图3 接受RHC治疗者效应值敏感性分析等高曲线图

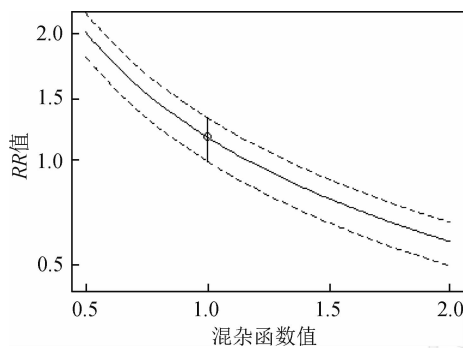


图4 RHC数据效应值RR的敏感性分析校正

的效应值归因于未测量混杂的影响。假设不存在未测量的混杂, RHC 治疗的风险比为 1.17, 95% CI (1.06 ~ 1.28)。在图 3 左上角的方格 [$c(0) \geq 1$ 及 $c(1) \leq 1$] 中, 可以看到大多数情况下 $RR > 1$, 说明 RHC 治疗增加受试者死亡的风险; 从图 4 中可以看出, 当混杂函数值在 (1 ~ 1.28) 范围内变化时对应的 95% CI 仍然没有显著性意义。同时, 不难注意到, 要使真实的 RR 值 = 1.00, 需要混杂函数值 $c(0) = c(1) = 1.17$, 即可能存在某些未测量混杂使接受 RHC 的病例的死亡风险是不接受 RHC 的死亡风险的 1.17 倍, 考虑到将 50 个变量组成的倾向性评分也仅仅使 RR 值的估计从 1.24 降到 1.17, 从专业上也不大容易再找到混杂因素能将 RR 值从 1.17 降到 1.00。

综上, RHC 能增加死亡风险的结论, 不太可能

是未测量混杂造成的。

(2) 边界因子敏感性分析法: 根据边界因子不等式可知敏感性分析参数 $RR_{XU} = RR_{UY} \geq 1.62$ 时, 暴露与结局之间的效应值完全由未测量混杂导致与暴露无关。表 2 给出敏感性参数 RR_{XU} 和 RR_{UY} 不同取值对应的效应值及 95% CI。当 $(RR_{XU}, RR_{UY}) = (1.6, 1.2)$ 或 $(RR_{XU}, RR_{UY}) = (1.2, 1.6)$ 时, 效应值 > 1 , 但其 95% CI (0.99 ~ 1.2) 下限 < 1 , 所以当未测量混杂分别与处理因素和结局间的最小效应值 > 1.2 时, 可以完全掩盖处理因素的效应大小。假设边界因子 2 个敏感性参数取值同混杂函数参数大小一致, $RR_{XU} = RR_{UY} = 1.17$, 校正后的 $RR \approx 1.14$ (95% CI: 1.03 ~ 1.24), 校正后的效应值稍微减弱一部分, 说明原始效应值含有未测量混杂导致的偏倚, 但 RHC 治疗与患者发生死亡结局之间仍存在正向因果关联; 只有当 $RR_{XU} = RR_{UY} \approx 1.6$ 强度时, 校正后的 $RR \approx 1$ 才能完全掩盖处理与结局间的效应大小。然而, 实际分析是在校正已知混杂的基础上进行的, 仍有如此强度的未测量混杂存在可能性很小, 所以该方法同样支持原文结论即 RHC 治疗增加患者死亡风险。

由混杂函数法分析结果可知, 当混杂函数参数 $c(0) = c(1) = 1.17$ 时暴露与结局之间的真实关联大小 $RR_{XY}^{true} = 1.00$; 根据边界因子法可知, 当 2 个参数 $RR_{XU} = RR_{UY}$ 且至少为 1.62 时, 才能做出同样的解释即观测的效应值 $RR_{XY}^{obs} = 1.17$ 完全因混杂因素导致与处理无关。虽然两种分析结果均支持原文中的结论危重病例接受 RHC 治疗后死亡风险增加, 但在校正未测量混杂因素影响方面, 混杂函数法在完全解释处理效应值时易理解, 能更加灵敏的识别因果效应中由未测量混杂导致的效应偏倚, 且混杂函数法只需一个参数 $c(a)$ 而边界因子法需要同时指定 2 个参数。研究者可以根据图 2 或图 4 所示信息直观的查看当 95% CI 包含 1.00 没有显著性意义时对应的混杂参数值, 分析混杂参数值处在什么范围内处理因素真实的效应值没有意义。所以混杂函数法在校正未测量混杂影响时更容易执行且结果更加可靠。

