Supporting Information for "Survival mediation analysis with the death-truncated mediator: The completeness of the survival mediation parameter"

An-Shun Tai[1*], Chun-An Tsai[1**], Sheng-Hsuan Lin[1***]

1. Institute of Statistics, National Chiao Tung University, Hsin-Chu, Taiwan. 1001 University Road, Hsinchu, Taiwan 300

***Corresponding author**
Sheng-Hsuan Lin, MD, ScD
Institute of Statistics, National Chiao Tung University, Hsin-Chu, Taiwan
1001 University Road,
Hsinchu, Taiwan 300
Cell: 886-3-5712121 ext 56822
E-mail: shenglin@stat.nctu.edu.tw

Web Appendix A

1. Mediation analysis for survival time in the presence of previous death

For survival time T, TE, NDE, and NIE are traditionally defined on risk ratio scale through log hazard function as follows:

$\text{TE} = log(\lambda_{T(1,M(1))}(t)) - log(\lambda_{T(0,M(0))}(t))$,

$\text{NDE} = log(\lambda_{T(1,M(0))}(t)) - log(\lambda_{T(0,M(0))}(t))$, and

$\text{NIE} = log(\lambda_{T(1,M(1))}(t)) - log(\lambda_{T(1,M(0))}(t))$,

where $\lambda_{T(1,M(1))}(t)$ is hazard function with respect to survival time T.

Similar to the causal effects defined on survival status, the regular causal effects defined on survival time is the lack of completeness in the presence of previous death. Among the protected group P_P (i.e. $Y_p(0) = 0$ and $Y_p(1) = 1$), the counterfactual outcome of survival time $T(1, Y_p(1), M(0, Y_p(0)))$, which equal to $T(1, y_p = 1, M(0, y_p = 0))$, cannot be defined since it is the hypothetical status suppose it is without previous death but intervened by $M(0, y_p = 0)$ which is the mediator truncated by previous death. While $T(0, Y_p(0), M(0, Y_p(0)))$ is also intervened by the death-truncated mediator, it is the survival time suppose this individual is subject to a previous death $(y_p = 0)$ which implies that

$T(0, Y_p(0), M(0, Y_p(0)))$is always equal to zero. Consequently, we can define TE among P_P but not NDE and NIE. Following the argumentation above, we found that among the group P_H, all counterfactual values of Y are either well-defined or zero. Among the group P_D, all counterfactuals are zero.

2. Completeness for total direct effect and pure indirect effect in the presence of protective effect

As the alternative for effect decomposition, the total direct effect (TDE) and pure indirect effect (PIE) are defined as

TDE $= Y\left(1, Y_p(1), M(1, Y_p(1))\right) - Y\left(0, Y_p(0), M(1, Y_p(1))\right)$, and

PIE $= Y\left(0, Y_p(0), M(1, Y_p(1))\right) - Y\left(0, Y_p(0), M(0, Y_p(0))\right)$.

By assuming the protective effect of the exposure, we subsequently discuss the completeness for the definitions of TDE and PIE under P_S, P_P, and P_D, respectively. Following the argumentation for NIE and NDE in the manuscript, it is clearly known that TDE and PIE are well-defined among P_S and P_D. Here, we shown that TDE and PIE are also well-defined among P_P. Among the protected group P_P (i.e. $Y_p(0) = 0$ and $Y_p(1) = 1$), by definition, the counterfactual outcome $Y(0, Y_p(0), M(1, Y_p(1)))$ and $Y(0, Y_p(0), M(0, Y_p(0)))$, which equal to $Y(0, y_p = 0, M(1, y_p = 1))$ and $Y(0, y_p = 0, M(0, y_p = 0))$, are always equal to zero. Moreover, $Y(1, Y_p(1), M(1, Y_p(1))) = Y(1, y_p = 1, M(1, y_p = 1))$ can be defined. Therefore, TDE and PIE are well-defined among protected group P_P. Consequently, we concluded that the definitions of TDE and PIE are completeness in the presence of protective effect.

3. The completeness of $\mathbf{NDE_{dt}}$ and $\mathbf{NIE_{dt}}$

Based on the proposed definitions of (3), we proved $\mathrm{NDE_{dt}}$ and $\mathrm{NIE_{dt}}$ are completeness. Among the always-survivor group P_S (i.e. $Y_p(1) = 1$ and $Y_p(0) = 1$), we have

$$\begin{aligned}\mathrm{NDE_{dt}} &= Y\left(1, Y_p(1), M(0, Y_p(1))\right) - Y\left(0, Y_p(0), M(0, Y_p(0))\right) \\ &= Y\left(1,1, M(0,1)\right) - Y\left(0,0, M(0,1)\right) \text{ and}\end{aligned}$$

$$\begin{aligned}\mathrm{NIE_{dt}} &= Y\left(1, Y_p(1), M(1, Y_p(1))\right) - Y\left(1, Y_p(1), M(0, Y_p(1))\right) \\ &= Y\left(1,1, M(1,1)\right) - Y\left(1,0, M(0,1)\right).\end{aligned}$$

