

Mediation analysis for a survival outcome with
time-varying exposures, mediators, and
confounders

Sheng-Hsuan Lin*

Jessica G. Young†

Roger Logan‡

Tyler J. VanderWeele**

*Department of Biostatistics, Columbia Mailman School of Public Health, sh1517@mail.harvard.edu

†Department of Population Medicine, Harvard Medical School, Boston, MA, jessica.gerald.young@gmail.com

‡Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, rwlogan@hsph.harvard.edu

**Department of Epidemiology and Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, tvanderw@hsph.harvard.edu

This working paper is hosted by The Berkeley Electronic Press (bepress) and may not be commercially reproduced without the permission of the copyright holder.

<http://biostats.bepress.com/harvardbiostat/paper203>

Copyright ©2016 by the authors.

Mediation analysis for a survival outcome with time-varying exposures, mediators, and confounders

Sheng-Hsuan Lin, Jessica G. Young, Roger Logan, and Tyler J. VanderWeele

Abstract

We propose an approach to conduct mediation analysis for survival data with time-varying exposures, mediators, and confounders. We identify certain interventional direct and indirect effects through a survival mediational g-formula and describe the required assumptions. We also provide a feasible parametric approach along with an algorithm and software to estimate these effects. We apply this method to analyze the Framingham Heart Study data to investigate the causal mechanism of smoking on mortality through coronary artery disease. The risk ratio of smoking 30 cigarettes per day for ten years compared with no smoking on mortality is 2.34 (95 % CI = (1.44, 3.70)). Of the overall effect, 7.91% (95% CI: = 1.36%, 19.32%) is mediated by coronary artery disease. The survival mediational g-formula constitutes a powerful tool for conducting mediation analysis with longitudinal data.

1 **Mediation analysis for a survival outcome with time-varying exposures,**
2 **mediators, and confounders**

3 **Authors:**

4 **Sheng-Hsuan Lin**

5 Email: shl517@mail.harvard.edu

6 Department of Biostatistics, Columbia Mailman School of Public Health, New York, NY

7 **Jessica G. Young**

8 Email: jessica.gerald.young@gmail.com

9 Department of Population Medicine, Harvard Medical School, Boston, MA

10 **Roger Logan**

11 Email: rwlogan@hsph.harvard.edu

12 Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA

13 **Tyler J. VanderWeele**

14 Email: tvanderw@hsph.harvard.edu

15 Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA

16 Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA

17

18



1 **Summary**

2 We propose an approach to conduct mediation analysis for survival data with time-varying
3 exposures, mediators, and confounders. We identify certain interventional direct and indirect
4 effects through a survival mediational g-formula and describe the required assumptions. We also
5 provide a feasible parametric approach along with an algorithm and software to estimate these
6 effects. We apply this method to analyze the Framingham Heart Study data to investigate the
7 causal mechanism of smoking on mortality through coronary artery disease. The risk ratio of
8 smoking 30 cigarettes per day for ten years compared with no smoking on mortality is 2.34 (95 %
9 CI = (1.44, 3.70)). Of the overall effect, 7.91% (95% CI: = 1.36%, 19.32%) is mediated by
10 coronary artery disease. The survival mediational g-formula constitutes a powerful tool for
11 conducting mediation analysis with longitudinal data.

12

13 **Key words:** longitudinal studies, mediation analysis, mechanism investigation, path analysis,
14 time-varying, survival

15

16 **1. Introduction**

17 In decomposing the total effect into direct and indirect effects, mediation analysis is
18 essential for investigating pathways or mechanisms in epidemiology and the social sciences.
19 Causal mediation analysis defines both direct and indirect effects based on counterfactual models,
20 extending traditional mediation analysis to settings involving nonlinearities and interaction
21 (Pearl, 2001; Robins and Greenland, 1992). Numerous methodological approaches based on
22 causal mediation analysis have been developed in recent years for the estimation of natural direct
23 effect (NDE) and natural indirect effect (NIE) which sum to a total exposure effect. These
24 approaches allow different outcome scales, including the risk difference, the odds ratio, and
25 scales for time-to-event data (Tchetgen and Shpitser, 2012; Valeri and VanderWeele, 2013; van

1 der Laan and Petersen, 2008; VanderWeele and Vansteelandt, 2009; VanderWeele, 2010;
2 VanderWeele and Vansteelandt, 2010).

3 Most of the approaches mentioned above only consider time-fixed exposures and mediators.
4 When conducting causal mediation analysis with longitudinal data, using only one single
5 exposure and mediator ignores the exposures or mediators at other time-points, thus neglecting
6 valuable information at other times. Further, in many settings the question of scientific interest
7 may correspond to the effect of a time-varying exposure and how it is explained by time-varying
8 patterns of a mediating variable.

9 A key challenge in the estimation of NDE and NIE in the context of time-varying exposures
10 and mediators is that, in this setting, time-varying confounders for the effect of the mediator on
11 the outcome at a given time t will be affected by past exposure. In this case, Avin et al. (2005)
12 showed that natural effects cannot generally be identified. VanderWeele and Tchetgen Tchetgen
13 proposed using so called “interventional effects” and derived a mediational g -formula to partially
14 address this problem (VanderWeele and Tchetgen Tchetgen, 2016 (accepted)). This method
15 decomposes the total effect into newly defined interventional direct effect (IDE) and
16 interventional indirect effect (IIE) (which are also called “randomly interventional analogues of
17 natural direct and indirect effects” in some literature (Lin et al., 2016; VanderWeele, Vansteelandt
18 and Robins, 2014)). These effects may generally be identified in the presence of time-varying
19 confounders affected by prior exposure. Lin et al. proposed a parametric mediational g -formula
20 estimator along with a corresponding Statistical Analysis System (SAS) macro for practical
21 implementation (Lin et al., 2016).

22 A limitation of the mediational g -formula as proposed by VanderWeele and Tchetgen
23 Tchetgen is that it requires that the outcome of interest is a variable measured only at the end of
24 follow up and does not immediately generalize to a survival outcome. By contrast, Zheng and

1 van der Laan gave identification results and various estimators for IDE and IIE that are well-
2 defined when the outcome of interest is a survival outcome and mediators vary over time.
3 However, these authors only considered the case where the exposure and confounders are
4 restricted to be time-fixed (Zheng and van der Laan, 2012).

5 In the current study, we adapt ideas from both the mediational g-formula approach and
6 that of Zheng and van der Laan to offer an approach that allows estimation of IDE and IIE of
7 time-varying exposures through time-varying mediators that are well-defined for a survival
8 outcome and, further, that may be identified when time-varying confounders are affected by past
9 exposure. We define IDE and IIE, provide conditions under which they may be identified in an
10 observational study and propose a parametric approach to estimate these effects, along with a
11 corresponding SAS macro for practical implementation. We then apply this method to the case
12 study of the Framingham Heart Study (FHS) data to investigate the effect of time-varying
13 smoking patterns on all-cause survival mediated by time-varying coronary artery disease (CAD)
14 status.

15 The paper is organized as follows: Section 2 introduces the notation and definitions of the
16 direct and indirect effects of interest. Section 3 presents the non-parametric identification of the
17 direct and indirect effects, and the required assumptions. Section 3 also shows that our
18 identifying formula reduces to that provided by Zheng and van der Laan (Zheng and van der
19 Laan, 2012) when both the exposure and confounders are restricted to be time-fixed. A feasible
20 parametric approach along with an algorithm and software are also provided in Section 4. Section
21 5 describes the analytical procedure and provides the estimation results for the FHS data. Section
22 6 concludes by discussing the strengths and limitations of the study.

23

24 **2. Notation and definitions**

1 **2.1 Observed data**

2 Consider an observational study where the exposures, mediators, confounders, and
3 survival outcomes that vary over time in longitudinal data with measurements at times 1 to T. Let
4 $(A(1), \dots, A(T))$, $(M(1), \dots, M(T))$, and $(L(1), \dots, L(T))$, denote random variables corresponding to
5 the measured time-varying exposures, mediators, and confounders, in measurement periods 1, ...,
6 T, respectively, with initial baseline confounders, V . Let $S(t) = 1$ indicate survival by period and
7 $S(t) = 0$ death by period $t = 1, \dots, T$. When $S(t) = 0$, the following survival variables (i.e. $S(k)$, k
8 $\in \{t+1, t+2, \dots, T\}$) are all equal to zero, and the following variables other than $S(k)$ (i.e. $A(k)$,
9 $M(k)$, $L(k)$, $k \in \{t+1, t+2, \dots, T\}$) are all undefined. The assumed temporal order of variables in
10 each measurement period is depicted in the causal diagram in Figure 1.

