eAppendix for: "SAS macro for causal mediation analysis with survival data"

Linda Valeri and Tyler J. VanderWeele

1 Causal effects under the counterfactual framework and their estimators

We let T_a and M_a denote respectively the values of the time-to-event outcome and mediator that would have been observed had the exposure A been set to level a. We let T_{am} denote the value of the time-to-event outcome that would have been observed had the exposure, A, and mediator, M, been set to levels a and m, respectively. The average controlled direct effect comparing exposure level a to a^* and fixing the mediator to level m on the mean survival time ratio scale is defined by $CDE_{a,a^*}(m) = E[T_{am}]/E[T_{a^*m}]$. The average natural direct effect is then defined by $NDE_{a,a^*}(a^*) = E[T_{aM_{a^*}}]/E[T_{a^*M_{a^*}}]$. The average natural indirect effect can be defined as $NIE_{a,a^*}(a) = E[T_{aM_a}]/E[T_{aM_{a^*}}]$, which compares the effect of the mediator at levels M_a and M_{a^*} on the survival outcome when exposure A is set to a. Controlled direct effects and natural direct and indirect effects within strata of $C = c$ are then defined by: $CDE_{a,a^*|c}(m) = E[T_{am}|c]/E[T_{a^*m}|c]$, $NDE_{a,a^*|c}(a^*) = E[T_{aM_{a^*}}|c]/E[T_{a^*M_{a^*}}|c]$ and $NIE_{a,a^*|c}(a) = E[T_{aM_a}|c]/E[T_{aM_{a^*}}|c]$ respectively. For an arbitrary time-to-event variable V, let $\lambda_V(t)$ and $\lambda_V(t|c)$ denote the hazard or hazard conditional on covariates c at time t, that is the instantaneous rate of the event conditional on $V \geq t$. The causal effects can also be defined on the hazard ratio scale, replacing $E[\cdot]$ with $\lambda(\cdot)$.

If we let $X \perp T|Z$ denote that X is independent of T conditional on Z, then the identification assumptions for the causal effects previously defined can be expressed formally in terms of counterfactual independence as (i) $T_{am} \perp A|C$, (ii) $T_{am} \perp M|\{A, C\}$, (iii) $M_a \perp A|C$, and (iv) $T_{am} \perp M_{a^*}|C$. Assumptions (i) and (ii) suffice to identify controlled direct effects; assumptions (i)-(iv) suffice to identify natural direct and indirect effects (Pearl, 2001; VanderWeele and Vansteelandt, 2009). The intuitive interpretation of these assumptions follows from the theory of causal diagrams (Pearl, 2001). Alternative identification assumptions have also been proposed (Imai 2010; Hafeman and VanderWeele, 2011). However, it has been shown that the intuitive graphical interpretation of these alternative assumptions are in fact equivalent (Shpitser and VanderWeele, 2011). Technical examples can be constructed where one set of identification assumptions holds and another does not, but on a causal diagram corresponding to a set of non-parametric structural equations, whenever one set of the assumptions among those in VanderWeele and Vansteelandt (2009), Imai (2010), and Hafeman and VanderWeele (2011) holds, the others will also.

Models and Estimators for Causal Effects: Continuous Mediator and Time-to-event Outcome

Let M be a continuous mediator following a linear model, A be an exposure and $\boldsymbol{C}$ be additional covariates. Assume that the outcome T is a failure time variable following a Cox-proportional hazard model or an accelerated failure time (AFT) model. We can define the mediator regression as

$$E(M|A,C) = \beta_0 + \beta_1 a + \beta_2' c \tag{1}$$

We define the outcome model either as Cox proportional hazard:

$$\lambda_T(t|a,m,c) = \lambda_T(t|0,0,0)e^{\gamma_1 a + \gamma_2 m + \gamma_3 am + \gamma_4' c} \tag{2}$$

or as accelerated failure time model:

$$log(T) = \theta_0 + \theta_1 a + \theta_2 m + \theta_3 am + \theta_4' c + \nu\epsilon \tag{3}$$

where ϵ follows the extreme value distribution and ν is a shape parameter taking value $\nu = 1$ if the exponential distribution is assumed and allowed to take different values if the Weibull distribution is assumed.