表2 校正未测量混杂后的效应值估计及其95%CI

RR_{XU}	$RR_{UY}=0.8$	$RR_{UY}=1.0$	$RR_{UY}=1.2$	$RR_{UY}=1.6$	$RR_{UY}=2.0$	$RR_{UY}=2.2$
0.8	1.10(0.99 ~ 1.20)	1.17(1.06 ~ 1.28)	1.22(1.10 ~ 1.33)	1.28(1.16 ~ 1.40)	1.32(1.19 ~ 1.44)	1.33(1.20 ~ 1.45)
1.0	1.17(1.06 ~ 1.28)	1.17(1.06 ~ 1.28)	1.17(1.06 ~ 1.28)	1.17(1.06 ~ 1.28)	1.17(1.06 ~ 1.28)	1.17(1.06 ~ 1.28)
1.2	1.22(1.10 ~ 1.33)	1.17(1.06 ~ 1.28)	1.14(1.03 ~ 1.24)	1.10(0.99 ~ 1.20)	1.07(0.97 ~ 1.17)	1.06(0.96 ~ 1.16)
1.6	1.28(1.16 ~ 1.40)	1.17(1.06 ~ 1.28)	1.10(0.99 ~ 1.20)	1.01(0.91 ~ 1.10)	0.95(0.86 ~ 1.04)	0.93(0.84 ~ 1.02)
2.0	1.32(1.19 ~ 1.44)	1.17(1.06 ~ 1.28)	1.07(0.97 ~ 1.17)	0.95(0.86 ~ 1.04)	0.88(0.80 ~ 0.96)	0.85(0.77 ~ 0.93)
2.2	1.33(1.20 ~ 1.45)	1.17(1.06 ~ 1.28)	1.06(0.96 ~ 1.16)	0.93(0.84 ~ 1.02)	0.85(0.77 ~ 0.93)	0.82(0.74 ~ 0.90)

讨 论

真实世界研究中,如果没有考虑相关混杂因素的非随机效应,导致对因果效应值被错误估计,可能会对患者安全方面或制定相关卫生政策产生严重的负面影响。真实世界研究正确估计因果效应值的关键任务是评估因果结论对未观测到的混杂的敏感度^[16]。虽然不能直接测量未测量混杂效应大小,但可以通过敏感性分析估计它的潜在影响^[17]。敏感性分析被认为是在研究设计、数据收集和数据分析中尽力减少、控制或消除此类混杂所导致偏倚的最后一道防线。以往文献中提到的敏感性分析方法常依赖于得不到精确验证的假设,如只存在一个二分类的混杂变量,暴露与混杂因素之间没有交互作用等。

混杂函数法在分析混杂对因果效应值影响时非常灵敏。其优点包括不对混杂因素做任何限制性假设,考虑研究中全部混杂对因果效应值的影响不只是未测量混杂,以及混杂函数参数 $c(a)$ 将未观测到的混杂大小度量。混杂函数法通过调整参数 $c(a)$ 不同大小分析其对因果效应值的大小和方向上改变的影响(图1)。虽然混杂函数法并不总是比其他方法合适,但这种方法在当数据集中可用变量有限的情况下是一种较有用的替代方法^[10]。缺点是混杂函数需要对各组之间的差异进行整体观察,但具体属于哪个方向更合适尚不能清楚的指定;不同混杂参数值在实际分析中是否具有实际临床意义需要由专业人员判断,具有主观性。

边界因子敏感性分析法的特点是对研究中涉及到的混杂不做任何假设,如不做只存在一个二分类的未测量混杂变量或者未测量混杂与暴露之间不存在交互作用等假设;通过构造边界因子不等式,将观测到的效应值或其95%CI除以边界因子,得到真实因果效应的估计值及95%CI。但由于该过程需要同时定义2个敏感性参数 RR_{xu} 和 RR_{xy} ,在分析应用时同其他方法相比较复杂。同样该方法需由专业人员对未测量混杂的取值大小合理性进行讨论与判断。

因此,真实世界研究中,研究人员应该充分考虑混杂因素可能造成的影响,并在最终统计分析报告补充敏感性分析结果,从而更加精确的评估暴露与结局之间真实因果效应大小。另外,本文主要研究结局变量是二分类变量。当感兴趣的效应值不再是RR或OR而是效应之差RD时,需对风险差做敏感性分析,本文介绍的2种方法也均适用,只需对方法中设置的参数和计算方法做相应的变换。关于未测