11 Our direct and indirect effects will be defined as contrasts in the mean of the survival
12 indicator by time T had all subjects followed different hypothetical interventions on the time-
13 varying exposure, mediator, as well as past survival. In the next section we will develop
14 counterfactual notation to formally define these effects.

15

16 **2.2 Counterfactual notation and the consistency assumption**

17 For any random variable W and realized value w , let $W(t_1:t_2) = (W(t_1), W(t_1+1), \dots, W(t_2))$
18 and $w(t_1:t_2) = (w(t_1), w(t_1+1), \dots, w(t_2))$, where t_1 and t_2 are all positive integers and $t_1 < t_2$.

19 Let $a(1:T)$ and $a(1:T)^*$ denote two different hypothetical time-varying exposure
20 interventions such that we set $A(1:T)$ to $a(1:T)$ and $A(1:T)$ to $a(1:T)^*$, respectively; for example,
21 the hypothetical interventions “always expose” would be $a(1:T) = (1, \dots, 1)$ and “never expose”
22 would be $a(1:T)^* = (0, \dots, 0)$. Let $S(t)_{a(1:t)m(1:t)}$ be the counterfactual value of $S(t)$, given the
23 previous exposures $A(1:t)$ are set to $a(1:t)$ and the previous mediators $M(1:t)$ are set to $m(1:t)$. Let
24 $S(t)_{a(1:t)m(1:t)s(1:t-1)}$ be the counterfactual value of $S(t)$, given the previous exposures $A(1:t)$ are set to

1 a(1:t), the previous mediators M(1:t) are set to m(1:t), and S(1:t-1) are set to s(1:t-1); for this
 2 latter counterfactual, we are conceiving an intervention on past survival. Another way forward in
 3 this context might be to address this setting using principal stratification (Frangakis and Rubin,
 4 2002; Tchetgen Tchetgen, 2014), but this is difficult in the context of mediation. So we assume
 5 that the intervention on survival is possible as does prior literature on this topic (Zheng and van
 6 der Laan, 2012).

7 Let $M(t)_{a(1:t)}$ and $S(t)_{a(1:t)}$ be the counterfactual value of M(t) and S(t), respectively, given
 8 $A(1:t)$ are set to a(1:t). Let $M(t)_{a(1:t)m(1:t)s(1:t-1)}$ be the counterfactual value of M(t), given the
 9 previous exposures, A(1:t), are set to a(1:t), the previous mediators, M(1:t-1), are set to m(1:t-1),
 10 and S(1:t-1) is set to s(1:t-1).

11 We will make consistency assumption (Pearl, 2009; VanderWeele and Vansteelandt, 2009;
 12 VanderWeele, 2009a). Under this assumption,

13

14 $S(t)_{a(1:t)} = S(t)$ and $M(t)_{a(1:t)} = M(t)$ given $A(1:t) = a(1:t)$ (Consistency assumption 1);

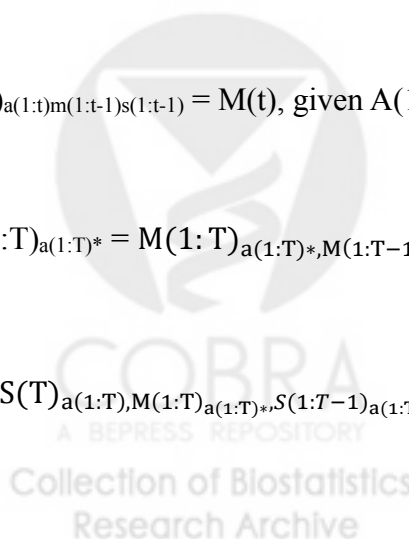
15 $S(t)_{a(1:t)m(1:t)} = S(t)$ given $A(1:t) = a(1:t)$ and $M(1:t) = m(1:t)$ (Consistency assumption 2);

16 $S(t)_{a(1:t)m(1:t)s(1:t-1)} = S(t)$ given $A(1:t) = a(1:t)$, $M(1:t) = m(1:t)$, and $S(t-1) = s(t-1)$
 17 (Consistency assumption 3);

18 $M(t)_{a(1:t)m(1:t-1)s(1:t-1)} = M(t)$, given $A(1:t) = a(1:t)$, $M(1:t-1) = m(1:t-1)$, and $S(t-1) = s(t-1)$
 19 (Consistency assumption 4);

20 $M(1:T)_{a(1:T)*} = M(1:T)_{a(1:T)*, M(1:T-1)_{a(1:T-1)*}, S(T-1)_{a(1:T-1)*, M(1:T-1)_{a(1:T-1)*}}$
 21 (Consistency assumption 5);

22 and $S(T)_{a(1:T), M(1:T)_{a(1:T)*}, S(1:T-1)_{a(1:T-1), M(1:T-1)_{a(1:T-1)*}}} = S(T)_{a(1:T), M(1:T)_{a(1:T)*}}$
 23 (Consistency assumption 6).



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24

2.3 Why the traditional mediation parameter is undefined for survival outcomes and time-varying mediators

For questions about the investigation of mechanisms, the direct and indirect effects are traditionally represented by NDE and NIE, respectively, which are mathematically defined below:

$$\text{NDE} \equiv E[Y_{a(1:T),M(1:T)_{a^*}} - Y_{a(1:T)^*,M(1:T)_{a(1:T)^*}} | V] \quad (\text{Definition 1-1})$$

$$\text{NIE} \equiv E[Y_{a(1:T),M(1:T)_{a(1:T)}} - Y_{a(1:T),M(1:T)_{a(1:T)^*}}]. \quad (\text{Definition 1-2})$$

By replacing the Y with the survival outcome at the end of follow up, the NDE and NIE would be defined as (Pearl, 2001; Robins and Greenland, 1992; VanderWeele and Vansteelandt, 2009):

$$\text{NDE} \equiv E[S(T)_{a(1:T),M(1:T)_{a^*}} - S(T)_{a(1:T)^*,M(1:T)_{a(1:T)^*}} | V] \quad (\text{Definition 1-3})$$

$$\text{NIE} \equiv E[S(T)_{a(1:T),M(1:T)_{a(1:T)}} - S(T)_{a(1:T),M(1:T)_{a(1:T)^*}}] \quad (\text{Definition 1-4}).$$

For simplicity of expression, we define the mediation parameter $\Phi(a(1:T), a(1:T)^*)$ below:

$$\Phi(a(1:T), a(1:T)^*) \equiv E[S(T)_{a(1:T),M(1:T)_{a(1:T)^*}}]. \quad (\text{Definition 1-5}).$$

Thus, we can rewrite the definitions 3 and 4 as follows:

$$\text{NDE} \equiv \Phi(a(1:T), a(1:T)^*) - \Phi(a(1:T)^*, a(1:T)^*) \quad (\text{Definition 1-6})$$

$$\text{NIE} \equiv \Phi(a(1:T), a(1:T)) - \Phi(a(1:T), a(1:T)^*). \quad (\text{Definition 1-7})$$

This definition of NIE captures all effects of A(1:T) on S(T) through M(1:T).

