

·基础理论与方法·

纵向研究中控制时依混杂的 G 方法

梁洁 和思敏 陈淑婷 王彤

山西医科大学公共卫生学院卫生统计学教研室,太原 030012

通信作者:王彤,Email: tongwang@sxmu.edu.cn

【摘要】 传统分析方法不能有效地控制纵向研究中的时依混杂以得到无偏因果效应估计值。本研究解释了纵向研究中正确控制时依混杂的必要性,概述了现有控制时依混杂的 G 方法——参数 g-formula、逆概率加权和 G 估计,并通过比较它们的优缺点和适用情况,为研究者在纵向研究中估计因果效应提供参考。

【关键词】 时依混杂; 参数 g-formula; 逆概率加权; G 估计

基金项目: 国家自然科学基金(81872715)

G methods for handling time-varying confounding in the longitudinal study

Liang Jie, He Simin, Chen Shuting, Wang Tong

Department of Health Statistics, School of Public Health, Shanxi Medical University, Taiyuan 030012, China

Corresponding author: Wang Tong, Email: tongwang@sxmu.edu.cn

【Abstract】 The conventional analytical methods cannot effectively adjust for time-varying confounding that occur in a longitudinal study and thus cannot correctly estimate the causal effects. This study explains the necessity of precisely controlling time-varying confounding and outlines G methods, including parametric g-formula, inverse probability of weighting, and G-estimation. We also compare the methods above to provide a reference for correctly estimating causal effects in the longitudinal study.

【Key words】 Time-varying confounding; Parametric g-formula; Inverse probability of weighting; G-estimation

Fund program: National Natural Science Foundation of China (81872715)

在观察性研究中,估计暴露与结局间因果效应的关键是控制混杂。当暴露不随时间变化(time-fixed treatment)时,例如接种疫苗/外科手术等仅在基线发生的处理因素、基因型/性别等不随时间变化的因素、年龄等随时间有确定性变化的因素,采用匹配、分层、回归校正、倾向性评分等传统分析方法即可实现混杂的有效控制。当前,很多大型人群队列研究都试图推断基线暴露与预设的一次随访结局的关系,其中一个常被忽视的假设是基线暴露及其可能的混杂是不随时间变化的。然而在流行病学研究中,许多感兴趣的暴露是随时间变化(time-varying treatment)的,例如药物剂量、用药

方案、吸烟状态、空气污染程度、社会经济状态等,此时要估计时变暴露的因果总效应,就必须考虑时依混杂的有效控制。

一、时依混杂

当感兴趣的暴露随时间变化时,可能存在随时间变化的、受过去暴露水平影响并影响着后续暴露水平的混杂,即时依混杂^[1]。例如,在研究抗反转录病毒治疗(anti-retroviral therapy, ART)($X=1$ 为 ART, $X=0$ 为其他治疗)对 HIV 感染者的 CD4⁺T 淋巴细胞(CD4)计数(Y)的影响时(例 1),HIV 载量(C)为时依混杂($C=1$ 为 HIV 病毒载量 >200 拷贝/ml, $C=0$ 则相反)。在研究起始时受试者的基线 HIV 病毒

DOI: 10.3760/cma.j.cn112338-20200731-01001

收稿日期 2020-07-31 本文编辑 李银鸽

引用本文:梁洁,和思敏,陈淑婷,等.纵向研究中控制时依混杂的 G 方法[J].中华流行病学杂志,2021,42(10):1871-1875. DOI: 10.3760/cma.j.cn112338-20200731-01001.