Among the harmed group P_H (i.e. $Y_p(1) = 0$ and $Y_p(0) = 1$), we have

$$\mathrm{NDE_{dt}} = Y\big(1,0,M(0,0)\big) - Y\big(0,1,M(0,1)\big)$$

$$= -Y\big(0,1,M(0,1)\big) \text{ and}$$

$$\mathrm{NIE_{dt}} = Y\big(1,0,M(1,0)\big) - Y\big(1,0,M(0,0)\big) = 0.$$

Among the protected group P_P (i.e. $Y_p(1) = 1$ and $Y_p(0) = 0$), we have

$$\mathrm{NDE_{dt}} = Y\big(1,1,M(0,1)\big) - Y\big(0,0,M(0,0)\big)$$

$$= Y\big(1,1,M(0,1)\big) \text{ and}$$

$$\mathrm{NIE_{dt}} = Y\big(1,1,M(1,1)\big) - Y\big(1,1,M(0,1)\big).$$

Among the doomed group the P_D(i.e. $Y_p(1) = 0$ and $Y_p(0) = 0$), we have

$$\mathrm{NDE_{dt}} = Y\big(1,0,M(0,0)\big) - Y\big(0,0,M(0,0)\big) = 0 \text{ and}$$

$$\mathrm{NIE_{dt}} = Y\big(1,0,M(1,0)\big) - Y\big(1,0,M(0,0)\big) = 0.$$

Since $M(1,1)$ and $M(0,1)$ are always meaningful, $\mathrm{NDE_{dt}}$ and $\mathrm{NIE_{dt}}$ are well-defined among four groups. Consequently, the proposed formulations of death-truncated causal effects are completeness.

4. Proof of *Theorem 1*

Among the group P_S:

$$\psi(1,1) = E\Big(Y\big(1,Y_p(1),M(1,Y_p(1))\big)\Big) = E\Big(Y\big(1,1,M(1,1)\big)\Big) = \phi(1,1)$$

$$\psi(0,0) = E\Big(Y\big(0,Y_p(0),M(0,Y_p(0))\big)\Big) = E\Big(Y\big(0,0,M(0,0)\big)\Big) = \phi(0,0)$$

$$\psi(1,0) = E\Big(Y\big(1,Y_p(1),M(0,Y_p(0))\big)\Big) = E\Big(Y\big(1,1,M(0,1)\big)\Big)$$

$$= E\Big(Y\big(1,Y_p(1),M(0,Y_p(1))\big)\Big) = \phi(1,0)$$

$$\therefore \mathrm{NDE} = \psi(1,0) - \psi(0,0) = \phi(1,0) - \phi(0,0) = \mathrm{NDE_{dt}}$$

$$\mathrm{NIE} = \psi(1,1) - \psi(1,0) = \phi(1,1) - \phi(1,0) = \mathrm{NIE_{dt}}$$

Among the group P_H:

$$\psi(1,1) = E\Big(Y\big(1,Y_p(1),M(1,Y_p(1))\big)\Big) = E\Big(Y\big(1,0,M(1,0)\big)\Big) = \phi(1,1)$$

$$\psi(0,0) = E\Big(Y\big(0,Y_p(0),M(0,Y_p(0))\big)\Big) = E\Big(Y\big(0,1,M(0,1)\big)\Big) = \phi(0,0)$$

$$\psi(1,0) = E\Big(Y\big(1,Y_p(1),M(0,Y_p(0))\big)\Big) = E\Big(Y\big(1,0,M(0,1)\big)\Big)$$

$$= E\left(Y\big(1, Y_p(1), M(0, Y_p(1))\big)\right) = \phi(1,0)$$

$\therefore$ NDE $= \psi(1,0) - \psi(0,0) = \phi(1,0) - \phi(0,0) = \text{NDE}_{\text{dt}}$

NIE $= \psi(1,1) - \psi(1,0) = \phi(1,1) - \phi(1,0) = \text{NIE}_{\text{dt}}$

Among the group P_p:

$$\psi(1,1) = E\left(Y\big(1, Y_p(1), M(1, Y_p(1))\big)\right) = E\left(Y\big(1,0, M(1,0)\big)\right) = \phi(1,1)$$

$$\psi(0,0) = E\left(Y\big(0, Y_p(0), M(0, Y_p(0))\big)\right) = E\left(Y\big(0,0, M(0,0)\big)\right) = \phi(0,0)$$

$$\psi(1,0) = E\left(Y\big(1, Y_p(1), M(0, Y_p(0))\big)\right) = E\left(Y\big(1,0, M(0,0)\big)\right)$$

$$= E\left(Y\big(1, Y_p(1), M(0, Y_p(1))\big)\right) = \phi(1,0)$$

$\therefore$ NDE $= \psi(1,0) - \psi(0,0) = \phi(1,0) - \phi(0,0) = \text{NDE}_{\text{dt}}$

NIE $= \psi(1,1) - \psi(1,0) = \phi(1,1) - \phi(1,0) = \text{NIE}_{\text{dt}}$

Thus, within three survival groups, the death-truncated causal effects, NDE_{dt} and NIE_{dt}, are identical to regular causal effects, NDE and NIE, respectively.