1 These traditional definitions present two complications in the setting of meditational
2 analysis with a survival outcome with time-varying exposures, mediators, and confounders. First,
3 when the confounders vary over time and are both affected by prior exposure and also affect
4 subsequent values of both the mediator and outcome, natural direct and indirect effects are not
5 identified from data regardless of whether data is available on the time-varying confounders or
6 not (Avin, Shpitser and Pearl, 2005; VanderWeele and Tchetgen Tchetgen, 2016 (accepted)). Here
7 we will deal with the first problem by changing the definition of the direct and indirect effects to
8 their interventional analogues (i.e., IDE and IIE) (Lin et al., 2016; VanderWeele and Tchetgen
9 Tchetgen, 2016 (accepted)) wherein the mediator for each exposed individual is not set to its
10 level in the absence of exposure but rather to a value randomly chosen from the distribution of
11 what would be observed if everyone had been unexposed. This slightly altered definition allows
12 for identification of the effects of interest, as has been done in prior literature for binary and
13 continuous outcomes at the end of follow-up (Lin et al., 2016; VanderWeele and Tchetgen
14 Tchetgen, 2016 (accepted)). We will describe the definition of IDE and IIE in section 2.6.

15 A second complication that arises in the survival setting is that due to death before end of
16 follow-up the mediator trajectory under alternative settings of the exposure may be undefined.
17 More specifically, the definition of NDE and NIE can be undefined for individuals when the
18 exposure benefits the survival status at time T-1, not through the mediators, which can be
19 expressed by following two equations jointly:

20 $S(T - 1)_{a(1:T-1), M(1:T-1)_{a(1:T-1)*}} = 1$ (Equation 1)

21 $S(T - 1)_{a(1:T-1)*, M(1:T-1)_{a(1:T-1)*}} = 0$ (Equation 2).

22
23 Based on Equation 2, $M(1:T)_{a(1:T)*}$ is undefined. Consequently, the mediation parameter $\Phi(a(1:T),$
24 $a(1:T)*)$ is also undefined. In addition, NDE and NIE, which are both defined by $\Phi(a(1:T),$

1 $a(1:T)^*$, are also undefined. Further discussion is provided in Appendix 5 in the Online
 2 Supplement.

3
 4 **2.4 An alternative definition of the mediation parameter that is well-defined**
 5 **for a survival outcome with time-varying mediators**

6 One way to address the undefined NDE and NIE is further altering the definition of the
 7 mediation parameter (Definition 5) (Zheng and van der Laan, 2012) so that the effects are always
 8 defined. We do so using the following notation. We define $M(t)^{**}$ and $S(t)^{**}$ sequentially. When
 9 time = 1, $M(1)^{**}$ is defined as $M(1)_{a(1)^*}$ and $S(1)^{**}$ as $S(1)_{a(1), M(1)^*}$. When time = 2, $M(2)^{**}$ is
 10 defined as $M(2)_{a(1:2)^*, M(1)^{**}, S(1)^{**}}$, and $S(2)^{**}$ as $S(2)_{a(1:2), M(1:2)^{**}, S(1)^{**}}$. We continue this definition
 11 process iteratively. For time = t , $M(t)^{**}$ and $S(t)^{**}$ are defined below.

12
$$M(t)^{**} \equiv M(t)_{a(1:t)^*, M(1:t-1)^{**}, S(1:t-1)^{**}} \quad (\text{Definition 2-1})$$

13
$$S(t)^{**} \equiv S(t)_{a(1:t), M(1:t)^{**}, S(1:t-1)^{**}} \quad (\text{Definition 2-2}).$$

14
 15 Since $M(t)$ is undefined when $S(t-1) = 0$, under the counterfactual model, $M(t)_{s(t-1)=0}$ is also
 16 undefined. Consequently, $S(t)_{m(t), s(t-1)=1}$ is undefined when $m(t)$ is undefined. Since $S(t)$ is always
 17 equal to zero when $S(t-1) = 0$, $S(t)_{m(t), s(t-1)=0}$ is always equal to zero even when $m(t)$ is undefined.

18 We saw in the previous section that, by Equation 2 and Consistency Assumptions 5 and 6,
 19 $M(t)_{a(1:t)^*}$ is undefined; in words, under the hypothetical intervention defining the counterfactual
 20 corresponding to the traditional mediation parameter (Definition 1-5), the mediator under that
 21 intervention can be undefined at time t for a subject still surviving at t under that intervention. In
 22 contrast to $M(t)_{a(1:t)^*}$, the alternative $M(t)^{**}$ as we have just defined above is undefined if and
 23 only if $S(t-1)^{**} = 0$; in words, under the hypothetical intervention defining our alternative
 24 counterfactual, the mediator under that intervention can only be undefined at time t if a subject

1 has failed by t. Therefore, $S(t)^{**}$ is still defined even when $M(t)^{**}$ is undefined since in that
 2 condition, $S(t-1)^{**} = 0$, leading to $S(t)^{**} = 0$; in words, regardless of the intervention under
 3 consideration, if a subject has failed by time t-1, she/he has failed by time t.

4 By the above, we define the alternative mediation parameter ($\Phi(a(1:T), a(1:T)^*)$), NDE, and
 5 NIE as follows:

6 $\Phi(a(1:T), a(1:T)^*) \equiv E[S(T)^{**}]$ (Definition 3-1)

7 $NDE \equiv \Phi(a(1:T), a(1:T)^*) - \Phi(a(1:T)^*, a(1:T)^*)$ (Definition 3-2)

8 $NIE \equiv \Phi(a(1:T), a(1:T)) - \Phi(a(1:T), a(1:T)^*)$. (Definition 3-3)

9

10 Thus at time 1, the natural direct and indirect effects are given explicitly as

11

12 $NDE(1) \equiv E[S(1)_{a(1), M(1)_{a(1)^*}}] - E[S(1)_{a(1)^*, M(1)_{a(1)^*}}]$ (Definition 3-4)

13 $NIE(1) \equiv E[S(1)_{a(1), M(1)_{a(1)}}] - E[S(1)_{a(1), M(1)_{a(1)^*}}]$ (Definition 3-5)

14

15 and at time 2, the natural direct and indirect effects are given explicitly by

16 $NDE(2) \equiv E[S(2)_{a(1), M(1)_{a(1)^*}, S(1)_{a(1), M(1)_{a(1)^*}, a(2)}, M(2)_{a(1)^*, M(1)_{a(1)^*}, S(1)_{a(1), M(1)_{a(1)^*}, a(2)^*}] -$

17 $E[S(2)_{a(1)^*, M(1)_{a(1)^*}, S(1)_{a(1)^*, M(1)_{a(1)^*}, a(2)^*}, M(2)_{a(1)^*, M(1)_{a(1)^*}, S(1)_{a(1)^*, M(1)_{a(1)^*}, a(2)^*}]$

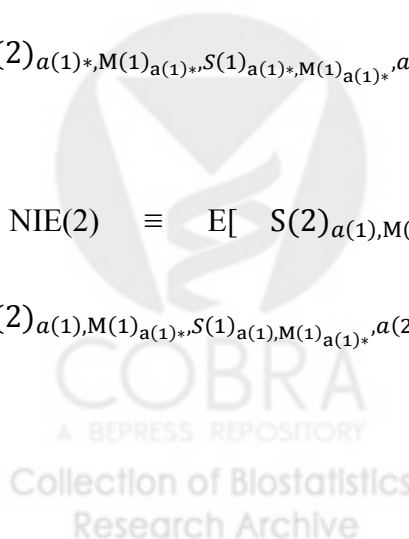
18 (Definition 3-6)

19 $NIE(2) \equiv E[S(2)_{a(1), M(1)_{a(1)}, S(1)_{a(1), M(1)_{a(1)}, a(2)}, M(2)_{a(1), M(1)_{a(1)}, S(1)_{a(1), M(1)_{a(1)}, a(2)}] -$

20 $E[S(2)_{a(1), M(1)_{a(1)^*}, S(1)_{a(1), M(1)_{a(1)^*}, a(2)}, M(2)_{a(1)^*, M(1)_{a(1)^*}, S(1)_{a(1), M(1)_{a(1)^*}, a(2)^*}]$

21 (Definition 3-7).