载量均 >200 拷贝/ml ($C_0=1$), 而第一次随访时的 HIV 病毒载量(C_1)会受到研究起始时治疗方案(X_0)的影响, 同时其又影响着该次随访时的治疗方案(X_1), 成为 X_0 到 Y 的一条因果通路上的中介变量(图 1)。此外, 当存在同时影响 HIV 病毒载量(C)和 CD4 计数(Y)的未测量协变量 U_0 时, C_1 是 X_0 、 U_0 和 Y 因果通路上的对撞因子(collider)(图 1)。在这些复杂情形下, 若采用传统分析方法直接控制 C_1 , 很可能产生过度校正偏倚(控制了中介变量 C_1)或对撞机偏倚(collider bias, 控制了对撞变量 C_1), 从而无法得到无偏因果效应估计值。

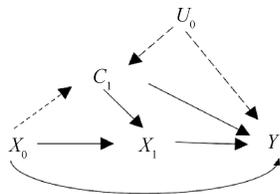


图 1 抗反转录病毒治疗与 HIV 感染者 CD4 计数的因果图

在临床流行病学纵向研究中, 许多临床体征或生化指标不仅影响着后续的暴露水平, 其本身也受过去暴露水平的影响。例如在研究哮喘抢救药物与呼气流量率峰值(peak expiratory flow rate, PEFR)的因果关系中, 某阶段哮喘症状(时依混杂)出现的频率及严重程度与先前用药剂量密切相关, 同时也影响着后续用药水平^[2]; 在探索他汀类药物对冠心病一级预防效果的研究中, 存在受先前是否服用他汀类药物影响的时依混杂, 如血清胆固醇水平、DBP 值等^[3]; 在 β 受体阻滞剂与急性心梗死亡之间的因果关系研究中, 血压值是受先前用药影响的时依混杂^[4]。因此, 纵向研究中采用有效的统计方法控制时依混杂是其因果推断的关键。

二、方法和原理

目前发展了 3 种控制时依混杂的统计方法: 参数 g-formula、逆概率加权(inverse probability of weighting, IPW)和 G 估计, 统称为 G 方法。不同于传统方法, G 方法不会通过将混杂因素固定在某个值或某个水平而实现控制混杂的目的, 因此可有效避免过度校正偏倚和对撞机偏倚, 从而得到无偏因果效应估计值^[5]。令 $X(t)$ 为时变二分类暴露, $C(t)$ 为二分类时依混杂, 假设 $C(t)$ 在固定的时间点 $t=0, \dots, k$ 处改变, $\bar{X}(t) = \{X(0), \dots, X(t)\}$ 和 $\bar{C}(t) = \{C(0), \dots, C(t)\}$ 分别代表每次随访时的暴露状态和混杂水平。

1. 参数 g-formula: 参数 g-formula 利用了标准化的思想, 在多个时间点上构建 \bar{X} 、 \bar{C} 和 Y 的联合分布, 模拟受试者在暴露和非暴露情况下产生的潜在结果, 并估计总体平均潜在结局 $E(Y^{\bar{x}}) = \sum_{c \in \bar{c}} \{E[Y|\bar{x}, \bar{c}] \prod_{t=0}^k f[(c_t|\bar{X}_t = \bar{x}_{t-1}, \bar{C}_t = \bar{c}_{t-1})]\}$, 例 1 的 $E(Y^{\bar{x}})$ 可简单表示为: $E(Y^{x_0, x_1}) = \sum_{c_1} E(Y|x_0, c_1, x_1)P(c_1|x_0)$, 通过比较暴露组和非暴露组的潜在预期结局估计处理效应^[6-7]。吴诗蓝等^[8]在探究膳食纤维摄入量(fiber ≤ 30 g/d 或 >30 g/d)是否影响冠心病发病风险的纵向研究中, 每次随访时糖尿病是否发生既受到过去膳食纤维摄入量的影响, 又可能使受试者未来的膳食行为发生改变, 是需要控制的时依混杂, 采用参数 g-formula 得到结果为 fiber >30 g/d 的人患冠心病的风险是 fiber ≤ 30 g/d 患病风险的 0.86 倍(95%CI: 0.78~0.99); 而当采用传统回归校正方法时, 得到的结果为 0.74(95%CI: 0.48~1.14), 显然其 95%CI 更宽, 更易产生有偏估计结果^[8]。