量混杂导致的效应偏倚问题仍需要大量研究与分析,只有完全充分了解研究中存在的混杂才能做出真实有效的结论,为促进科学发展提供更好的基础。

利益冲突 所有作者均声明不存在利益冲突

参 考 文 献

- [1] 耿直. 观察性研究与混杂因素[J]. 统计与信息论坛, 2004(5): 13-17. DOI: 10.3969/j.issn.1007-3116.2004.05.003.
- [2] 中国临床医学真实世界研究施行规范专家委员会. 中国临床医学真实世界研究施行规范[J]. 解放军医学杂志, 2018, 43(1): 1-6. DOI: 10.11855/j.issn.0577-7402.2018.01.01.
- [3] Stürmer T, Joshi M, Glynn RJ, et al. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods[J]. J Clin Epidemiol, 2006, 59(5): 437-447. DOI: 10.1016/j.jclinepi.2005.07.004.
- [4] Juurlink DN, Mamdani MM, Kopp A, et al. The risk of suicide with selective serotonin reuptake inhibitors in the elderly[J]. Am J Psychiatry, 2006, 163(5): 813-821. DOI: 10.1176/ajp.2006.163.5.813.
- [5] 钱维, 叶小飞, 王超, 等. 药品不良反应信号检测中混杂因素的控制方法[J]. 中国药物警戒, 2010, 7(3): 142-144. DOI: 10.3969/j.issn.1672-8629.2010.03.004.
- [6] Qian W, Yie XF, Wang C, et al. Methods of controlling Confounding Factors in Adverse Drug Reaction Signal Detection[J]. Chin J Pharmacovigilance, 2010, 7(3): 142-144. DOI: 10.3969/j.issn.1672-8629.2010.03.004.
- [7] Rosenbaum PR, Rubin DB. Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome[J]. J Royal Stat Soc Series B, 1983, 33: 212-218. DOI: 10.2307/2345524.
- [8] Vanderweele TJ, Arah OA. Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments, and confounders[J]. Epidemiology, 2011, 22(1): 42-52. DOI: 10.1097/EDE.0b013e3181f74493.
- [9] Ding P, Vanderweele TJ. Sensitivity analysis without assumptions[J]. Epidemiology, 2016, 27(3): 368-377. DOI: 10.1097/ede.0000000000000457.
- [10] Kasza J, Wolfe R, Schuster T. Assessing the impact of unmeasured confounding for binary outcomes using confounding functions[J]. Int J Epidemiol, 2017, 46(4): 1303-1311. DOI: 10.1093/ije/dyx023.
- [11] 苗旺, 刘春辰, 耿直. 因果推断的统计方法[J]. 中国科学: 数学, 2018, 48(12): 1753-1778. DOI: 10.1360/N012018-00055.
- [12] Miao W, Liu CC, Geng Z. Statistical approaches for causal inference[J]. Sci Chin: Series A, 2018, 48(12): 1753-1778. DOI: 10.1360/N012018-00055.
- [13] Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies[J]. J Educ Psychol, 1974, 66: 688-701. DOI: 10.1037/h0037350.
- [14] Connors AF Jr., Speroff T, Dawson NV, et al. The effectiveness of right heart catheterization in the initial care of critically ill patients. SUPPORT Investigators[J]. JAMA, 1996, 276(11): 889-897. DOI: 10.1001/jama.1996.03540110043030.
- [15] 李智文, 刘建蒙, 任爱国. 基于个体的标准化法——倾向评分加权[J]. 中华流行病学杂志, 2010, 31(2): 223-226. DOI: 10.3760/cma.j.issn.0254-6450.2010.02.024.
- [16] Li ZW, Liu JM, Ren AG. Introduction to an individual-based standardization method-propensity score weighting[J]. Chin J Epidemiol, 2010, 31(2): 223-226. DOI: 10.3760/cma.j.issn.0254-6450.2010.02.024.
- [17] Stürmer T, Rothman KJ, Glynn RJ. Insights into different results from different causal contrasts in the presence of effect-measure modification[J]. Pharmacoepidemiol Drug Saf, 2006, 15(10): 698-709. DOI: 10.1002/pds.1231.
- [18] Austin PC. The use of propensity score methods with survival or time-to-event outcomes: reporting measures of effect similar to those used in randomized experiments[J]. Statist Med, 2014, 33(7): 1242-1258. DOI: 10.1002/sim.5984.
- [19] Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics[J]. Pharmacoepidemiology Drug Saf, 2006, 15(5): 291-303. DOI: 10.1002/pds.1200.
- [20] Groenwold RH, Nelson DB, Nichol KL, et al. Sensitivity analyses to estimate the potential impact of unmeasured confounding in causal research[J]. Int J Epidemiol, 2010, 39(1): 107-117. DOI: 10.1093/ije/dyp332.

(收稿日期: 2019-02-27)

(本文编辑: 李银鸽)