Web Appendix B

1. NPSEM

$\varepsilon_A \coprod \varepsilon_Y \Rightarrow A \coprod \left(Y_p(a), Y(a,1,m)\right)$ (Assumption 1)

$\varepsilon_M \coprod \varepsilon_Y \Rightarrow M \coprod Y(a,1,m)|A = a,\ Y_p = 1$ (Assumption 2)

$\varepsilon_M \coprod \varepsilon_A \Rightarrow M\left(a^*, Y_p = 1\right) \coprod A$ (Assumption 3)

$\varepsilon_M \coprod \varepsilon_Y \Rightarrow M\left(a^*, Y_p = 1\right) \coprod \left(Y_p(a), Y(a,1,m)\right)$ (Assumption 4)

$\varepsilon_M \coprod \varepsilon_Y \Rightarrow M\left(a^*, Y_p = 1\right) \coprod Y_p|A$ (Assumption 5)

2. Proof of *Theorem 2*

$$\phi(a,a^*) \equiv \mathrm{E}\left(Y\left(a, Y_p(a), M(a^*, Y_p(a))\right)\right)$$

$$= \mathrm{E}\left(Y\left(a, Y_p(a), M(a^*, Y_p(a))\right)\middle|Y_p(a) = 1\right)P\left(Y_p(a) = 1\right)$$

$$+\mathrm{E}\left(Y\left(a, Y_p(a), M(a^*, Y_p(a))\right)\middle|Y_p(a) = 0\right)P\left(Y_p(a) = 0\right)$$

$$= \mathrm{E}\left(Y\left(a, y_p = 1, M(a^*, y_p = 1)\right)|Y_p(a) = 1\right)P\left(Y_p(a) = 1\right) \quad (\because \boldsymbol{Y(Y_p(a) = 0) = 0})$$

$$= \mathrm{E}\left(Y\left(a, y_p = 1, M(a^*, y_p = 1)\right)Y_p(a)\right) \quad (\because \textbf{B is binary} \Rightarrow \boldsymbol{E(BY) = E(Y|B = 1)P(B = 1)})$$

$$= \int_c E\left(Y_p(a)Y(a,1,m)|C = c\right)f(c)dc$$

$$= \int_{m,c} E\left(Y_p(a)Y(a,1,m)|M(a^*, y_p = 1) = m, C = c\right)f\left(M(a^*, y_p = 1) = m\right)f(c)dmdc$$

$$= \int_{m,c} E\left(Y_p(a)Y(a,1,m)|C = c\right)f\left(M\left(a^*, Y_p = 1\right) = m\right)f(c)dmdc \quad (\textbf{by } \boldsymbol{Assumption\ 4.4})$$

$$= \int_{m,c} E\left(Y_p(a)Y(a,1,m)|A = a, C = c\right)f\left(M(a^*, Y_p = 1) = m\right)f(c)dmdc$$

$$(\textbf{by } \boldsymbol{Assumption\ 4.1})$$

$$= \int_{m,c} E\left(Y_pY(a,1,m)|A = a, C = c\right)f\left(M(a^*, Y_p = 1) = m\right)f(c)dmdc \quad (\textbf{by } \boldsymbol{consistency})$$

$$= \int_{m,c} E\left(Y(a,1,m)|A = 1, Y_p = 1, C = c\right)f\left(Y_p = 1\middle|A = a, C = c\right)$$

$$f\left(M\left(a^*, Y_p = 1\right) = m\right) f(c)dmdc$$

$$= \int_{m,c} E\left(Y(a,1,m)|A = a, Y_p = 1, M = m, C = c\right)f\left(Y_p = 1\middle|A = a, C = c\right)$$

$$f\left(M\left(a^*, Y_p = 1\right) = m\right)f(c)dmdc$$

$$(\textbf{by } \boldsymbol{Assumption\ 4.2})$$

$$= \int_{m,c} E\left(Y(a,1,m)|A = a, Y_p = 1, M = m, C = c\right)f\left(Y_p = 1\middle|A = a, C = c\right)$$

$$f\left(M\left(a^*, Y_p = 1\right) = m\middle|A = a^*, Y_p = 1, C = c\right)f(c)dmdc$$

$$(\textbf{by } \boldsymbol{Assumptions\ 4.3\ \&4.5})$$

$$= \int_{m,c} E\left(Y \middle| A = a, Y_p = 1, M = m\right)\ f\left(M = m \middle| A = a^*, Y_p = 1, C = c\right)$$

$$f\left(Y_p = 1 \middle| A = a, C = c\right) f(c) dmdc$$

(by ***consistency***)

Thus, $\phi(a, a^*)$ can be identified as $Q(a, a^*)$ for any a and a^*.

3. Proof of *Theorem 3*

$$\psi(a, a^*) \equiv \mathrm{E}\left(Y\left(a, Y_p(a), M(a^*, Y_p(a^*))\right)\right)$$

$$= \mathrm{E}\left(Y\left(a, Y_p(a), M(a^*, Y_p(a^*))\right) \middle| Y_p(a) = 1\right) P\left(Y_p(a) = 1\right)$$

$$+ \mathrm{E}\left(Y\left(a, Y_p(a), M(a^*, Y_p(a^*))\right) \middle| Y_p(a) = 0\right) P\left(Y_p(a) = 0\right)$$

$$= \mathrm{E}\left(Y\left(a, Y_p(a) = 1, M(a^*, Y_p(a^*) = 1)\right) | Y_p(a) = 1\right) P\left(Y_p(a) = 1\right) \quad (\because \boldsymbol{Y(Y_p(a) = 0) = 0})$$

$$= \mathrm{E}\left(Y\left(a, y_p = 1, M(a^*, y_p = 1)\right) | Y_p(a) = 1\right) P\left(Y_p(a) = 1\right)$$