2. IPW: 在没有未测量混杂(no unmeasured confounding)假设下, 给每个观测者分配权重来重新获得一个混杂在组间均衡分布的伪人群, 在该人群中拟合处理因素和结局间的边际结构模型(marginal structure models, MSMs)即可得到因果效应无偏估计值^[9-10]。IPW 的计算包括不稳定权重:

$$W = \prod_{t=0}^k \frac{1}{f[X_t|\bar{X}_{t-1}, \bar{C}_t]}$$

和稳定权重: $SW = \prod_{t=0}^k \frac{f[X_t|\bar{X}_{t-1}]}{f[X_t|\bar{X}_{t-1}, \bar{C}_t]}$ (例 1 中: $W = \frac{1}{f[x_1|x_0, c_1]}, SW = \frac{f[x_1|x_0]}{f[x_1|x_0, c_1]}$)。之所以称为不稳定权重是由于, 可能有受试者在 \bar{X}_{k-1}, \bar{C}_k 条件下被分配进入暴露组的估计概率极低, 致其 W 极大, 使因果效应估计量的标准误增大, 甚至产生偏倚^[11-12]。SW 将分子替换为个体实际接受暴露的边际概率, 故基于 SW 的因果效应估计值标准误显著降低^[11]。Danaei 等^[3]在评估他汀类药物在冠心病一级预防中的作用时, 使用 IPW 和 MSMs 分析, 发现服用他汀类药物能够降低冠心病发病风险 $HR=0.79$ (95%CI: 0.41~1.41), 而用传统回归校正方法得到的结论为 $HR=1.31$ (95%CI: 1.04~1.66), 这与 RCT 试验得到的结论 $HR=0.72$ (95%CI: 0.65~0.79) 相悖^[13]。

3. G 估计:将暴露效应定义为比较同一个体或一组受试者接受暴露和非暴露的潜在结局^[5,14]。首先,在先前暴露状态和混杂水平下,第 t 次随访暴露与非暴露的效应估计值为 $E(Y_i^{\bar{x}_i-1, x_i} - Y_i^{\bar{x}_i-1, 0} | \bar{X}_i = \bar{x}_i, \bar{C}_i = \bar{c}_i)$, 随后采用结构嵌套模型 (structural nested mean models, SNMs) 来参数化处理组受试者的平均因果效应: $g\{E(Y_i^{\bar{x}_i-1, x_i} | \bar{X}_i = \bar{x}_i, \bar{C}_i = \bar{c}_i)\} = g\{E(Y_i^{\bar{x}_i-1, 0} | \bar{X}_i = \bar{x}_i, \bar{C}_i = \bar{c}_i)\} + \gamma(\bar{x}_i, \bar{c}_i; \psi)$, 其中 $g(\cdot)$ 为链接函数 (如 identity 链接、log 链接、logit 链接), $\gamma(\bar{x}_i, \bar{c}_i; \psi)$ 是波动函数 (如线性波动函数), 其满足 $\gamma(\bar{x}_i, \bar{c}_i; \psi) = 0 \begin{cases} \text{if } x_i = 0 \\ \text{if } \psi = 0 \end{cases}$, 只有 ψ 是需要估计的未知有限维参数^[15-16]。以例 1 为例, 采用 identity 链接和线性波动函数, 有: $E(Y_i^{x_0, x_1} - Y_i^{x_0, 0} | x_0, c_1, x_1) = x_1(\psi_1 + \psi_2 x_0 + \psi_3 c_1 + \psi_4 x_0 c_1)$ 和 $E(Y_i^{x_0, 0} - Y_i^{0, 0} | x_0) = \psi_0 x_0$ 。根据条件可交换性假设, 每个时间点的暴露与结局间的条件协方差为 0, 则可估计出系数 ψ 。Shakiba 等^[17] 在一项量化肥胖和冠心病发病风险间因果关系的研究中使用 G 估计控制时依混杂 (包括发生高血压/糖尿病/脑卒中、血清胆固醇/TG、吸烟饮酒状态), 得到腹型肥胖患病风险更高 $HR=1.65(95\%CI: 1.35\sim 1.92)$ 的结论, 而仅调整了基线混杂的传统 Cox 回归方法得到的结论为 $HR=0.93(95\%CI: 0.77\sim 1.12)$ 。