22



1 2.5 Relation to path-specific effects

2 The distinctions between these definitions and the traditional definitions can be illustrated
 3 by considering path-specific effects (Avin et al., 2005; VanderWeele and Tchetgen Tchetgen,
 4 2016 (accepted); Zheng and van der Laan, 2012). Under the traditional definition, an indirect
 5 effect consists of two groups of path-specific effects: first, the path of exposures affecting the
 6 mediators through earlier survival history, and second, the path of exposures affecting the
 7 mediators not through earlier survival history. Under our definition for the alternative mediation
 8 parameter (Definition 3-1), the indirect effect includes the second path-specific effect, and the
 9 first path-specific effect is included in the direct effect. Consider the simplest setting as an
 10 example: given $T = 2$, and V , $A(2)$, $L(1)$, and $L(2)$ are all empty (Figure 2). According to
 11 traditional definition, NIE includes (a) $A \rightarrow M(2) \rightarrow S(2)$, (b) $A \rightarrow M(1) \rightarrow M(2) \rightarrow S(2)$, (c)
 12 $A \rightarrow M(1) \rightarrow S(1) \rightarrow M(2) \rightarrow S(2)$, (d) $A \rightarrow M(1) \rightarrow S(1) \rightarrow S(2)$, (e) $A \rightarrow M(1) \rightarrow S(2)$, and (f)
 13 $A \rightarrow S(1) \rightarrow M(2) \rightarrow S(2)$; NDE includes (a) $A \rightarrow S(1) \rightarrow S(2)$ and (b) $A \rightarrow S(2)$. From the eight possible
 14 path-specific effects, $A \rightarrow S(1) \rightarrow M(2) \rightarrow S(2)$ is traditionally included as part of NIE since $M(2)$ is
 15 involved. Under our definition (Definition 3-1 to 3-3), $A \rightarrow S(1) \rightarrow M(2) \rightarrow S(2)$ is included in the
 16 direct effect since the exposure affects survival history ($S(1)$) first, not through the earlier
 17 mediator ($M(1)$). For the natural effects, the path ($A \rightarrow S(1) \rightarrow M(2) \rightarrow S(2)$) can be measured by
 18 $E[S(2)_{a,M(1)_{a^*},S(1)_{a,M(1)_{a^*}},M(2)_{a^*,M(1)_{a^*},S(1)_{a,M(1)_{a^*}}}] - E[S(2)_{a,M(1)_{a^*},S(1)_{a,M(1)_{a^*}},M(2)_{a^*,M(1)_{a^*},S(1)_{a^*,M(1)_{a^*}}}]$.

19 Now, when exposure affects survival status at time 1 beneficially (mathematically expressed as
 20 $(S(1)_{a,M(1)_{a^*}}, S(1)_{a^*}) = (1,0)$), the quantity $\Phi(a,a^*) (= E[S(2)_{a,M(1)_{a^*},S(1)=1,M(2)_{a^*,M(1)_{a^*},S(1)=0}])$ is
 21 undefined, and consequently, the path, $A \rightarrow S(1) \rightarrow M(2) \rightarrow S(2)$, is undefined. On the other hand,
 22 when exposure has monotonically harmful effect on survival status at time 1 (mathematically
 23 expressed as $(S(1)_{a,M(1)_{a^*}}, S(1)_{a^*}) \in \{(1,1), (0,0), (0,1)\}$), the path is always equal to zero (when

1 $S(1)_{a,M(1)_{a^*}} = 0$, $E[S(2)_{a,M(1)_{a^*},S(1)=0,M(2)_{a^*,M(1)_{a^*},S(1)=0}}] = 0$ and $\Phi(a,a^*) =$
 2 $E[S(2)_{a,M(1)_{a^*},S(1)=0,M(2)_{a^*,M(1)_{a^*},S(1)_{a^*,M(1)_{a^*}}}] = 0$; when $S(1)_{a,M(1)_{a^*}} = 1$ and $S(1)_{a^*} = 1$, $\Phi(a,a^*)$ are
 3 all equal to $E[S(2)_{a,M(1)_{a^*},S(1)=1,M(2)_{a^*,M(1)_{a^*},S(1)=1}}]$. In conclusion, the the traditional and
 4 alternative mediation parameters will coincide unless the tradition parameter is undefined in
 5 which case their difference is also of course undefined. Thus, the path-specific effect might be
 6 thought to be a reasonable alternative definition for the direct and indirect effects.

7

8 **2.6 Randomized interventional analogues for identification**

9 Following arguments of Avin et al., even the effects above cannot be identified by empirical
 10 data. We can, as prior literature, proceed by defining effects that correspond reasonably closely to
 11 natural direct and indirect effects, but that are still defined in the context of survival outcomes,
 12 and that also, as will be seen in the next section can be identified with data. We do so using the
 13 following notation, which is similar to Definitions 2-1 and 2-2. We define $M(t)^*$, $G(t)^*$, and $S(t)^*$
 14 sequentially. When time = 1, $M(1)^*$ is defined as $M(1)_{a(1)^*}$, $G(1)^*$ as a random draw of $M(1)^*$,
 15 and $S(1)^*$ as $S(1)_{a(1), G(1)^*}$. When time = 2, $M(2)^*$ is defined as $M(2)_{a(1:2)^*,G(1)^*,S(1)^*}$, $G(2)^*$ as a
 16 random draw of $M(2)^*$, and $S(2)^*$ as $S(2)_{a(1:2),G(1:2)^*,S(1)^*}$. We continue this definition process
 17 iteratively. For time = t, $M(t)^*$, $G(t)^*$, and $S(t)^*$ are defined below:

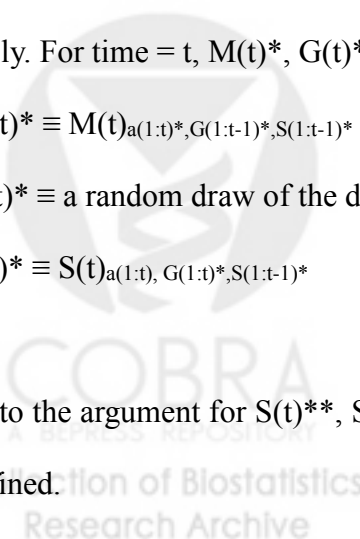
18 $M(t)^* \equiv M(t)_{a(1:t)^*,G(1:t-1)^*,S(1:t-1)^*}$ (Definition 4-1)

19 $G(t)^* \equiv$ a random draw of the distribution of $M(t)^*$ (Definition 4-2)

20 $S(t)^* \equiv S(t)_{a(1:t), G(1:t)^*,S(1:t-1)^*}$ (Definition 4-3).

21

22 Similar to the argument for $S(t)^{**}$, $S(t)$ is also always well-defined and equal to zero when $M(t)^*$
 23 is undefined.



1 By the above, we define the alternative mediation parameter, $\Psi(a(1:T), a(1:T)^*)$, as well as
 2 IDE and IIE as following:

3 $\Psi(a(1:T) = E[S(T)^*]$ (Definition 5-1)

4 $IDE = \Psi(a(1:T), a(1:T)^*) - \Psi(a(1:T)^*, a(1:T)^*)$ (Definition 5-2)

5 $IIE = \Psi(a(1:T), a(1:T)) - \Psi(a(1:T), a(1:T)^*)$. (Definition 5-3)

6

7 Thus at time 1, the interventional direct and indirect effects are given explicitly as

8 $IDE(1) = E[S(1)_{a(1),G(1)_{a(1)^*}}] - E[S(1)_{a(1)^*,G(1)_{a(1)^*}}]$ (Definition 5-4)

9 $IIE(1) = E[S(1)_{a(1),G(1)_{a(1)}}] - E[S(1)_{a(1),G(1)_{a(1)^*}}]$ (Definition 5-5)

10

11 and at time 2, the interventional direct and indirect effects are given explicitly by

12

13 $IDE(2) = E[S(2)_{a(1),G(1)_{a(1)^*},S(1)_{a(1),G(1)_{a(1)^*},a(2), G(2)_{a(1)^*,G(1)_{a(1)^*},S(1)_{a(1),G(1)_{a(1)^*},a(2)^*} }] -$

14 $E[S(2)_{a(1)^*,G(1)_{a(1)^*},S(1)_{a(1)^*,G(1)_{a(1)^*},a(2)^*, G(2)_{a(1)^*,G(1)_{a(1)^*},S(1)_{a(1)^*,G(1)_{a(1)^*},a(2)^*}]$

15 (Definition 5-6)

16 $IIE(2) = E[S(2)_{a(1),G(1)_{a(1)},S(1)_{a(1),G(1)_{a(1)},a(2), G(2)_{a(1),G(1)_{a(1)},S(1)_{a(1),G(1)_{a(1)},a(2)}] -$

17 $E[S(2)_{a(1),G(1)_{a(1)^*},S(1)_{a(1),G(1)_{a(1)^*},a(2), G(2)_{a(1)^*,G(1)_{a(1)^*},S(1)_{a(1),G(1)_{a(1)^*},a(2)^*}]$.