(∵ **decreasing monotonicity assumption for $\boldsymbol{Y_p}$ and $\boldsymbol{a \geq a^*}$**
$\boldsymbol{\Rightarrow Y_p(a^*) \geq Y_p(a) = 1}$
$\boldsymbol{\Rightarrow Y_p(a^*) = Y_p(a) = 1}$)

$$= \mathrm{E}\left(Y\left(a, y_p = 1, M(a^*, y_p = 1)\right) Y_p(a)\right) \quad (\because \textbf{B is binary} \Rightarrow \boldsymbol{E(BY) = E(Y|B = 1)P(B = 1)})$$

$$= \int_c E\left(Y_p(a) Y(a, 1, m) | C = c\right) f(c) dc$$

$$= \int_{m,c} E\left(Y_p(a) Y(a, 1, m) | M(a^*, y_p = 1) = m, C = c\right) f\left(M(a^*, y_p = 1) = m\right) f(c) dmdc$$

$$= \int_{m,c} E\left(Y_p(a) Y(a, 1, m) | C = c\right) f\left(M\left(a^*, Y_p = 1\right) = m\right) f(c) dmdc \quad (\textbf{by } \boldsymbol{Assumption\ 4.4})$$

$$= \int_{m,c} E\left(Y_p(a) Y(a, 1, m) | A = a, C = c\right) f\left(M(a^*, Y_p = 1) = m\right) f(c) dmdc$$

(by ***Assumption* 4. 1**)

$$= \int_{m,c} E\left(Y_p Y(a, 1, m) | A = a, C = c\right) f\left(M(a^*, Y_p = 1) = m\right) f(c) dmdc \quad (\textbf{by } \boldsymbol{consistency})$$

$$= \int_{m,c} E\left(Y(a, 1, m) | A = 1, Y_p = 1, C = c\right) f\left(Y_p = 1 \middle| A = a, C = c\right)$$

$$f\left(M\left(a^*, Y_p = 1\right) = m\right)\ f(c) dmdc$$

$$= \int_{m,c} E\left(Y(a, 1, m) | A = a, Y_p = 1, M = m, C = c\right) f\left(Y_p = 1 \middle| A = a, C = c\right)$$

$$f\left(M\left(a^*, Y_p = 1\right) = m\right) f(c) dmdc$$

(by ***Assumption* 4. 2**)

$$= \int_{m,c} E\left(Y(a, 1, m) | A = a, Y_p = 1, M = m, C = c\right) f\left(Y_p = 1 \middle| A = a, C = c\right)$$

$$f\left(M\left(a^*, Y_p = 1\right) = m \middle| A = a^*, Y_p = 1, C = c\right) f(c) dmdc$$

(by ***Assumptions* 4. 3** &**4. 5**)

$$= \int_{m,c} E\left(Y \middle| A = a, Y_p = 1, M = m\right) \ f\left(M = m \middle| A = a^*, Y_p = 1, C = c\right)$$

$$f\left(Y_p = 1 \middle| A = a, C = c\right) f(c) dmdc$$

(**by** $\boldsymbol{consistency}$)

Thus, $\psi(a, a^*)$ can be identified as $Q(a, a^*)$ for $a \geq a^*$.

4. Proof of *Lemma 1*

$$Q_T(a, a^*) = \ log\, \lambda\left(T\left(a, Y_p(a), M\left(a^*, Y_p(a)\right)\right); t\right)$$

$$= log \frac{f\left(T\left(a, Y_p(a), M\left(a^*, Y_p(a)\right)\right); t\right)}{S\left(T\left(a, Y_p(a), M\left(a^*, Y_p(a)\right)\right); t\right)}$$

$$= \ log \frac{\int_{c,m} f\left(T = t \middle| a, m, c, y_p = 1\right) f\left(y_p = 1 \middle| a, c\right) f\left(M = m \middle| a^*, c, y_p = 1\right) f(c) dmdc}{\int_{c,m} S\left(T = t \middle| a, m, c, y_p = 1\right) f\left(y_p = 1 \middle| a, c\right) f\left(M = m \middle| a^*, c, y_p = 1\right) f(c) dmdc}$$

$$= \ log \frac{\int_{c,m} e^{-\Lambda\left(t \middle| a, m, c, y_p = 1\right)} \lambda\left(t \middle| a, m, c, y_p = 1\right) f\left(y_p = 1 \middle| a, c\right) f\left(M = m \middle| a^*, c, y_p = 1\right) f(c) dmdc}{\int_{c,m} e^{-\Lambda\left(t \middle| a, m, c, y_p = 1\right)} f\left(y_p = 1 \middle| a, c\right) f\left(M = m \middle| a^*, c, y_p = 1\right) f(c) dmdc}$$

($\because S(t) = exp\left(-\Lambda(t)\right)$ and $f(t) = S(t)\lambda(t) = \ exp\left(-\Lambda(t)\right)\lambda(t)$)

$$\approx \ log \frac{\int_{c,m} \lambda\left(t \middle| a, m, c, y_p = 1\right) f\left(y_p = 1 \middle| a, c\right) f\left(M = m \middle| a^*, c, y_p = 1\right) f(c) dmdc}{\int_{c,m} f\left(y_p = 1 \middle| a, c\right) f\left(M = m \middle| a^*, c, y_p = 1\right) f(c) dmdc}$$

(by assuming outcome is rare, $e^{-\Lambda\left(t \middle| a, m, c, y_p = 1\right)} \approx 1$)

$= log(\vartheta_T^1(a, a^*)/\vartheta_T^2(a, a^*))$,
where

$\vartheta_T^1(a, a^*) = \int_{c,m} \lambda\left(t \middle| a, m, c, y_p = 1\right) f\left(y_p = 1 \middle| a, c\right) f\left(M = m \middle| a^*, c, y_p = 1\right) f(c) dmdc$, and

$\vartheta_T^2(a, a^*) = \int_{c,m} f\left(y_p = 1 \middle| a, c\right) f\left(M = m \middle| a^*, c, y_p = 1\right) f(c) dmdc$.