If the no-unmeasured confounding assumptions hold, the model for the continuous mediator and for the outcome are correctly specified, and the Cox regression is employed, then controlled direct effect, natural direct and natural indirect effects estimators on the hazard ratio scale are given by (VanderWeele, 2011):

$$\begin{aligned}
\lambda_{T_{am}}(t|c)/\lambda_{T_{a^*m}}(t|c) &= exp\{(\gamma_1 + \gamma_3 m)(a - a^*)\} \\
\lambda_{T_{aM_{a^*}}}(t|c)/\lambda_{T_{a^*M_{a^*}}}(t|c) &\approx exp[\{\gamma_1 + \gamma_3(\beta_0 + \beta_1 a^* + \beta_2' c + \gamma_2\sigma^2)\}(a - a^*) + 0.5\gamma_3^2\sigma^2(a^2 - a^{*2})] \\
\lambda_{T_{aM_a}}(t|c)/\lambda_{T_{aM_{a^*}}}(t|c) &\approx exp\{(\gamma_2\beta_1 + \gamma_3\beta_1 a)(a - a^*)\}
\end{aligned}$$

The approximations to estimate natural direct and natural indirect effects apply if the outcome is rare at the end of follow-up.

If the no-unmeasured confounding assumptions hold, the model for the continuous mediator and for the outcome are correctly specified, and the AFT model is employed, then controlled direct effect, natural direct and natural indirect effects estimators on the mean survival time ratio scale are given by (VanderWeele, 2011):

$$\begin{aligned}
E[T_{am}]/E[T_{a^*m}] &= exp\{(\theta_1 + \theta_3 m)(a - a^*)\} \\
E[T_{aM_{a^*}}]/E[T_{a^*M_{a^*}}] &= exp[\{\theta_1 + \theta_3(\beta_0 + \beta_1 a^* + \beta_2' c + \theta_2\sigma^2)\}(a - a^*) + 0.5\theta_3^2\sigma^2(a^2 - a^{*2})] \\
E[T_{aM_a}]/E[T_{aM_{a^*}}] &= exp\{(\theta_2\beta_1 + \theta_3\beta_1 a)(a - a^*)\}
\end{aligned}$$

Note that under the AFT the rare outcome assumption is not required to estimate natural direct and indirect effects.

Models and Estimators for Causal Effects: Binary Mediator and Time-to-event Outcome

Let T denote the time-to-event outcome modeled either as in (2) or (3). Let M be a binary mediator following a logistic model.

We can define the mediator regression as:

$$logit\{P(M=1|A,C)\} = \beta_0 + \beta_1 a + \beta_2' c \qquad (4)$$

If the no-unmeasured confounding assumptions hold, the model for the binary mediator and for the outcome are correctly specified, and the Cox regression is employed, then controlled direct effect, natural direct and natural indirect effects estimators on the hazard ratio scale are given by (see eAppendix section 4 for proof):

$$\begin{aligned}
\lambda_{T_{am}}(t|c)/\lambda_{T_{a^*m}}(t|c) &= exp\{(\gamma_1+\gamma_3 m)(a-a^*)\} \\
\lambda_{T_{aM_{a^*}}}(t|c)/\lambda_{T_{a^*M_{a^*}}}(t|c) &\approx \frac{\exp(\gamma_1 a)\{1+\exp(\gamma_2+\gamma_3 a+\beta_0+\beta_1 a^*+\beta_2' c)\}}{\exp(\gamma_1 a^*)\{1+\exp(\gamma_2+\gamma_3 a^*+\beta_0+\beta_1 a^*+\beta_2' c)\}} \\
\lambda_{T_{aM_a}}(t|c)/\lambda_{T_{aM_{a^*}}}(t|c) &\approx \frac{\{1+\exp(\beta_0+\beta_1 a^*+\beta_2' c)\}\{1+\exp(\gamma_2+\gamma_3 a+\beta_0+\beta_1 a+\beta_2' c)\}}{\{1+\exp(\beta_0+\beta_1 a+\beta_2' c)\}\{1+\exp(\gamma_2+\gamma_3 a+\beta_0+\beta_1 a^*+\beta_2' c)\}}
\end{aligned}$$

The approximations to estimate natural direct and natural indirect effects apply if the outcome is rare at the end of follow-up.

