

三臂非劣效临床试验基于干预依从性的效应估计方法*

吴研鹏¹ 陈平雁² 吴莹^{2△}

【摘要】目的 临床试验中干预不依从问题通常无法避免,分析方法不当会导致结果偏倚,非劣效研究尤其如此。本研究针对三臂非劣效临床试验,建立 Bayes 因果模型,实现对干预依从人群的效应估计。**方法** 基于主分层思想,按依从性类型将试验人群分层,从而转化为不同依从性人群的结局混合分布识别问题。通过构建 Bayes 模型和使用 DA (data augmentation) 算法,实现结局参数后验分布估计和统计推断。通过模拟研究,比较本研究提出方法与传统的意向性治疗 (intention-to-treat, ITT)、遵循研究方案 (per-protocol, PP) 及实际接受治疗 (as-treated, AT) 分析的统计性能。**结果** 当依从类型与结局相关时,传统的 ITT、PP 及 AT 分析均存在较大偏倚;本研究提出方法在依从类型与结局相关和不相关两种情况下,均能够将偏倚控制在较小的范围。**结论** 对于存在较严重干预不依从问题的非劣效临床试验,本研究提出方法较传统分析可以更好地控制分析偏倚。

【关键词】 不依从 三臂非劣效性临床试验 Bayes 因果模型 DA 算法

【中图分类号】 R195 **【文献标识码】** A **DOI** 10.11783/j.issn.1002-3674.2024.02.011

Estimating Treatment Effects in Three-arm Non-inferiority Clinical Trials based on Compliance of Active Treatments

Wu Yanpeng, Chen Pingyan, Wu Ying (Key Laboratory of Public Health Safety of Ministry of Education, Department of Epidemiology, School of Public Health, Fudan University(200032), Shanghai)

【Abstract】 Objective Non-compliance of active treatments occurred in clinical trials is usually unavoidable, and improper use of standard approaches may lead to biased results, especially for non-inferiority trials. Thus, for three-arm non-inferiority clinical trials, we established a Bayes causal model to estimate causal effects in the presence of non-compliance.

Methods Based on the framework of principal stratification, population was stratified according to types of compliance, and the issue was transformed into mixed-distribution identification. Bayes causal model was constructed and data augmentation (DA) algorithm were employed to calculate the posterior distribution of parameters of interest and complete statistical inference. Through simulation, we evaluated performances of our approach, compared with traditional methods including ITT, PP and AT.

Results The method of ITT, PP, and AT all had a significant bias when the type of compliance associated with outcomes. The method proposed in this study both had a good performance whether the type of compliance associated with outcomes or not.

Conclusion For non-inferiority trials with a high proportion of non-compliance, the method in this article has a better control of the bias.

【Key words】 Non-compliance; Three-arm non-inferiority trials; Bayes causal model; DA algorithm

受试者不按照随机分配方案接受干预是随机对照临床试验中不依从问题之一,即所谓干预不依从^[1]。针对这一问题的统计分析,无论采用意向性治疗 (intention-to-treat, ITT) 分析,实际接受治疗 (as-treated, AT) 分析,还是遵循研究方案 (per-protocol, PP) 分析,均存在一定局限性^[2]。从因果推断框架出发的统计方法,如工具变量法^[3]和依从者平均因果效应 (complier average causal effect, CACE) 方法^[4-6],针对两臂设计能得出更可靠的效应估计。针对“金标准”的三臂非劣效临床试验^[7],即包含试验组、阳性对照组和安慰剂组的试验,有关干预不依从虽开展了一些研究,但仍有待完善^[8-9]。本研究采用 ICH E9(R1) 提出的主层策略 (principal stratum strategy)^[10],基于