Web Appendix C

1. Regression-based estimator for continuous mediator *M*

We assumed that given $A = a^*$, $C = c$, and $Y_p = 1$, M follows a regression model with mean $\beta_0 + \beta_A a^* + \beta_C c$ and variance σ_M^2. Thus, we have

$$Q(a, a^*)$$
$$= \int_{m,c} E\left[Y|A = a, Y_p = 1, M = m, C = c\right] f(Y_p = 1|A = a, C = c) \, f_{M|A,C}\left(M = m|A = a^*, Y_p = 1, C = c\right) f(c) dmdc$$
$$= \int_c expit(\alpha_0 + \alpha_A a + \alpha_C c) \int_m expit(\theta_0 + \theta_A a + \theta_M m + \theta_C c) \, 1/(\sqrt{2\pi}\sigma_M)\exp(-\left(m - (\beta_0 + \beta_A a^* + \beta_C c)\right)^2/2\sigma_M^2) f(c) dmdc$$

Since above formula does not have closed form, we derived that approximate value by using Monte-Carlo method.

2. Proof of *Theorem 4*

$$Q(a, a^*)$$
$$= \iint_{c,m} E\left(Y\middle|A = a, Y_p = 1, M = m, C = c\right) f\left(M = m\middle|A = a^*, Y_p = 1, C = c\right) \times f\left(Y_p = 1\middle|A = a, C = c\right) f(C = c)$$

$$= \iint_{c,m} E\left(Y\middle|A = a, Y_p = 1, M = m, C = c\right) f\left(M = m\middle|A = a^*, Y_p = 1, C = c\right) \frac{f\left(Y_p = 1, A = a, C = c\right)}{f(A = a|C = c)}$$

$$= \iiint_{y,c,m} y f\left(y\middle|A = a, Y_p = 1, M = m, C = c\right) \times \frac{f\left(M = m\middle|A = a^*, Y_p = 1, C = c\right)}{f\left(M = m\middle|A = a, Y_p = 1, C = c\right) f(A = a|C = c)} f\left(M = m, Y_p = 1, A = a, C = c\right)$$

$$= \iiint_{y,c,m} y \frac{f\left(M = m\middle|A = a^*, Y_p = 1, C = c\right)}{f\left(M = m\middle|A = a, Y_p = 1, C = c\right) f(A = 1|C = c)} \times f\left(Y = y, M = m, Y_p = 1, A = a, C = c\right)$$

$$= \iint_{a,y_p} \iiint_{y,c,m} y \frac{I(A=a)}{f(A = a|C = c)} \frac{f\left(M = m\middle|A = a^*, Y_p = 1, C = c\right) I(Y_p = 1)}{f\left(M = m\middle|A = a, Y_p = 1, C = c\right)} \times f\left(Y = y, M = m, Y_p = 1, A = a, C = c\right)$$

$$= E\left[\frac{f\left(M = m\middle|A = a^*, Y_p = 1, C\right) I(A=a) I(Y_p = 1)}{f\left(M = m\middle|A = a, Y_p = 1, C\right) f(A = a|C)} Y\right]$$

3. Proof of *Theorem 5*

$$\mathrm{E}_{Y,M|A,C,Y_p=1}[U_{TE}(\boldsymbol{\lambda})]$$
$$= \mathrm{E}_{Y,M|A,C,Y_p=1}[w(a,c,\widehat{\boldsymbol{\alpha}}) \times \Gamma_{TE}(a,c,\boldsymbol{\lambda})\{Y - g(\mu_{TE}(\boldsymbol{\lambda};a,c))\}]$$
$$= \Gamma_{TE}(a,c,\boldsymbol{\lambda})\mathrm{E}_{Y,M|A,C,Y_p=1}[w(a,c,\widehat{\boldsymbol{\alpha}})(Y - g(\mu_{TE}(\boldsymbol{\lambda};a,c)))]$$
$$= \Gamma_{TE}(a,c,\boldsymbol{\lambda})\{\mathrm{E}_{Y,M|A,C,Y_p=1}[w(a,c,\widehat{\boldsymbol{\alpha}})Y] - \mathrm{E}_{Y,M|A,C,Y_p=1}[w(a,c,\widehat{\boldsymbol{\alpha}})g(\mu_{TE}(\boldsymbol{\lambda};a,c))]\}$$
$$= \Gamma_{TE}(a,c,\boldsymbol{\lambda})\{\int_m E\big(Y|A=a,M=m,C=c,Y_p=1\big)f\big(M=m\big|A=a,C=c,Y_p=1\big)$$
$$f\big(Y_p=1|A=a,C=c\big) - g(\mu_{TE}(\boldsymbol{\lambda};a,c))P_{y_p=1}\big(Y_p=1|A=a,C=c\big)\}$$
$$= \Gamma_{TE}(a,c,\boldsymbol{\lambda})\Big(\varphi(a,a) - g(\mu_{TE}(\boldsymbol{\lambda};a,c))P_{y_p=1}\big(Y_p=1|A=a,C=c\big)\Big) = 0$$