三、方法比较

本研究将从原理、适用范围、估计的有效性和稳定性、计算复杂程度等方面比较 3 种方法的差异, 并用表 1 简要概括了 3 种方法的优缺点。

1. 原理: 参数 g-formula 可以看作标准化方法在多个随访时间点的拓展, 通过预测缺失的反事实结局, 比较在不同的假设干预下人群的结局是否存在差异; IPW 通过对调查人群加权, 获得一个不受混杂因素影响的、伪随机化的人群^[1,18]。这 2 种方法是直观的, 相对于 G 估计更好理解, 但是通过 G 估计得到的因果效应结论更容易外推到其他人中^[18]。

2. 零悖论: 因果推断的零假设为暴露和结局间不存在因果关联, 当零假设为真时错误地拒绝零假设被称为零悖论^[19]。参数 g-formula 需要估计给定协变量和处理因素的平均结局效应 $E[Y | \bar{X} = \bar{x}, \bar{C} = \bar{c}]$, 和给定历史协变量和处理因素时该次随访协变量的分布 $f[(C_i = c_i | \bar{X}_{i-1} = \bar{x}_{i-1}, \bar{C}_{i-1} = \bar{c}_{i-1})]$ 。当先

前的暴露对之后的混杂有非零影响时, 基于参数 g-formula 得到的结论会与零假设不一致, 即产生了零悖论^[1,18]。而其他 2 种估计方法不存在零悖论, 零假设对 IPW 而言, 意味着加权后的 MSMs 中估计的暴露系数为 0; 零假设在 G 估计中, 对应着波动函数的有限维参数 $\psi = 0$ ^[18]。因此零悖论也被称为 g-null 悖论, 其只会发生在控制时依混杂的因果推断中。避免零悖论的方法是仅考虑已知对结局有影响的暴露, 或基于零假设不成立的先验信息^[20]。

3. 估计的有效性和稳定性: 参数 g-formula 估计的有效性和稳定性主要依赖于模型是否被正确设定, 尤其是在不能排除零假设的前提下, 而一旦结局模型 (以结局为预测变量建模) 和混杂模型 (以混杂为预测变量建模) 均被正确设定, 参数 g-formula 的效率和稳定性都很高^[18]。基于 IPW 加权建立伪人群方法的效率和稳定性都低于模型正确设定时的参数 g-formula, 且随着混杂维数的增加, IPW 的变异性增加, 会给数据集中少数受试者非常大的权重, 导致其低效和不稳定性进一步增加^[18,21]。G 估计的效率和稳定性通常介于参数 g-formula 和 IPW 之间, 与 MSMs 相比, SNMs 可以纳入暴露和混杂的交互项, 这可能是 G 估计的效率高于 IPW 的原因^[18]。

4. 适用范围: 参数 g-formula 可用于估计较复杂的联合暴露 (例如吸烟和每日 ≥ 30 min 的体育锻炼对健康的联合效应), 但当暴露是连续变量/具有多个水平的分类变量/具有多个随访时间点的二分类变量, 且无感兴趣的暴露水平时, 参数 g-formula 要估计每一种可能处理组合的平均结局, 当有太多可能的假设处理水平可供比较时, 计算量很大, 模型解释变得很困难, 同时也会增加模型误设的概率^[1,18]。当某一混杂水平下所有受试者均为暴露或均为非暴露, 即违背了正性假设 (给定的混杂水平下, 没有接受某个处理水平概率为零的受试者) 时, 不能用 IPW 方法^[1]; IPW 难以处理连续性暴露, 且由于 MSMs 是边际的, 因此无法纳入暴露和混杂的交互项。G 估计可以纳入暴露和混杂的交互项^[1,19]。参数 g-formula 和 IPW 均可以用于连续性、分类的结局变量以及生存分析中; G 估计在生存分析中采用的模型为基于 SNMs 的结构嵌套累计失效时间模型^[22]。