Bayes 因果模型框架^[6, 11],构建三臂非劣效试验中的 CACE 主估计量,针对干预依从人群进行效应估计,以期更准确地估计干预效应和进行非劣效评价。

原理与方法

1. 基本原理

本研究将干预依从性分为依从和不依从两类,不考虑部分依从情况。试验中,受试者的完整信息包括 Z 、 D 、 T 、 Y 和 X 五方面: Z 表示随机化分配结果, $Z=0$ 为安慰剂组, $Z=1$ 为阳性对照组, $Z=2$ 为试验组; D 表示受试者潜在依从特征, $D_i(Z_i)=1$,表示受试者 i 对其分配结果 Z_i 依从, $D_i(Z_i)=0$ 则表示不依从,受试者 i 的完整依从特征由 $D_i=(D_i(Z_i=0), D_i(Z_i=1), D_i(Z_i=2))$ 刻画,由于只能观测到受试者在实际分组情况下的依从性结果,故其他两种分配结果下的依从性缺失; T 为受试者的实际用药情况,由 Z 和 D 可推断出 T , $T=0, 1, 2$ 分别表示实际服用的是安慰剂、阳性对照药、试验药; Y 表示最终结局; X 表示可能与潜在结

* 基金项目:国家自然科学基金(81703322, 82273732);广东省医学科研基金资助(A2019438)

1. 复旦大学公共卫生学院流行病学教研室公共卫生安全教育部重点实验室(200032)

2. 南方医科大学公共卫生学院生物统计学系

△通信作者:吴莹, E-mail: wuying19890321@gmail.com

局和依从类型相关的协变量集。基于主分层思想的 CACE 法是对依从人群进行效应估计,本研究将依从人群定义为满足 $D=(1,1,1)$ 。在三臂非劣效试验中,同时设定两主估计量,分别为阳性对照组相对安慰组的效应 $CACE_{10}$ 、试验组相对安慰组的效应 $CACE_{20}$,表示为:

$$CACE_{10} = E[\bar{y}_1 - \bar{y}_0 | D=(1,1,1)] \quad (1)$$

$$CACE_{20} = E[\bar{y}_2 - \bar{y}_0 | D=(1,1,1)] \quad (2)$$

式中, $D=(1,1,1)$ 表示潜在特征为依从人群, $\bar{y}_0, \bar{y}_1, \bar{y}_2$ 分别表示依从者在安慰剂组、阳性对照组、试验组中的平均潜在效应。基于非劣效设计中的阳性对照效应保留法^[12],针对事先设定的保留比例 θ ,估计出 $\frac{CACE_{20}}{CACE_{10}}$,通过与 θ 进行比较,完成试验药的非劣效评价。但由于依从特征 D 属于不可观测的潜变量,因此 $CACE_{20}$ 和 $CACE_{10}$ 无法直接识别,需要借助模型和假设进行估计。

2. 模型假设

本研究基于因果推断经典假设^[6],并结合三臂临床试验设计特点做如下模型假设:

假设①:实施随机化分组,即 $(Y, D, X, T) \perp Z$;

假设②:受试者间互相独立,即 $Y_i(Z, T) = Y_i(Z_i, T_i)$;

假设③:受试者结局只与实际接受组有关,即 $Y_i(Z_i, T_i, D_i) = Y_i(T_i, D_i)$;

假设④:试验组和阳性对照组不能相互转组,安慰剂组不能转组到试验组或阳性对照组,即 $T(Z=1) \neq 2, T(Z=2) \neq 1, T(Z=0) = 0$;

假设⑤:依从性特征存在以下单调性, $P(D(1)=1 | D(2)=1) = 1$,即对阳性对照组依从的受试者一定对试验药依从。

假设①~③为主分层思想中常用因果假设,假设④简化试验人群不同依从特征类数,假设⑤为强假设,通过单调性对干预组依从性进行约束,进而使得混合模型的参数估计可部分确认。基于假设①~⑤,将潜在依从特征 D 简化为三种依从类型,用 C 表示,模型中可能的依从类型见表 1。

表 1 五种假设下的可能依从类型

依从类型(C)	D			T		
	D(0)	D(1)	D(2)	T(0)	T(1)	T(2)
0 (仅对安慰剂依从)	1	0	0	0	0	0
1 (对安慰剂和阳性对照药依从)	1	1	0	0	1	0
2 (对各分配结果均依从)	1	1	1	0	1	2