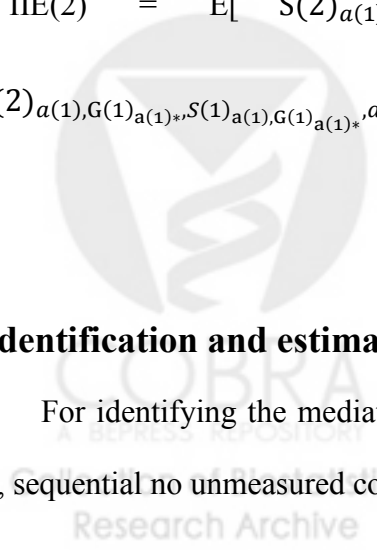
18 (Definition 5-7)

19

20 **3. Identification and estimation of survival mediational g-formula**

21 For identifying the mediation parameter, $\Psi(a(1:T), a(1:T)^*)$, we can make the following

22 four, sequential no unmeasured confounding assumptions for $t = 1, 2, \dots, T$:



- 1 1. $S(T)_{a(1:t), m(1:t), s(1:t-1)=1, G(t+1:T)^*, S(t:T)^*} \perp A(t) | v, A(1:t-1)=a(1:t-1), M(1:t-1)=m(1:t-1), L(1:t-1), S(1:t-$
2 $1)^*=1, G(1:t)^*=m(1:t)$ (no unmeasured exposure-outcome confounding conditional on the past
3 covariates, while \perp indicates independence) (Assumption 1)
- 4 2. $S(T)_{a(1:t), m(1:t), s(1:t-1)=1, G(t+1:T)^*, S(t:T)^*} \perp M(t) | v, A(1:t)=a(1:t), M(1:t-1)=m(1:t-1), L(1:t-1), S(1:t-$
5 $1)^*=1, G(1:t)^*=m(1:t)$ (no unmeasured mediator-outcome confounding conditional on the past
6 covariates) (Assumption 2)
- 7 3. $M(t)_{a(1:t)^*, m(1:t-1), s(1:t-1)=1} \perp S(t-1) | v, A(1:t-1), M(1:t-1), L(1:t-1), S(1:t-2)=1$ (no unmeasured
8 mediator-previous survival confounding conditional on the past covariates) (Assumption 3)
- 9 4. $M(t)_{a(1:t)^*, m(1:t-1), s(1:t-1)=1} \perp A(t) | A(1:t-1), M(1:t-1), L(1:t-1), S(1:t-1)=1, v$ (no unmeasured exposure-
10 mediator confounding conditional on the past covariates) (Assumption 4)

11

12 It is worth noting that a common cause of survival status at different time points is
13 allowed by assumption 3. We illustrate this in Appendix 4 of the Online Supplement. Under the
14 four assumptions, the mediation parameter $\Psi(a(1:T), a(1:T)^*)$ can be identified as

15 $\Psi(a(1:T), a(1:T)^*) = Q(a(1:T), a(1:T)^*)$ (Equation 3),

16 where

17 $Q(a(1:T), a(1:T)^*) = \sum_{v, m(1:T)} \sum_{l(1:T)} \prod_{t=1}^T E[S(t) | a(1:t), m(1:t), l(1:t), S(t-1) = 1, v] \times$
18 $\prod_{t=1}^{T-1} P(l(t) | a(1:t), m(1:t), l(1:t-1), S(t-1) = 1, v) \times$
19 $\sum_{l'(1:T-1)} \prod_{t=1}^T \Pr(m(t) | a^*(1:t), m(1:t-1), l'(1:t-1), S(1:t-1) = 1, v) \times$
20 $\Pr(l'(t-1) | a^*(1:t-1), m(1:t-1), l'(1:t-2), S(1:t-2) = 1, v) \Pr(v)$ (Equation 4)

21

22 The $Q(a(1:T), a(1:T)^*)$ can also be expressed by the counting process notation, $N(t)$, and the
23 continuous time-varying mediators and confounders as follows.

24 $Q(a(1:T), a(1:T)^*)$

$$\begin{aligned}
&= \int_{m(1:T), l(1:T)} \prod_{t=1}^T E[1 - N(t) | a(1:t), m(1:t), l(1:t), N(t-1) = 0, v] \times \\
&\prod_{t=1}^{T-1} f_{L(t)}(l(t) | a(1:t), m(1:t), l(1:t-1), N(t-1) = 0, v) dl(1:T) \times \\
&\int_{l'(1:T-1)} \prod_{t=1}^T f_{M(t)}(m(t) | a^*(1:t), m(1:t-1), l'(1:t-1), N(1:t-1) = 0, v) \times \\
&f_{L'(t-1)}(l'(t-1) | a^*(1:t-1), m(1:t-1), l'(1:t-2), N(1:t-2) = 0, v) dm(1:T) dl'(1:T-1)
\end{aligned}$$

(Equation 5)

The proof is given in Appendices 1 and 3 of the Online Supplement. We refer to this final expression $Q(a(1:T), a(1:T)^*)$ as the survival mediational g-formula (sMGF, Equations 4 and 5). Consequently, the IDE and IIE can be identified non-parametrically by the following equations:

$$IDE = Q(a(1:T), a(1:T)^*) - Q(a(1:T)^*, a(1:T)^*) \quad \text{(Equation 6)}$$

$$IIE = Q(a(1:T), a(1:T)) - Q(a(1:T), a(1:T)^*) \quad \text{(Equation 7)}$$

Intuitively, the sMGF (Equations 1 and 2) can be understood as a weighted average. Each weighed term

$$\begin{aligned}
&(\sum_{m(1:T)} \sum_{l(1:T)} \prod_{t=1}^T E[S(t) | a(1:t), m(1:t), l(1:t), S(t-1) = 1, v] \times \\
&\prod_{t=1}^{T-1} P(l(t) | a(1:t), m(1:t), l(1:t-1), S(t-1) = 1, v)
\end{aligned}$$

(Formula 1))

is the g-formula for the probability of survival by time T had all subjects undergone time-varying intervention set $A(1:t)$ to $a(1:t)$ and set $M(1:t)$ to some $m(1:t)$ through all surviving times $t=1, \dots, T$ without censoring (Hernán, 2004; Robins, 1986; Taubman et al., 2009). The weights for each $m(1:t)$ in the previous sum are

$$\sum_{l'(1:T-1)} \prod_{t=1}^T \Pr(m(t)|a^*(1:t), m(1:t-1), l'(1:t-1), S(1:t-1) = 1, v) \times \Pr(l'(t-1)|a^*(1:t-1), m(1:t-1), l'(1:t-2), S(1:t-2) = 1, v)$$

(Formula 2),

which, given our assumptions, identifies the joint distribution of the counterfactual mediators, $M(1:T)^*_{a(1:T)^*}$, for survivors; i.e. those with $S(1:T-1)^* = 1$.

When the set of mediators is empty, the survival mediational g-formula reduces to the following form:

$$E[S(T)_{a(1:T)}] = \sum_{l(1:T)} \prod_{t=1}^T E[S(t)|a(1:t), l(1:t), S(t-1) = 1, v] \times \prod_{t=1}^{T-1} P(l(t)|a(1:t), l(1:t-1), S(t-1) = 1, v)$$

(Equation 8)

which is the standard g-formula associated with an intervention $a(1:T)$. Thus we can conclude that the survival mediational g-formula, similar to the mediational g-formula (VanderWeele and Tchetgen Tchetgen, 2016 (accepted)), is a generalized form of the g-formula that applies when assessing mediation is of interest.