上述 G 方法对随访次数均不做限制, 但随着随访次数增加, 失访比例可能增大。虽然 G 方法能够

将失访视为另一种可能的治疗方式,从而通过二次加权均衡失访偏倚,但其需要假设失访是完全随机发生的。而在实际研究中,失访常与暴露或结局存在一定关联,因此在设计阶段采取一系列措施提高随访率是十分必要的,例如纳入主动性较高更易跟踪随访的受试者、设计的问卷尽可能易于理解和完成、确保在所有研究小组中保持高水平的随访等。此外,研究者应结合临床实际、文献报道、或进行预调查,判断研究纳入的暴露和协变量是否随时间变化,即判断是否存在需要控制的时依混杂,虽然在理论上 G 方法可以估计不随时间变化的暴露因素与结局间的因果效应,但其复杂程度要远高于传统的匹配、分层、倾向性评分等方法。本研究仅讨论了包含单个随访时间点结局事件(即随访终点结局事件 Y)的因果效应估计,但研究者可根据实际研究问题将 G 方法拓展到包含了多个随访时间点结局事件(Y_t)的因果效应估计中^[23]。

5. 计算和程序实现:参数 g-formula 计算较复杂,因为涉及了蒙特卡洛模拟和 bootstrap,可用 Stata 软件中的“gformula”命令或 R 软件中的“gformula”软件包实现;IPW 可在大多数统计软件中轻松实现,可参考 R 软件包“ipw”;G 估计计算较复杂,可通过 Stata 软件中的“stgest”命令、SAS 软件中的“SURV”宏程序或 R 软件中的“DTRreg”软件包实现。

四、实例分析

以 2016 年 Naimi 等^[5]的研究报告中关于 ART 治疗对 HIV 感染者 CD4 计数影响的前瞻性研究数据为例,比较传统回归分析方法与 3 种 G 方法估计效应的准确性(表 2)。该研究调查了基线(t_0)和第一次随访(t_1)时的 ART 状态,分别用 X_0 、 X_1 表示($X_i=$

1 为接受 ART、 $X_i=0$ 为未接受 ART);基线的 HIV 病毒载量(C_0)均 >200 拷贝/ml,在第一次随访时第二次治疗前测得的 HIV 病毒载量(C_1)为时依混杂($C_i=0$ 表示 ≤ 200 拷贝/ml, $C_i=1$ 则相反),随访结束时测得的 CD4 计数(个/ μl)为结局 Y (因果图可参考图 1)。分别用传统线性回归方法,包括只考虑基线治疗状态 X_0 和同时考虑两次治疗状态 X_0 和 X_1 的校正了 C_i 的线性回归模型,以及 3 种可以控制时依混杂 C_i 的 G 方法,估计 ART 与 CD4 计数间的因果关系。表 2 中 β_0 、 β_1 分别代表 X_0 、 X_1 的回归系数, $\Delta\hat{y}_{\text{mean}}$ 代表基线和第一次随访时均接受 ART($X_0=1$ 且 $X_1=1$)与均未接受 ART($X_0=0$ 且 $X_1=0$)分别对应的 Y 均值的差异。表 2 中显示, G 方法估计的 β_0 和 β_1 很接近,即在 t_0 和 t_1 时的 ART 对 CD4 计数的影响效应相同,而传统回归方法估计的 β_0 和 β_1 有较大差异;且 G 方法估计的 $\Delta\hat{y}_{\text{mean}}$ 均为 50.0 个/ μl ,即两次均接受 ART 的患者 CD4 计数平均增加 50.0 个/ μl 。传统回归方法则低估了 ART 与 CD4 计数间的因果效应,尤其是在只关注基线暴露对结局的影响效应时。