3. 参数估计

本研究基于贝叶斯模型,首先对估计参数构建似然函数,估计参数包括两部分:结局变量的分布参数 π_1 和依从性类型的分布参数 π_2 。基于假设①~⑤,对于 π_1 ,个体 i 有概率函数或者密度函数:

$$f(Y_i | C_i, T_i, Z_i, X_i, \pi_1) = f\{Y_i | C_i, T_i(Z_i, C_i), X_i, \pi_1\} = f\{Y(T_i) = Y_i | C_i, X_i, \pi_1\} \quad (3)$$

对 π_2 ,有概率函数:

$$f(C_i | Z_i, X_i, \pi_2) = f(C_i | X_i, \pi_2) \quad (4)$$

则试验人数为 n 下的模型似然函数表示如下:

$$L(\pi_1, \pi_2) = \prod_i^n L_i(\pi_1, \pi_2 | Y_i, T_i, C_i, X_i) = \prod_i^n f\{Y(T_i) = Y_i | C_i, X_i, \pi_1\} f(C_i | X_i, \pi_2) \quad (5)$$

由于公式(5)中依从性类型 C_i 不可观测,可根据 Z_i 和 T_i 的信息,判断可能取值,表示为 $C_{obs,i}$ 。比如在 $Z_i=1$ 和 $T_i=1$ 下, $C_{obs,i} = \{1, 2\}$ 。进一步将个体 i 的结局分布函数看作可能 $C_{obs,i} = c$ 取值下的有限混合分布函数,则

$$L(\pi_1, \pi_2) = \prod_i^n \sum_{c \in C_{obs,i}} f\{Y(T_i) = Y_i | C_i = c, X_i, \pi_1\} f(c | C_{obs,i}, X_i, \pi_2) \quad (6)$$

假设相同依从性类型的受试者具有相同的潜在特征(协变量集均衡),可将 $f\{Y(T_i) = Y_i | C_i = c, X_i, \pi_1\}$ 简称为 f_{ct}^i ,表示依从类型 $C=c$,实际接受处理 $T=t$ 的受试者结局概率密度函数; $f(c | Z_i, C_{obs,i}, X_i, \pi_2)$ 可简称为 ρ_c^i ,表示人群中依从类型为 $C=c$ 的受试者比例;根据 (Z_i, T_i) 组合判断 $C_{obs,i}$ 可能取值,并总结混合分布情况如表 2。

表 2 五种假设下的受试者 i 在所有 Z 和 T 组合下的似然函数

Z_i	T_i	依从类型 C_i			混合分布
		0	1	2	
0	0	$\rho_0^i f_{00}^i$	$\rho_1^i f_{10}^i$	$\rho_2^i f_{20}^i$	$(\rho_0^i f_{00}^i + \rho_1^i f_{10}^i + \rho_2^i f_{20}^i) / (\rho_0^i + \rho_1^i + \rho_2^i)$
1	1	0	$\rho_1^i f_{11}^i$	$\rho_2^i f_{21}^i$	$(\rho_1^i f_{11}^i + \rho_2^i f_{21}^i) / (\rho_1^i + \rho_2^i)$
0	0	f_{00}^i	0	0	f_{00}^i
2	2	0	0	f_{22}^i	f_{22}^i
0	0	$\rho_0^i f_{00}^i$	$\rho_1^i f_{10}^i$	0	$(\rho_0^i f_{00}^i + \rho_1^i f_{10}^i) / (\rho_0^i + \rho_1^i)$

基于表 2 和公式(6),可得出具体式如下

$$L(\pi_1, \pi_2) = \prod_{C_{obs,i}=\{1,2\}} \frac{\rho_1^i f_{11}^i + \rho_2^i f_{21}^i}{\rho_1^i + \rho_2^i} \times \prod_{C_{obs,i}=\{0\}} f_{00}^i \times \prod_{C_{obs,i}=\{2\}} f_{22}^i \times \prod_{C_{obs,i}=\{0,1\}} \frac{\rho_0^i f_{00}^i + \rho_1^i f_{10}^i}{\rho_0^i + \rho_1^i} \times \prod_{C_{obs,i}=\{0,1,2\}} \frac{\rho_0^i f_{00}^i + \rho_1^i f_{10}^i + \rho_2^i f_{20}^i}{\rho_0^i + \rho_1^i + \rho_2^i} \quad (7)$$