4. Proof of *Theorem 6*

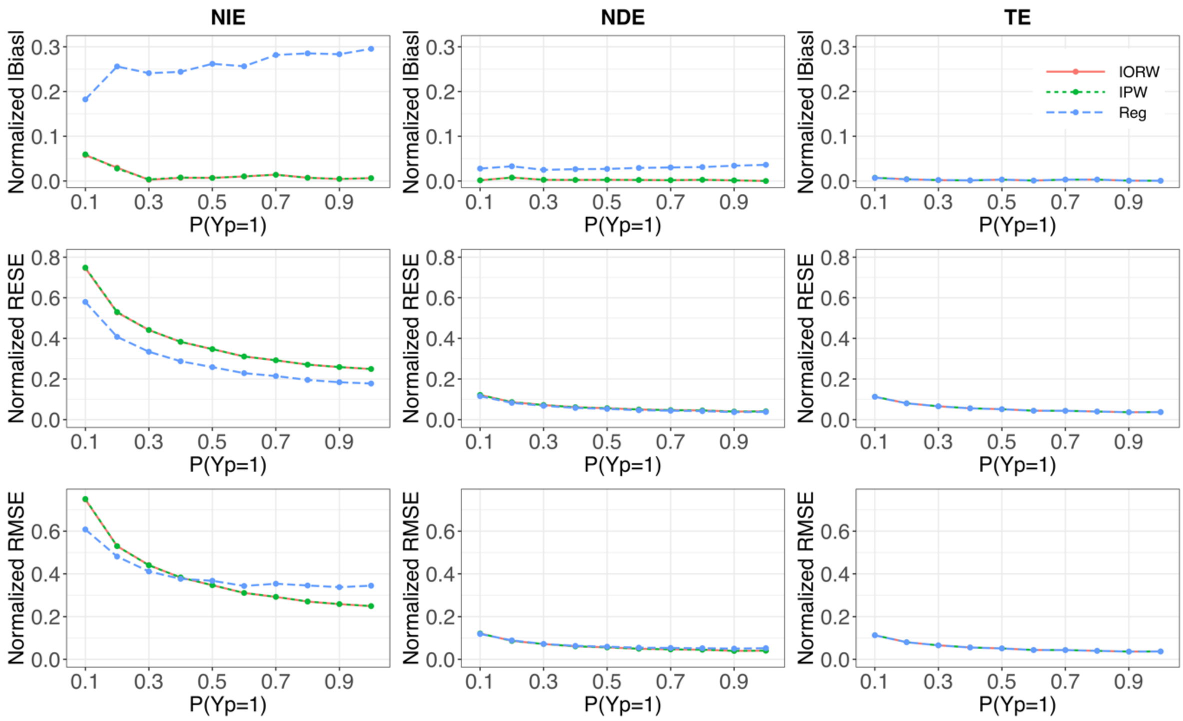
$$\mathrm{E}_{Y,M|A,C,Y_p=1}[U_{DE}(\boldsymbol{\nu})]$$
$$= \mathrm{E}_{Y,M|A,C,Y_p=1}\Big[\mathrm{OR}\big(M=m,A=a\big|C,y_p=1,\widehat{\boldsymbol{\alpha}},\widehat{\boldsymbol{\kappa}}\big)^{-1} \times \Gamma_{DE}(a,c,\boldsymbol{\nu})\{Y - g(\mu_{DE}(\boldsymbol{\nu};a,c))\}\Big]$$
$$= \Gamma_{DE}(a,c,\boldsymbol{\nu})\mathrm{E}_{Y,M|A,C,Y_p=1}\Big[\mathrm{OR}\big(M=m,A=a\big|C,y_p=1,\widehat{\boldsymbol{\alpha}},\widehat{\boldsymbol{\kappa}}\big)^{-1}\{Y - g(\mu_{DE}(\boldsymbol{\nu};a,c))\}\Big]$$
$$= \Gamma_{DE}(a,c,\boldsymbol{\nu})\{\mathrm{E}_{Y,M|A,C,Y_p=1}\Big[\mathrm{OR}\big(M=m,A=a\big|C,y_p=1,\widehat{\boldsymbol{\alpha}},\widehat{\boldsymbol{\kappa}}\big)^{-1}Y\Big]$$
$$-\mathrm{E}_{Y,M|A,C,Y_p=1}\Big[\mathrm{OR}\big(M=m,A=a\big|C,y_p=1,\widehat{\boldsymbol{\alpha}},\widehat{\boldsymbol{\kappa}}\big)^{-1}g(\mu_{DE}(\boldsymbol{\nu};a,c))\Big]\}$$
$$= \Gamma_{DE}(a,c,\boldsymbol{\nu})\frac{f_{M|A,C,Y_p}\big(M=m_0\big|A,Y_p=1,C\big)}{f_{M|A,C,Y_p}\big(M=m_0\big|A=a^*,Y_p=1,C\big)}\Big[\phi(a,0)$$
$$- g(\mu_{DE}(\boldsymbol{\nu};a,c))P_{y_p=1}\big(Y_p=1|A=a,C=c\big)\Big] = 0$$