五、总结

在纵向研究中,常常存在受过去暴露影响的时依混杂,但在实践中,通常被忽略或者直接使用传统方法进行调整,得到有偏因果效应估计值。G 方法是在对暴露和混杂进行重复测量的纵向数据中,调整受过去暴露影响的时依混杂的方法。在一项关于 2000–2016 年控制时依混杂的相关研究文献的系统回顾中,使用 G 方法的文献虽然在逐年增加,但也仅占总文献量的 20%,且其中有 60.9% 集中在 HIV、心肺疾病、肾病或精神健康领域,80.3% 采用的方法是 IPW^[24]。因此,本研究可为不同领域的研究人员在处理纵向数据的因果推断实践中提

表 1 G 方法的优缺点

方法	优点	缺点
参数 g-formula	可估计联合暴露;效率和稳定性高	计算复杂、程序复杂;易受模型误设的影响;只能在零假设被认为不成立的前提下使用
IPW 结合 MSMs	易于理解,容易实现	不满足正性假设时不能用该方法;估计不稳定;不能处理交互
G 估计	不满足正性假设时可以使用;可以处理交互	计算复杂、程序复杂;方法较难、解释不易

表 2 不同方法估计的参数

估计值	传统线性回归模型		g-formula	IPW	G 估计	真实值
	只纳入基线 X_0	同时纳入 X_0 、 X_1				
β_0	18.00	17.99	25.02	25.02	24.98	25.00
β_1	-	24.97	24.99	25.04	24.96	25.00
$\Delta\hat{y}_{\text{mean}}$	27.10	42.60	50.00	50.00	50.00	50.00

注:传统线性回归模型: $y = \alpha_0 + \beta_0 x_0 + \alpha_1 c_1$; $y = \alpha_0 + \beta_0 x_0 + \beta_1 x_1 + \beta_{01}(x_0 * x_1) + \alpha_1 c_1$; G 方法的估计模型参见方法和原理