其中, ρ_c^i 表示个体 i 的依从类型为 c 的概率, $C_{obs,i}$ 表示个体 i 的依从类型可能取值, f_{ct}^i 表示对应的结局密度函数。用 $p(\eta)$ 表示 $\eta = (\pi_1, \pi_2)$ 的先验分布,则 η 的后验分布为

$$f(\eta | Z, T, Y, C_{obs,i}, X) \propto p(\eta) \prod_{C_{obs,i}=\{1,2\}} \frac{\rho_1^i f_{11}^i + \rho_2^i f_{21}^i}{\rho_1^i + \rho_2^i} \times \prod_{C_{obs,i}=\{0\}} f_{00}^i \times \prod_{C_{obs,i}=\{2\}} f_{22}^i \times \prod_{C_{obs,i}=\{0,1\}} \frac{\rho_0^i f_{00}^i + \rho_1^i f_{10}^i}{\rho_0^i + \rho_1^i} \times \prod_{C_{obs,i}=\{0,1,2\}} \frac{\rho_0^i f_{00}^i + \rho_1^i f_{10}^i + \rho_2^i f_{20}^i}{\rho_0^i + \rho_1^i + \rho_2^i} \quad (8)$$

通过对结局参数向量 π_1 的后验估计完成非劣效评价。

本研究采用 Data Augmentation (DA) 算法^[13], 先验设置为无信息先验。通过对依从性 C 的多项分布参数进行补足和更新, 进而识别各个依从类型下的结局分布, 从而估计分布参数的后验分布。根据最终参数的后验分布结果, 计算主估计量 $CACE_{20}$ 和 $CACE_{10}$, 完成统计推断:

$$CACE_{20} = \int Yf_{22}dY - \int Yf_{20}dY \quad (9)$$

$$CACE_{10} = \int Yf_{21}dY - \int Yf_{20}dY \quad (10)$$

模拟试验

通过统计模拟, 在不同样本量及依从率情况下, 比较本研究提出的方法与传统的 ITT、PP、AT 分析方法的估计性能。

1. 参数设置

基于模型假设下的可观测依从性由 Z 和 T 定义, 如表 2 中 ($Z=1, T=0$) 和 ($Z=2, T=0$) 为观测到的干预不依从组合; 通过控制表 1 中依从类型 C 为 0, 1, 2 对应的比例 ρ_0, ρ_1, ρ_2 参数值, 将可观测不依从率控制为 5% 至 20%, 如下四种情况:

(1) $(\rho_0, \rho_1, \rho_2) = (0.05, 0.05, 0.9)$, 不依从率 5%;

(2) $(\rho_0, \rho_1, \rho_2) = (0.1, 0.1, 0.8)$, 不依从率 10%;

(3) $(\rho_0, \rho_1, \rho_2) = (0.15, 0.15, 0.7)$, 不依从率 15%;

(4) $(\rho_0, \rho_1, \rho_2) = (0.2, 0.2, 0.6)$, 不依从率 20%;

干预不依从发生机制考虑两种情况: ①结局与依从性类型相关, 且假设成正相关关系 (高优指标), 即

同处理组中, 表现依从的人群, 其结局期望值更高; ②结局与依从性类型无关。假设安慰组、阳性对照组、试验组结局服从正态分布, μ_{ct} 和 σ_{ct} 分别对应表示依从类型 $C=c$, 实际接受处理 $T=t$ 的受试者结局均值和标准差。其依从人群在三组的结局均值分别为 $\mu_{20} = 10, \mu_{21} = 20, \mu_{22} = 18$; 当结局与依从性相关时, $(\mu_{11}, \mu_{10}, \mu_{00}) = (15, 8, 6)$; 结局与依从性无关, 只与实际接受组有关, 则 $(\mu_{11}, \mu_{10}, \mu_{00}) = (20, 10, 10)$; 所有标准差 $\sigma_{ct} = 2$ 。样本量考虑 360 例 (每组 120 例) 和 600 例 (每组 200 例) 两种情况, 以上每种模拟参数组合运行 1000 次。模拟试验采用 R3.6.3 软件编程实现。