and

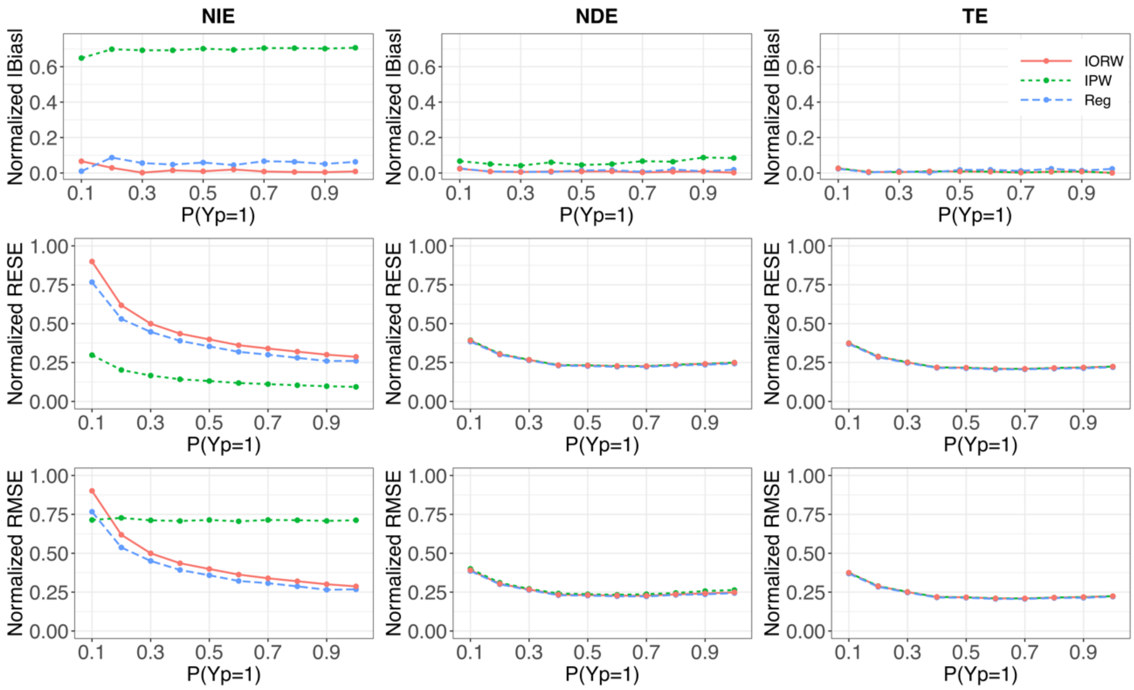
$$\mathrm{E}_{Y,M|A,C,Y_p=1}\Big[\mathrm{OR}\big(M=m,A=a\big|C,y_p=1,\widehat{\boldsymbol{\alpha}},\widehat{\boldsymbol{\kappa}}\big)^{-1}g(\mu_{DE}(\boldsymbol{\nu};a,c))\Big]$$
$$= g(\mu_{DE}(\boldsymbol{\nu};a,c))P_{y_p=1}\big(Y_p=1|A=a,C=c\big) \times$$
$$\int_{y,m}\frac{f(M|A=a^*,C,Y_p=1)f(M=m_0|A,C,Y_p=1)}{f(M|A,C,Y_p=1)f(M=m_0|A=a^*,C,Y_p=1)} \times f\big(y,m|A,C,Y_p=1\big)$$
$$= g(\mu_{DE}(\boldsymbol{\nu};a,c))\frac{f(M=m_0|A,C,Y_p=1)P_{y_p=1}(Y_p=1|A=a,C=c)}{f(M=m_0|A=a^*,C,Y_p=1)} \times$$
$$\int_{y,m}\frac{f(M|A=a^*,C,Y_p=1)}{f(M|A=a,C,Y_p=1)} \times f\big(m|A=a,C=c,Y_p=1\big)f\big(y|M=m,A=a,C=c,Y_p=1\big)$$
$$= g(\mu_{DE}(\boldsymbol{\nu};a,c))\frac{f(M=m_0|A,C,Y_p=1)P_{y_p=1}(Y_p=1|A=a,C=c)}{f(M=m_0|A=a^*,C,Y_p=1)}$$

Web Appendix D

1. Results of simulation study 1



Web Figure 1. Estimated causal effects performance among different methods for scenario 2. The row of the panel represents different measurements, and the column represents different causal effects. The x-axis of each plot is the probability of $Y_p = 1$, and the y-axis represents the quantity of measurements. The complete case approach, IORW, IPW, and Reg methods are colored by red, green, blue, and purple, respectively. Abbreviations: NIE, nature indirect effect; NDE, nature direct effect; TE, total effect; RESE, rooted empirical standard error; RMSE rooted mean square error; IORW, inverse odds ratio weighting; IPW, inverse probability weighting; Reg, regression-based method.



Web Figure 2. Estimated causal effects performance among different methods for scenario 3. The row of the panel represents different measurements, and the column represents different causal effects. The x-axis of each plot is the probability of $Y_p = 1$, and the y-axis represents the quantity of measurements. The complete case approach, IORW, IPW, and Reg methods are colored by red, green, blue, and purple, respectively. Abbreviations: NIE, nature indirect effect; NDE, nature direct effect; TE, total effect; RESE, rooted empirical standard error; RMSE rooted mean square error; IORW, inverse odds ratio weighting; IPW, inverse probability weighting; Reg, regression-based method.