When $S(T-1)$ is always equal to one, this expression reduces to

$$\sum_{m(1:T)} \sum_{l(1:T)} E[S(T)|a(1:T), m(1:T), l(1:T), v] \times \prod_{t=1}^{T-1} P(l(t)|a(1:t), m(1:t), l(1:t-1), v) \times \sum_{l'(1:T-1)} \prod_{t=1}^T \Pr(m(t)|a^*(1:t), m(1:t-1), l'(1:t-1), v) \times \Pr(l'(t-1)|a^*(1:t-1), m(1:t-1), l'(1:t-2), v)$$

(Formula 3)

which is the mediational g-formula provided by VanderWeele and Tchetgen Tchetgen's work (VanderWeele and Tchetgen Tchetgen, 2016 (accepted)), given that the outcome of interest

1 (denoted as Y) is survival status at the end of follow up ($S(T)$).

2 If neither the confounders nor the exposure is time-varying (i.e. $A(2:T)$ and $L(1:T)$ are
3 empty) (Figure 3), the survival mediational g-formula reduces to

$$4 \sum_{m(1:T)} \prod_{t=1}^T E[S(t)|a, m(1:t), S(t-1) = 1, v] \times \Pr(m(t)|a^*, m(1:t-1), S(t-1) = 1, v)$$

5 (Formula 4)

6
7 which is the identifying formula given by Zheng's and van der Laan (equation 6 and equation 24)
8 (Zheng and van der Laan, 2012). However, our survival mediational g-formula can still be used,
9 unlike the approach of Zheng and van der Laan, even if the exposures and confounders are time-
10 varying.

11 Furthermore, under the monotonicity assumption:

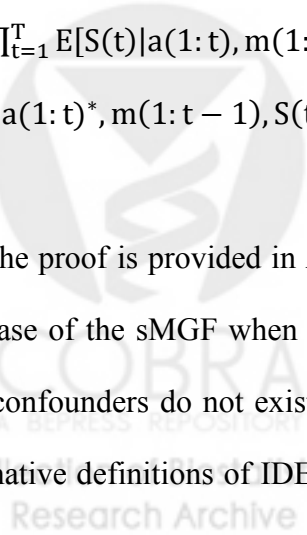
$$12 S(t)_{a,m(1:t-1)} \leq S(t)_{a^*,m(1:t-1)} \text{ for all individuals where } t = 1, 2, \dots, T-1 \quad (\text{monotonicity}$$

13 assumption),

14
15 the traditional definition of mediation parameter $\Phi(a(1:T), a(1:T)^*)$ can be defined and identified
16 by the following expression:

$$17 \sum_{m(1:T)} \prod_{t=1}^T E[S(t)|a(1:t), m(1:t), S(t-1) = 1, v] \times$$
$$18 \Pr(m(t)|a(1:t)^*, m(1:t-1), S(t-1) = 1, v) \quad (\text{Formula 5})$$

19
20 The proof is provided in Appendix 2 of the Online Supplement. This expression is also a
21 special case of the sMGF when both $A(2:T)$ and $L(1:T)$ are empty. In other words, given time-
22 varying confounders do not exist and exposures have monotonically harmful effect on survival,
23 the alternative definitions of IDE and IIE are the same as the traditional definitions of NDE and



1 NIE, respectively, and under these assumptions the traditional definitions will always be defined
2 (i.e., the difference between the traditional and alternative mediation parameters is equal to zero).

3

4 **4. Parametric approach for survival mediational g-formula**

5 We here describe a parametric approach to estimate the survival mediational g-formula
6 functional and develop a corresponding SAS macro. Since the survival mediational g-formula is a
7 generalized form of the g-formula, we create the algorithm and a macro for the survival
8 mediational g-formula based on the framework of the g-formula macro
9 (<http://www.hsph.harvard.edu/causal>) (Taubman et al., 2009). First, we specify parametric
10 regression models for the distribution of the time-varying exposures, mediators, confounders, and
11 survival variables. For each model, we include former covariates as the independent variables to
12 try to eliminate confounding and to ensure that the four assumptions (mentioned in Section 3)
13 hold. By fitting these models with data, maximum likelihood estimates for all parameters can be
14 obtained. Finally, these estimates are substituted for the parameters in the survival mediational g-
15 formula, deriving consistent estimates for IDE and IIE based on Monte-Carlo simulation.

16 The estimation algorithm is as follows:

17 (1) Fit parametric models for:

18 (1a) for $t > 1$, the joint density of the observed confounders, exposures, and mediators at t ,
19 given past covariate history and survival to $t-1$.

20 (1b) for all t , the probability of surviving by t given past covariate history and survival to
21 $t-1$.

22 (2) Estimate the joint distribution of the counterfactual mediators, $M(1:T)^*$, for survivors as
23 identified by Equation 4.

24 (2a) Set baseline ($t = 1$) covariates to the observed values for subject i . Recursively, for

- 1 each time $t = 1, \dots, T$ and each subject $i = 1, \dots, n$:
- 2 (2a.i) For $t \geq 1$, generate time t confounders, exposure, and mediator based on the
3 estimated model coefficients of (1a) and previously generated covariates
4 under the time-varying exposure intervention $a(1:t-1)$.
- 5 (2a.ii) Assign time t exposure under the intervention $a(1:T)$.
- 6 (2b) For each $t = 1, \dots, T$, randomly permute the n values of the joint mediators assigned
7 under intervention $a(1:T)$ in (2a). For each t , save this permutation for use in (3)
8 below (i.e., we obtain the random draw $G(1:T)^*$ of $M(1:T)^*$ amongst survivors in
9 this step).
- 10 (2c) Repeat (2a) replacing intervention $a(1:T)$ with $a(1:T)^*$.
- 11 (2d) Repeat (2b) replacing intervention $a(1:T)$ with $a(1:T)^*$ (i.e., we obtain the random
12 draw $G(1:T)^*$ of $M(1:T)^*$ amongst survivors in this step).
- 13 (3) Estimate $Q(a(1:T), a(1:T))$, $Q(a(1:T)^*, a(1:T))$, $Q(a(1:T), a(1:T)^*)$ and $Q(a(1:T)^*, a(1:T)^*)$ by
14 repeating the following for each $(a(1:T)_1, a(1:T)_2) = (a(1:T), a(1:T)), (a(1:T), a(1:T)^*), (a(1:T)^*,$
15 $a(1:T))$, and $(a(1:T)^*, a(1:T)^*)$:
- 16 (3a) Recursively for each time $t = 1, \dots, T$ and each subject $i = 1, \dots, n$:
- 17 (3a.i) Repeat (2a.i) but replacing “time-varying exposure intervention $a(1:T)$ through
18 $T-1$ ” with the joint “time-varying exposure and mediator intervention
19 $(a(1:T)_1, G(1:T)_{a(1:T)_2, S(T-1)=1})$ ”.
- 20 (3a.ii) Assign the time t mediator as the i th component of the permuted vector for
21 time t from (2b) (if $a(1:T)_2 = a(1:T)$) or (2d) (if $a(1:T)_2 = a(1:T)^*$).
- 22 (3a.iii) Assign time t exposure under the intervention $a(1:T)_1$.
- 23 (3a.iv) Estimate the the probability of surviving by t given past covariate history and
24 survival to $t-1$ given the generated history in (3.a.i) through (3.a.iii) based on

1 the estimated regression coefficients from (1b)
2 (3b) For each of the $i=1, \dots, n$, generated histories, take the product of the time t -specific
3 estimated probabilities in (3.a.iv) over all $t=1, \dots, T$
4 (3c) Estimate $Q(a(1:T)_1, a(1:T)_2)$ as the mean over the n products in (3b).
5

6 Estimates of the IDE and IIE are then calculated from the estimates of the four
7 $Q(a(1:T)_1, a(1:T)_2)$ in (3). 95% confidence intervals are calculated based on repeating the above
8 algorithm in 500 bootstrap samples of the original n observations. This algorithm can be
9 implemented with the `mgformula` macro, freely accessible with documentation at
10 <http://www.hsph.harvard.edu/causal/software/>.