2. 评价指标

包括偏倚 (bias), 均方误差 (MSE), 95% 可信区间覆盖率 (coverage rate)。试验组与安慰剂组, 阳性对照组与安慰剂组的疗效差分别记为 $\hat{\delta}_{20}$ 和 $\hat{\delta}_{10}$, 阳性组保留效应估计量表示为 $\hat{\theta}$ 。

3. 结果

表 3 和表 4 为依从性与结局相关情况下, 总样本量分别为 360 和 600 时各方法的统计性能。可以看出, 本研究方法在各参数组合下, 明显优于其他传统方法估计。随着不依从率增加, ITT, PP, AT 分析的平均误差和均方误差出现增加趋势, 其中 ITT 偏误最大, 而 CACE 估计表现稳定, 各种情况均小于或接近 0.01。在真值覆盖率上, CACE 大部分情况下接近 95%, 而其他方法偏离 95% 较远。随着样本量增加, CACE 估计表现稳健, 偏倚更小。

表 5 给出了依从性与结局无关情况下各方法的性能。总体来看, CACE 估计表现依旧稳定, 各种参数设置下平均误差均小于 0.01; PP 和 AT 的原理是基于接受的处理进行估计, 在不依从性与结局无关情况下表现出了无偏估计; 而 ITT 分析表现出了较大偏倚。

表 3 依从性与结局相关情况下的模拟试验结果比较 (样本量 = 360)

	bias				MSE				coverage rate			
	CACE	ITT	AT	PP	CACE	ITT	AT	PP	CACE	ITT	AT	PP
不依从率 5%												
$\hat{\delta}_{10}$	-0.079	-0.649	0.033	0.427	0.197	0.533	0.034	0.222	0.988	0.090	0.94	0.220
$\hat{\delta}_{20}$	-0.348	-0.807	0.287	0.681	0.335	0.764	0.107	0.495	0.962	0.024	0.082	0
$\hat{\theta}$	-0.003	-0.030	0.026	0.033	0.012	0.003	0.001	0.001	0.978	0.820	0.079	0.019
不依从率 10%												
$\hat{\delta}_{10}$	-0.128	-1.317	0.038	0.673	0.154	1.924	0.056	0.511	0.968	0.001	0.940	0.262
$\hat{\delta}_{20}$	-0.097	-1.599	0.598	1.232	0.135	2.756	0.391	1.555	0.970	0	0	0
$\hat{\theta}$	0.008	-0.061	0.057	0.065	0.004	0.007	0.004	0.005	0.955	0.570	0	0
不依从率 15%												
$\hat{\delta}_{10}$	-0.173	-1.911	0.037	0.794	0.216	3.903	0.074	0.698	0.934	0	0.154	0.001
$\hat{\delta}_{20}$	-0.048	-2.405	0.897	1.654	0.150	6.031	0.843	2.772	0.952	0	0	0
$\hat{\theta}$	0.012	-0.106	0.087	0.095	0.001	0.016	0.008	0.009	0.945	0.885	0.006	0
不依从率 20%												
$\hat{\delta}_{10}$	-0.155	-2.633	-0.067	0.737	0.252	7.201	0.097	0.628	0.941	0	0.819	0.178
$\hat{\delta}_{20}$	-0.003	-3.203	1.189	1.993	0.180	10.541	1.454	4.005	0.955	0	0	0
$\hat{\theta}$	0.015	-0.146	0.126	0.131	0.002	0.028	0.017	0.018	0.933	0.630	0	0

* 各分布下的均值设置为 $\mu_{20} = 10, \mu_{21} = 20, \mu_{22} = 18, \mu_{11} = 15, \mu_{10} = 8, \mu_{00} = 6, \sigma$ 均为 2; 每组样本量为 120 例。