供参考,由于G方法包括的3种方法的适用范围、估计的有效性和稳定性有差异,研究者应根据实际情况,选择适合的G方法控制时依混杂,从而获得无偏因果效应估计值。

利益冲突 所有作者均声明不存在利益冲突

参 考 文 献

- [1] Mansournia MA, Etminan M, Danaei G, et al. Handling time varying confounding in observational research[J]. *BMJ*, 2017, 359:j4587. DOI:10.1136/bmj.j4587.
- [2] Mortimer KM, Neugebauer R, van der Laan M, et al. An application of model-fitting procedures for marginal structural models[J]. *Am J Epidemiol*, 2005, 162(4): 382-388. DOI:10.1093/aje/kwi208.
- [3] Danaei G, Rodríguez LAG, Cantero OF, et al. Observational data for comparative effectiveness research: an emulation of randomised trials of statins and primary prevention of coronary heart disease[J]. *Stat Methods Med Res*, 2011, 22(1):70-96. DOI:10.1177/0962280211403603.
- [4] Williamson T, Ravani P. Marginal structural models in clinical research: when and how to use them? [J]. *Nephrol Dial Transplant*, 2017, 32 Suppl 2:ii84-90. DOI:10.1093/ndt/gfw341.
- [5] Naimi AI, Cole SR, Kennedy EH. An introduction to G methods[J]. *Int J Epidemiol*, 2016, 46(2): 756-762. DOI: 10.1093/ije/dyw323.
- [6] Robins JM. A new approach to causal inference in mortality studies with a sustained exposure period-application to control of the healthy worker survivor effect[J]. *Math Model*, 1986, 7(9/12): 1393-1512. DOI: 10.1016/0270-0255(86)90088-6.
- [7] Westreich D, Cole SR, Young JG, et al. The parametric g-formula to estimate the effect of highly active antiretroviral therapy on incident AIDS or death[J]. *Stat Med*, 2012, 31(18):2000-2009. DOI:10.1002/sim.5316.
- [8] 吴诗蓝,周价,李逊,等. Parametric g-formula方法在因果分析中的应用[J]. *中华流行病学杂志*, 2019, 40(10): 1310-1313. DOI: 10.3760/cma.j.issn.0254-6450.2019.10.025.
Wu SL, Zhou J, Li X, et al. Application of parametric g-formula in causal analysis[J]. *Chin J Epidemiol*, 2019, 40(10): 1310-1313. DOI: 10.3760/cma.j.issn.0254-6450.2019.10.025.
- [9] Robins JM, Hernán MÁ, Brumback B. Marginal structural models and causal inference in epidemiology[J]. *Epidemiology*, 2000, 11(5): 550-560. DOI: 10.1097/00001648-200009000-00011.
- [10] Imai K, Ratkovic M. Robust estimation of inverse probability weights for marginal structural models[J]. *J Am Stat Assoc*, 2015, 110(511):1013-1023. DOI:10.1080/01621459.2014.956872.
- [11] 郭威. 基于统计学习的逆概率加权方法研究及其在医学中的应用[D]. 上海:中国人民解放军海军军医大学, 2018. DOI:CNKI:CDMD:1.1018.084023.
- [12] Rubin DB. On principles for modeling propensity scores in medical research[J]. *Pharmacoepidemiol Drug Safe*, 2004, 13(12):855-857. DOI:10.1002/pds.968.
- [13] Taylor F, Huffman MD, Macedo AF, et al. Statins for the primary prevention of cardiovascular disease in the Elderly[J]. *Cochrane Database Syst Rev*, 2013, 2013: CD004816. DOI:10.1002/14651858.CD004816.pub5.
- [14] Picciotto S, Neophytou AM. G-estimation of structural nested models: recent applications in two subfields of epidemiology[J]. *Curr Epidemiol Rep*, 2016, 3(3): 242-251. DOI:10.1007/s40471-016-0081-9.
- [15] Robins JM. The analysis of randomized and non-randomized AIDS treatment trials using a new approach to causal inference in longitudinal studies[M]// Sechrest L, Freeman H, Mulley A. Health service research methodology: a focus on AIDS. Washington: U. S. Public Health Service, National Center for Health Services Research, 1989:113-159.
- [16] 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.
- [17] Shakiba M, Mansournia MA, Salari A, et al. Accounting for time-varying confounding in the relationship between obesity and coronary heart disease: analysis with G-estimation: the ARIC study[J]. *Am J Epidemiol*, 2018, 187(6):1319-1326. DOI:10.1093/aje/kwx360.
- [18] Daniel RM, Cousens SN, de Stavola BL, et al. Methods for dealing with time-dependent confounding[J]. *Stats Med*, 2013, 32(9):1584-1618. DOI:10.1002/sim.5686.
- [19] Robins JM, Hernán MA. Estimation of the causal effects of time-varying exposures[M]//Fitzmaurice G, Davidian M, Verbeke G, et al. Longitudinal Data Analysis. New York: Chapman and Hall/CRC Press, 2009: 553-599. DOI: 10.1201/9781420011579.ch23.
- [20] Wilsgaard T, Vangen-Lønne AM, Mathiesen E, et al. Hypothetical interventions and risk of myocardial infarction in a general population: application of the parametric g-formula in a longitudinal cohort study—the Tromsø Study[J]. *BMJ Open*, 2020, 10(5): e035584. DOI: 10.1136/bmjopen-2019-035584.
- [21] Kang JDY, Schafer JL. Demystifying double robustness: a comparison of alternative strategies for estimating a population mean from incomplete data (with discussion and rejoinder) [J]. *Stat Sci*, 2007, 22(4): 523-539. DOI: 10.1214/07-STS227.
- [22] Vansteelandt S, Joffe M. Structural nested models and g-estimation: the partially realized promise[J]. *Stat Sci*, 2014, 29(4):707-731. DOI:10.1214/14-STS493.
- [23] Clare PJ, Dobbins TA, Mattick RP. Causal models adjusting for time-varying confounding—a systematic review of the literature[J]. *Int J Epidemiol*, 2019, 48(1): 254-265. DOI: 10.1093/ije/dyy218.