11

12 **5. Analysis of Framingham Heart Study data**

13 In this section, we apply the survival mediational g -formula to the Framingham Heart Study
14 (FHS) dataset to investigate the causal mechanism of smoking on overall mortality mediated by
15 CAD. FHS is a longitudinal cohort study beginning in Framingham, Massachusetts in 1948. The
16 original cohort consisted of 5,209 participants between 30 and 62 years old without any symptom
17 of cardiovascular disease (CVD) at baseline. All the participants underwent examinations at the
18 beginning of the study and every two years during the follow up. For each examination, potential
19 CVD risk factors including socio-demographic data, lifestyle characteristics, detailed medical
20 history, physical examination data and blood samples were collected. Further details on the
21 design of FHS are available elsewhere (Dawber, Meadors and Moore Jr, 1951; Kannel et al.,
22 1961).

23 Exam 3 is specified as the first exam and exams 1 and 2 as pre-baseline covariates to
24 allow lag predictive models. As the analysis is intended only as an illustration we make some

1 simplifications in the analysis. We focus on only ten year follow up (i.e. exam 3 to exam 7; the
2 total exam number $T = 5$). Three exclusion criteria are listed below: (1) no record at baseline on
3 weight, height, smoking status, former smoking history, systolic blood pressure (SBP), or total
4 cholesterol level; (2) diagnosis of diabetes, cancer, or CVD at baseline; and (3) value for smoking
5 status or BMI missing more than once. After these exclusions, 3,116 participants are eligible for
6 analysis. For simplicity, we now refer to the original FHS exams 3,..., 7 as exams 1,..., 5.

7 The smoking status at all five exams, measured as self-reported average number of
8 cigarettes smoked per day, are exposures of interest ($A(1:5)$); mortality at the end of follow up is
9 the outcome of interest ($S(5)$); and the CAD status at all exams are the mediators of interest
10 ($M(1:5)$). For missing smoking status at one time point, we carry forward the last observed
11 smoking status for one exam period only. We considered "smoking 30 cigarettes per day" and "no
12 smoking" at all exams as two hypothetical intervention statuses, $A(1:T) = a(1:T)$ and $A(1:T) =$
13 $a^*(1:T)$. Time-varying covariates $L(1:T)$ include the exam number, the systolic blood pressure
14 (mm/hg), body mass index (kg/mm), and the usage of antihypertensive drugs. Baseline covariates
15 V include gender and age (years).

16 The parametric g-formula is used to estimate the total effect of smoking 30 cigarettes per
17 day (v.s. no smoking) on mortality at exam 5, by the g-formula SAS macro. The survival
18 mediational g-formula is applied to conduct mediation analysis with time-varying exposures,
19 mediators, and confounders by the mGFORMULA SAS macro. For conducting the survival
20 mediational g-formula, we specify model for the distribution of time-varying exposures,
21 mediators, confounders, and survival variables at each time point. We use current covariates and
22 covariates at one period back (one period lagged model) as the predictors. Specifically, for $t = 1,$
23 $2, \dots, 5$, we regress $S(t)$ on $A(t), M(t), L(t), A(t-1), M(t-1),$ and $L(t-1)$; regress $L(t)$ on $A(t), M(t),$
24 $A(t-1), M(t-1),$ and $L(t-1)$; regress $M(t)$ on $A(t), A(t-1), M(t-1),$ and $L(t-1)$; and regress $A(t)$ on

1 A(t-1), M(t-1), and L(t-1). All analyses are conducted using SAS 9.4 (Cary, NC).

2 The mortalities among the original group, a hypothetical group if everyone did not smoke
3 for 10 years, and a hypothetical group if everyone smoked 30 cigarettes per day for 10 years, are
4 4.52%, 3.37%, and 7.87 %, respectively (Table 1). The risk ratio (RR) is 2.34 (95 % CI = (1.44,
5 3.70)), which is close to the result calculated by mGFORMULA macro (RR = 2.30; 95 % CI =
6 (1.36, 2.88)). On the additive scale, smoking increases mortality by 3.96% directly not through
7 changing CAD, and by 0.34 % through changing CAD. The proportion of the total effect
8 explained by the mediation of CAD is 7.91 % (95 % CI= (1.36, 19.32)) (Table 2).

9

10 **6. Discussion**

11 This is the first paper providing a general framework for causal mediation analysis with a
12 survival outcome and with time-varying exposures, mediators, and confounders. We provided
13 identification results for the interventional direct and indirect effects and the required
14 assumptions, and used a parametric approach for the point and interval estimation. Based on the
15 g-formula macro, we have developed a feasible algorithm and a mGFORMULA SAS macro (Lin
16 et al., 2016; VanderWeele and Tchetgen Tchetgen, 2016 (accepted)). Similar to the g-formula
17 macro, we use Monte-Carlo simulation and bootstrapping for point and interval estimation,
18 respectively. Like other simulation-based methods (Imai, Keele and Tingley, 2010), our approach
19 has the advantage of allowing for very flexible models.

20 The survival mediational g-formula is currently the only method for investigating the
21 mechanisms of an effect on a survival outcome with time-varying exposures, mediators, and
22 confounders. Under traditional techniques for causal mediation analysis with survival outcome,
23 only one mediator is allowed, and this mediator should occur immediately after the exposure to
24 ensure that the identification assumption holds. Thus traditional techniques do not adequately

1 capture the indirect effect in settings in which the mediators vary over time (VanderWeele, 2015;
2 VanderWeele, 2009b). When both the exposure and confounders are fixed, our formula also
3 reduces to Zheng's and van der Laan's formula (Zheng and van der Laan, 2012), but our formula
4 can be used more generally when, in addition to the mediators, the exposures and confounders
5 also vary over time.

6 Several limitations of this study should be noted. First, our path-specific definition is
7 different from the traditional definition for direct and indirect effect. This is inevitable for causal
8 mediation analysis with time-varying mediators because the NDE/NIE (as well as IDE/IIE)
9 cannot be defined. Because the difference of traditional and alternative definitions is undefined or
10 zero, it may be considered reasonable to use this alternative definition. Second, the outcome only
11 focuses on survival probability at the end of follow up. Extending the outcome model to different
12 survival models such as the Cox proportional hazard model or the accelerated failure time model
13 could be developed in future research. Third, non-ignorable drop-out has not been adjusted for
14 here. In the illustration, we restricted our follow up to ten years to partially address this selection
15 bias. The relatively low proportion of loss to follow up (<20%) perhaps partially mitigates this
16 problem. Our method is also sensitive to the violation of model misspecification, which is the
17 trade-off in obtaining efficient estimates with a parametric approach. However, in our approach,
18 the allowance for using very flexible models (including splines) partially mitigates this issue.
19 Finally, the analysis is subject to potential violation of the confounding assumptions. Future
20 research could develop sensitivity analysis techniques for violations of these assumptions.

21 The survival mediational g-formula serves as a powerful and useful tool for mediation
22 analysis with longitudinal data and survival outcomes. When the outcome of interest is survival
23 variable, researchers can apply our method to disentangle the complicated causal mechanisms
24 arising from time-varying mediators, exposures, and confounders.