表 4 依从性与结局相关情况下的模拟试验结果比较(样本量=600)

	bias				MSE				coverage rate			
	CACE	ITT	AT	PP	CACE	ITT	AT	PP	CACE	ITT	AT	PP
不依从率 5%												
$\hat{\delta}_{10}$	-0.084	-0.657	0.028	0.423	0.098	0.497	0.022	0.204	0.970	0.011	0.936	0.044
$\hat{\delta}_{20}$	-0.127	-0.794	0.298	0.694	0.104	0.700	0.105	0.500	0.972	0	0.001	0
$\hat{\theta}$	-0.001	-0.028	0.028	0.034	0.001	0.002	0.001	0.001	0.973	0.712	0.007	0
不依从率 10%												
$\hat{\delta}_{10}$	-0.09	-1.303	0.049	0.679	0.088	1.812	0.035	0.495	0.957	0	0.944	0.087
$\hat{\delta}_{20}$	-0.017	-1.593	0.605	1.234	0.068	2.657	0.384	1.542	0.957	0	0	0
$\hat{\theta}$	0.009	-0.062	0.056	0.065	0.004	0.006	0.003	0.004	0.956	0.406	0	0
不依从率 15%												
$\hat{\delta}_{10}$	-0.068	-1.946	0.020	0.773	0.108	3.933	0.048	0.642	0.953	0	0.033	0
$\hat{\delta}_{20}$	0.022	-2.378	0.899	1.653	0.087	5.813	0.833	2.752	0.952	0	0	0
$\hat{\theta}$	0.009	-0.101	0.088	0.096	0.001	0.013	0.008	0.010	0.934	0.866	0	0
不依从率 20%												
$\hat{\delta}_{10}$	-0.109	-2.596	-0.049	0.749	0.156	6.909	0.056	0.610	0.939	0	0.748	0.03
$\hat{\delta}_{20}$	0.001	-3.195	1.204	2.001	0.109	10.376	1.476	4.026	0.953	0	0	0
$\hat{\theta}$	0.010	-0.149	0.125	0.131	0.001	0.026	0.016	0.017	0.925	0.437	0	0

* 每组样本量为 200 例,其余参数设置同表 3

表 5 依从性与结局不相关情况下的模拟试验结果比较(样本量=600)

	bias				MSE				coverage rate			
	CACE	ITT	AT	PP	CACE	ITT	AT	PP	CACE	ITT	AT	PP
不依从率 5%												
$\hat{\delta}_{10}$	0.003	-0.494	0.004	0.004	0.052	0.281	0.011	0.010	0.954	0.162	0.946	0.943
$\hat{\delta}_{20}$	0.012	-0.794	0.005	0.004	0.049	0.667	0.011	0.010	0.952	0.007	0.940	0.943
$\hat{\theta}$	0.001	-0.042	0.000	0.000	0.000	0.002	0.000	0.000	0.952	0.589	0.950	0.949
不依从率 10%												
$\hat{\delta}_{10}$	0.004	-0.998	0.005	0.006	0.069	1.052	0.011	0.010	0.969	0.001	0.946	0.955
$\hat{\delta}_{20}$	-0.008	-1.604	0.004	0.004	0.062	2.629	0.011	0.010	0.943	0	0.932	0.936
$\hat{\theta}$	-0.001	-0.089	0.000	0.000	0.000	0.009	0.000	0.000	0.971	0.227	0.956	0.955
不依从率 15%												
$\hat{\delta}_{10}$	0.007	-1.494	0.003	0.001	0.097	2.301	0.011	0.009	0.969	0	0.948	0.949
$\hat{\delta}_{20}$	-0.006	-2.387	0.004	0.002	0.084	5.777	0.013	0.011	0.943	0	0.944	0.936
$\hat{\theta}$	0.000	-0.14	0.000	0.000	0.001	0.021	0.000	0.000	0.976	0.051	0.946	0.947
不依从率 20%												
$\hat{\delta}_{10}$	0.028	-1.994	0.000	0.000	0.154	4.067	0.011	0.009	0.975	0	0.934	0.949
$\hat{\delta}_{20}$	0.024	-3.199	-0.003	-0.003	0.106	10.316	0.013	0.011	0.968	0	0.949	0.953
$\hat{\theta}$	0.002	-0.200	0.000	0.000	0.001	0.042	0.000	0.000	0.982	0.008	0.952	0.950