2. Results of simulation study 2

Study 2 also employed a sample of size 10,000 with one binary mediator through Cox's proportional hazards model. The data for each variable were simulated as follows:

$C \sim Bernoulli(p = 0.5)$

$A \sim Bernoulli(p = 0.5)$

$Y_p \sim Bernoulli(p_1),\ p_1 = expit(\alpha_0 + \alpha_A A + \alpha_C C)$

$M = \begin{cases} undefined,\ if\ Y_p = 0 \\ \sim Bernoulli(p_2),\ if\ Y_p = 1 \end{cases},\ p_2 = expit(\beta_0 + \beta_A A + \beta_C C)$

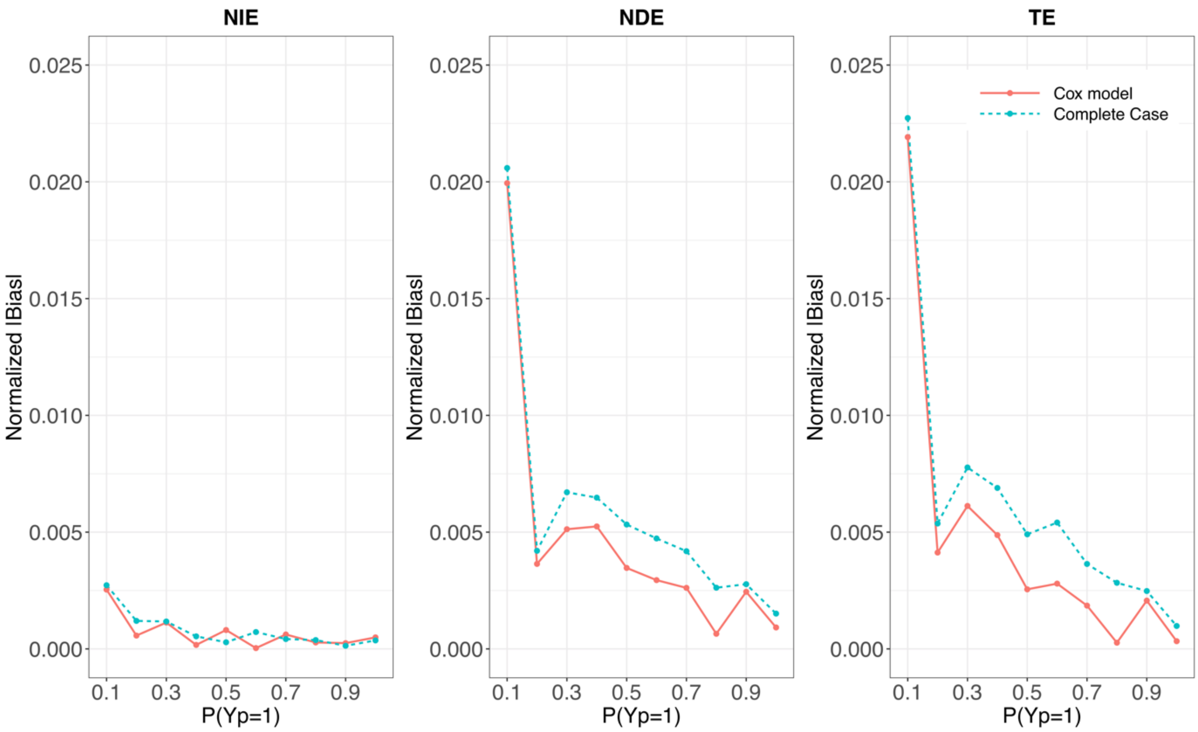
Under $Y_p = 1$, the event times (T) are generated according to a Weibull distribution as

$$T = -log(u) \,/\, (0.01 \times e^{\mu_T}),\ u \sim \text{Uniform}(0,1) \text{ where}$$

$$\mu_T = \gamma_A A + \gamma_M M + \gamma_C C,$$

The parameters were set as $\gamma_A = 0.2$, $\gamma_M = 0.5$, and $\gamma_C = 0.5$. The censoring times (C_T) are randomly drawn from an exponential distribution with a rate of 0.1. As a result, the observed survival times is defined as the minimum of T and C_T. If $Y_p = 0$, then $T = 0$.

The result illustrates in Web Figure 3 by comparing the complete case approach. The figure shows the normalized biases of the estimators for the survival mediation formulas for various probabilities of $Y_p = 1$. Consequently, the result reveals that the proposed Cox model performs better that the complete case approach. Notably, as mentioned in Section 5.1, we adopt the normalized absolute bias, which were divided by $P(Y_p = 1)$, to enable a fair comparison.



Web Figure 3. Absolute values of the normalized biases for the two methods for survival time. The x axis of each plot represents the probability of $Y_p = 1$, and the y axis represents the absolute value of the normalized biases. The proposed Cox model and the complete case approach are indicated by red and green lines, respectively.

Web Appendix E

Data description and data preprocessing

In REVEAL-HBV, a total of 23,820 residents aged 30–65 years from seven townships of Taiwan were recruited from 1991 to 1992 and followed up to 2008, with 477 incident cases of HCC developing subsequently. They provided written informed consent for the questionnaire interview, health examinations, biospecimen collection, and data linkage of health status with death certification profiles and national cancer registry. Blood samples collected at enrollment were tested for seromarkers, viral load of HBV, and HCV antibody. We regard survival status as the outcome of interest (Y), the status of HCV infection as the exposure (A), and age and gender as confounders (C). In this study, the population is restricted to HBV-positive patients and elevated viral load of HBV defined as viral load > 10,000 copies/ml is regarded as M. Yp is used to record the survival status before measuring the value of the mediator. The total number of HBV-positive patients is 4,155. After removing the NA values, there are totally 3,894 samples in the analysis. The frequencies of variables are shown in Web Table 1.

Web Table 1. Frequency table

Variable	Frequency
A:	1: 3696(94.92%); 0: 198(5.08%)
M:	1: 1434(36.83%); 0: 2021(51.9%); Truncation (i.e. Yp =0): 439(11.27%)
Yp:	1: 3455(88.73%); 0: 439(11.27%)
Y:	1: 3254(83.56%); 0: 640(16.44%)