1

2 **Reference**

- 3 Avin, C., Shpitser, I., and Pearl, J. (2005). Identifiability of path-specific effects. *Department of*
4 *Statistics, UCLA*.
- 5 Dawber, T. R., Meadors, G. F., and Moore Jr, F. E. (1951). Epidemiological Approaches to Heart
6 Disease: The Framingham Study*. *American Journal of Public Health and the Nations Health* **41**,
7 279-286.
- 8 Frangakis, C. E., and Rubin, D. B. (2002). Principal stratification in causal inference. *Biometrics* **58**,
9 21-29.
- 10 Hernán, M. (2004). A definition of causal effect for epidemiological research. *Journal of*
11 *Epidemiology and Community Health* **58**, 265-271.
- 12 Imai, K., Keele, L., and Tingley, D. (2010). A general approach to causal mediation analysis.
13 *Psychological methods* **15**, 309.
- 14 Kannel, W. B., Dawber, T. R., Kagan, A., Revotskie, N., and Stokes, J. (1961). Factors of Risk in the
15 Development of Coronary Heart Disease—Six-Year Follow-up ExperienceThe Framingham Study.
16 *Annals of internal medicine* **55**, 33-50.
- 17 Lin, S.-H., Young, J., Tchetgen Tchetgen, E., and VanderWeele, T. J. (2016). Parametric
18 mediational g-formula approach to mediation analysis with time-varying exposures, mediators
19 and confounders: an application to smoking, weight, and blood pressure. *Epidemiology*
20 *(accepted)*.
- 21 Pearl, J. (2001). Direct and indirect effects. In *Proceedings of the seventeenth conference on*
22 *uncertainty in artificial intelligence*, 411-420: Morgan Kaufmann Publishers Inc.
- 23 Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys* **3**, 96-146.
- 24 Robins, J. (1986). A new approach to causal inference in mortality studies with a sustained
25 exposure period—application to control of the healthy worker survivor effect. *Mathematical*
26 *Modelling* **7**, 1393-1512.
- 27 Robins, J. M., and Greenland, S. (1992). Identifiability and exchangeability for direct and indirect
28 effects. *Epidemiology*, 143-155.
- 29 Taubman, S. L., Robins, J. M., Mittleman, M. A., and Hernán, M. A. (2009). Intervening on risk
30 factors for coronary heart disease: an application of the parametric g-formula. *International*
31 *Journal of Epidemiology* **38**, 1599-1611.
- 32 Tchetgen, E. J. T., and Shpitser, I. (2012). Semiparametric theory for causal mediation analysis:
33 efficiency bounds, multiple robustness and sensitivity analysis. *The Annals of Statistics* **40**, 1816-
34 1845.
- 35 Tchetgen Tchetgen, E. J. (2014). Identification and estimation of survivor average causal effects.

1 *Statistics in medicine* **33**, 3601-3628.
2 Valeri, L., and VanderWeele, T. J. (2013). Mediation analysis allowing for exposure–mediator
3 interactions and causal interpretation: Theoretical assumptions and implementation with SAS
4 and SPSS macros. *Psychological methods* **18**, 137.
5 van der Laan, M. J., and Petersen, M. L. (2008). Direct effect models. *The international journal of*
6 *biostatistics* **4**.
7 VanderWeele, T. (2015). *Explanation in Causal Inference: Methods for Mediation and Interaction*:
8 Oxford University Press.
9 VanderWeele, T., and Vansteelandt, S. (2009). Conceptual issues concerning mediation,
10 interventions and composition. *Statistics and its Interface* **2**, 457-468.
11 VanderWeele, T. J. (2009a). Concerning the consistency assumption in causal inference.
12 *Epidemiology* **20**, 880-883.
13 VanderWeele, T. J. (2009b). Marginal structural models for the estimation of direct and indirect
14 effects. *Epidemiology* **20**, 18-26.
15 VanderWeele, T. J. (2010). Bias formulas for sensitivity analysis for direct and indirect effects.
16 *Epidemiology (Cambridge, Mass.)* **21**, 540.
17 VanderWeele, T. J., and Tchetgen Tchetgen, E. (2016 (accepted)). Mediation Analysis with Time-
18 Varying Exposures and Mediators. *Journal of the Royal Statistical Society: Series B (Statistical*
19 *Methodology)*.
20 VanderWeele, T. J., and Vansteelandt, S. (2010). Odds ratios for mediation analysis for a
21 dichotomous outcome. *American journal of epidemiology* **172**, 1339-1348.
22 VanderWeele, T. J., Vansteelandt, S., and Robins, J. M. (2014). Effect decomposition in the
23 presence of an exposure-induced mediator-outcome confounder. *Epidemiology* **25**, 300-306.
24 Zheng, W., and van der Laan, M. J. (2012). Causal mediation in a survival setting with time-
25 dependent mediators.

26
27
28



1 **Table 1.** Estimates of the overall effect of smoking 30 cigarettes per day for 10 year (compared
 2 with no smoking) on mortality.

Intervention	Mortality (%)	Difference	95% CI	Ratio	95% CI
No intervention	4.52	1.15	0.58, 2.09	1.34	1.16, 1.69
No smoking	3.37	Ref.		Ref.	
30 cigarettes/day	7.87	4.90	1.59, 8.18	2.34	1.44, 3.70

3 CI: confident interval;

4

5

6

7



- 1 **Table 2.**Mediation analysis for the effect of smoking 30 cigarettes per day for 10 years
 2 (compared with no smoking) on overall mortality, mediated by cardiovascular disease.

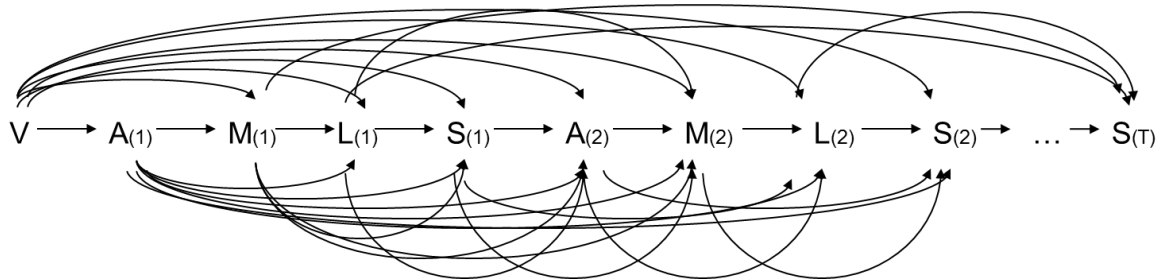
	Estimate	95% CI
Additive scale		
Total effect	4.3	1.37, 6.30
Direct effect	3.96	1.22, 6.06
Indirect effect	0.34	0.05, 0.96
Multiplicative scale		
Total effect	2.30	1.36, 2.88
Direct effect	2.20	1.33, 2.70
Indirect effect	1.05	1.01, 1.12
Proportion mediated (%)	7.91	1.36, 19.32

3 CI: confident interval;

4



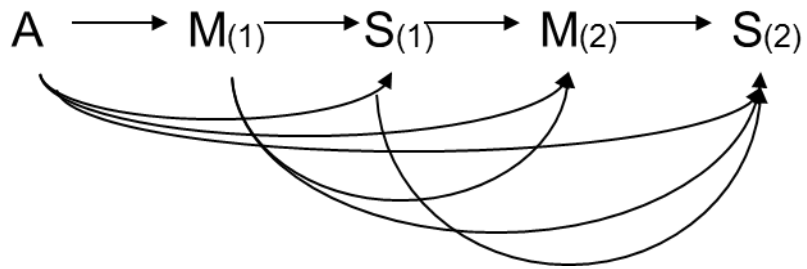
1 **Figure 1.** Time-varying mediation with variables of $A(t)$, $M(t)$, $L(t)$, and $S(t)$, for $t = 1$ to T .



2
3



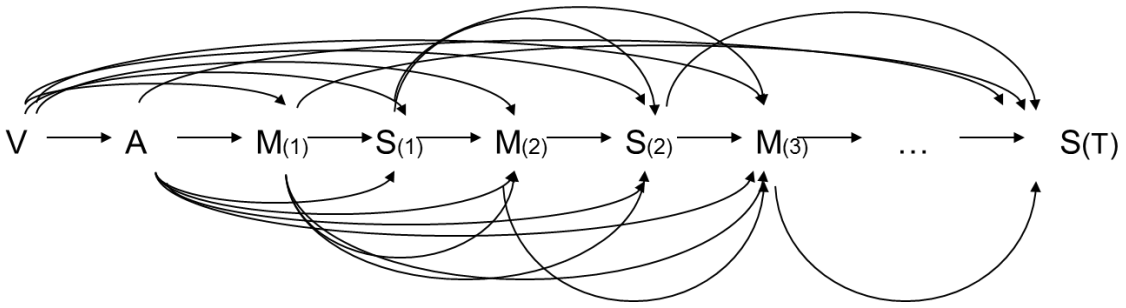
- 1 **Figure 2.** Simple model for survival mediation analysis with fixed exposure and time-varying
2 mediators with 2 time points.



- 3
4
5



1 **Figure 3.** Simple model for survival mediation analysis with fixed exposure and time-varying
2 mediators.



3
4
5
6