* 各分布下的均值设置为 $\mu_{20} = 10, \mu_{21} = 20, \mu_{22} = 18, \mu_{11} = 20, \mu_{10} = 10, \mu_{00} = 10, \sigma$ 均为 2; 每组样本量为 200 例。

实例分析

重度抑郁症是一种常见的精神疾病,目前常用治疗方案为抗抑郁药物治疗,然其疗效有限,且会产生耐药性和不良反应。在此背景下,研究人员研发了一种非药物疗法,即经颅直流电刺激(tDCS)治疗法,该疗法几乎没有副反应,且操作简易、携带方便和价格低廉。证明 tDCS 较抗抑郁药的非劣效性,将对临床精神病学产生巨大影响,因此,该研究开展了一项非劣效性、随机、双盲、安慰剂对照的三臂试验,即三臂非劣效临床试验。其中试验组为 tDCS (E),阳性对照组为常用抗抑郁药——草酸艾司西酞普兰(R),安慰剂对照组是伪 tDCS 刺激加安慰剂药丸(P),主要观察指标是 HDRS-17 抑郁评定量表评分的变化^[14]。本研究在此背景下假想和模拟一组数据,主要目的是对实例的结

果进行展示和解读。

该数据是基于正态分布结局模拟,估计目标层定义为完全依从人群的效应,主要估计量为试验组相对阳性对照组保留效应比例,本研究将预定阳性保留效应比例阈值 θ 为 0.8。表 6 结果显示本研究模型较传统策略分析结果效果更好,有效估计了真实效应值和阳性保留效应比例;因此该模型可以更好校正三臂非劣效设计的不依从偏倚。

表 6 传统分析和贝叶斯因果模型估计结果表

	试验组效应	阳性对照组效应	阳性保留效应*
贝叶斯因果模型	2.68(2.19, 3.20)	3.31(2.79, 3.87)	0.81(0.69, 0.93)
ITT 分析	2.40(2.15, 2.65)	2.77(2.50, 3.05)	0.87(0.76, 0.98)
AT 分析	4.05(3.58, 3.91)	4.61(4.14, 4.49)	0.88(0.83, 0.91)
PP 分析	3.75(3.90, 4.20)	4.32(4.44, 4.78)	0.87(0.84, 0.91)
样本参数真实值	2.40	3.20	0.80

* :ITT、AT、PP 分析的阳性保留效应估计和对应 95% 置信区间是基于 bootstrap 法计算得出。

讨论

三臂设计是公认的金标准非劣效设计,然其较双臂设计可能面临更复杂的依从性问题,传统分析方法难以控制偏倚。本研究通过建立 Bayes 因果模型,提供一种新的三臂非劣效试验不依从效应估计的思路和方法,为该设计进一步应用和推广做铺垫。

本研究基于 ICH E9(R1) 建议的主层策略,对三臂非劣效临床试验进行干预依从效应估计。由于干预依从人群对试验药组、阳性对照组和安慰剂组的潜在依从性相同,因此在临床试验不同分配组中,该人群具有完全可交换性,针对其估计的干预效应具有因果效应。因此干预依从效应能够反映干预依从人群在真实世界中的疗效,是评估临床试验药物非劣效的重要估计量。在不同模拟设置下,本研究方法均取得较好的干预依从效应估计,而传统 ITT、PP 和 AT 方法的效应估计受依从性影响较大。当依从性与结局无关时,本研究方法、PP 方法和 AT 方法均能取得无偏估计,而 ITT 偏倚较大;当依从性与结局相关时,三种传统方法均有较大偏倚,且有随不依从率增大而增大的趋势,而本研究仍然估计稳健。

本研究建立的 Bayes 因果模型,其实质是基于因果推断的主分层框架,根据人群依从性进行分层,使主层内人群可交换,从而运用 Bayes 方法,对未观测结局的后验分布推断,实现干预间的因果效应估计。然而定义主层的依从类型属于不可观测变量,实际人群结局是不同依从类型人群结局的混合分布,其对应混合模型的参数估计通常会面临不确定性问题,即不同参数可能形成同一混合分布,可能导致估计结果偏倚^[15]。对此,本研究在假设①~④的基础上,增加了单调性假设(假设⑤),使得部分主层可识别,从而降低了算法识别的不确定,提高了估计性能。此外,DA 算法实施中,经常遇到干预互换问题^[16],本研究是基于一定现实假设,根据不同依从性下的结局大小关系进行约束,即接受相同某组的处理,潜在依从性特征高(比如对各组分配表现依从)的结局期望值更高,从而避免干预互换问题,使得算法性能明显改善。然而,当结局与依从无关情况下或者不同分布的差距很小时,该方法估计时可能出现置空现象,即混合模型中某一分布的比例迭代为 0。这种情况下,暗示模型的混合分布类数不成立,应及时调整这种约束方法或者调整模型,更合适的解决干预互换问题的方法可进一步探索^[16]。另外在本研究的 Bayes 框架下,利用先验信息影响收敛方向,也可提高估计准确性。对于如何通过设置合理先验,影响后验分布收敛方向,仍需进一步研究^[17]。

同类似研究比较^[5,18],本研究主要针对干预依从

人群进行效应估计,其非劣效评价只能适用于干预依从人群。其他依从类型人群的效应也可同时估计,其视目标人群改变主估计量;当目标人群为包含所有依从类型的全部人群时,基于对应依从类型比例加权下的干预效应估计结果等同于 ITT 分析。本研究是基于一定假设,简化了人群依从类型。对于其他可能假设,CACE 估计虽然表现形式可能不同,但原理具有类似性。实际应用中,可合理改变假设,并进行敏感性分析,从而实现模型估计的稳健性评价。

模型估计采用的是 DA 算法,其基于 MCMC 采样实现,对多参数估计时需要较大计算量;本研究根据结局方差大小及 MCMC 收敛情况,得出模拟样本 1000 次即可准确估计,实际中可视具体情况合理调整模拟次数。另外本研究只模拟了在不同依从类型比例和样本量下,四种方法的效应估计准确度。除此之外,混合分布间的疗效差、方差是否齐性等对估计性能也具有一定影响,进一步的研究中,可增设依从性比例,疗效差和方差作为模拟参数进行比较。同时可通过这些模拟参数,推导 ITT、AT、PP 统计量与 CACE 主估计量的函数关系式,从而在因果框架下对传统估计方法作定量偏倚评价。

综上所述,本研究提出的方法较传统分析方法可以更清晰和更灵活地控制分析偏倚,可用于三臂非劣效试验存在干预不依从的效应估计。

参 考 文 献

- [1] Jo B. Bias Mechanisms in Intention-to-Treat Analysis With Data Subject to Treatment Noncompliance and Missing Outcomes. *Journal of Educational and Behavioral Statistics*, 2007, 33(2):158-185.
- [2] Dodd S, White IR, Williamson P. Nonadherence to treatment protocol in published randomised controlled trials; a review. *Trials*, 2012, 13(1):84.
- [3] Angrist JD, Imbens GW, Rubin DB. Identification of Causal Effects Using Instrumental Variables(with Discussion). *Journal of the American Statistical Association*, 1996, 91(434):444-472.
- [4] Frangakis CE, Rubin DB. Principal Stratification in Causal Inference. *Biometrics*, 2015, 58(1):21-29.
- [5] Ye C, Beyene J, Browne G, et al. Estimating treatment effects in randomised controlled trials with non-compliance; a simulation study. *BMJ Open*, 2014, 4(6):e5362.
- [6] Imbens GW, Rubin DB. Bayesian Inference for Causal Effects in Randomized Experiments with Noncompliance. *The Annals of statistics*, 1997, 25(1):305-327.
- [7] 国家药品监督管理局药品审评中心. 药物临床试验非劣效设计指导原则. 2020. <http://www.cde.org.cn/news.do?method=largeInfo&id=31c0a47beeb92f8f>.
- [8] Wu Y, Zhao L, Hou Y, et al. Correcting for non-compliance in randomized non-inferiority trials with active and placebo control using structural models. *Statistics in Medicine*, 2015, 34(6):950-965.

(下转第 222 页)