If the no-unmeasured confounding assumptions hold, the model for the binary mediator and for the outcome are correctly specified, and the AFT model is employed, then controlled direct effect, natural direct and natural indirect effects estimators on the mean survival ratio scale are given by (see eAppendix section 4):

$$\begin{aligned}
E[T_{am}]/E[T_{a^*m}] &= exp\{(\theta_1+\theta_3 m)(a-a^*)\} \\
E[T_{aM_{a^*}}]/E[T_{a^*M_{a^*}}] &= \frac{\exp(\theta_1 a)\{1+\exp(\theta_2+\theta_3 a+\beta_0+\beta_1 a^*+\beta_2' c)\}}{\exp(\theta_1 a^*)\{1+\exp(\theta_2+\theta_3 a^*+\beta_0+\beta_1 a^*+\beta_2' c)\}} \\
E[T_{aM_a}]/E[T_{aM_{a^*}}] &= \frac{\{1+\exp(\beta_0+\beta_1 a^*+\beta_2' c)\}\{1+\exp(\theta_2+\theta_3 a+\beta_0+\beta_1 a+\beta_2' c)\}}{\{1+\exp(\beta_0+\beta_1 a+\beta_2' c)\}\{1+\exp(\theta_2+\theta_3 a+\beta_0+\beta_1 a^*+\beta_2' c)\}}
\end{aligned}$$

Note that under the AFT the rare outcome assumption is not required to estimate natural direct and indirect effects.

2 Description of the SAS macro

The present software is designed to enable the investigator to easily implement mediation analysis in the presence of exposure-mediator interaction accounting for different types of outcomes (normal, dichotomous (with logit or log link), poisson, negative binomial, failure time (with Cox PH model or AFT Weibull or AFT exponential), can be assumed) and mediators of interest (normal or dichotomous with logit link). The logit link for dichotomous outcomes and Cox model for survival outcome should only be used if the outcome is rare. If the outcome is not rare and binary the log link can be used (though the outcome model may not always converge) and failure times could be modeled using AFT. For binary outcomes, in the case of logit link the direct and indirect effects are on the odds ratio scale and in the case of using the log link the direct and indirect effects are on the risk ratio scale. For survival outcomes, in the case of Cox model the direct and indirect effects are on the hazard ratio scale and in the case of using the AFT model the direct and indirect effects are on the mean survival ratio scale. This SAS macro, provides estimates, and confidence intervals for the direct and indirect effects defined in Valeri and VanderWeele (2013). The estimates assume the model assumptions are correct and the identifiability assumptions discussed in Valeri and VanderWeele (2013) hold.

In order to implement mediation analysis via the *mediation macro* in SAS the investigator first inputs the data which has to include the outcome, treatment and mediator variables. The investigator inputs also the covariates to be adjusted for in the model. Macro activation requires the investigator to save the macro script and input information in the statement

```
%mediation(data=,yvar=,avar=,mvar=,cvar=,a0=,a1=,m=,yreg=, mreg=,interaction=,cens=)
run;
```

First one inputs the name of the dataset, then the name of the outcome variable ($yvar =$), the treatment variable ($avar =$), the mediator variable ($mvar =$), the other covariates ($cvar =$). Then the investigator needs to specify the baseline level of the exposure a^* ($a0 =$), the new exposure level a ($a1 =$) and the level of mediator m at which the controlled direct effect is to be estimated. The user must also specify which types of regression have to be implemented. In particular either *linear*, *logistic*, *loglinear*, *poisson*, *negbin*, *survCox*, *survAFT_exp*, or *survAFT_weibull* can be specified in the option yreg. If *survCox*, *survAFT_exp*, or *survAFT_weibull* for failure time outcome are specified, the user need to specify the censoring variable ($cens =$) which has to take value 1 if the event is censored and 0 if the failure time is observed. For the option mreg either linear or logistic regressions are allowed.

Finally, the analyst needs to specify whether an exposure-mediator interaction is present (true or false). The software provides the following output: first the outcome and mediator regression output are provided. The output in the SAS macro is derived from the procedures of proc reg when the variable is continuous and proc logistic when the variable is binary. When the outcome is specified as poisson, negative binomial or log-linear the procedure proc genmod is employed. When the outcome is failure time and the Cox model is specified, the procedure phreg is employed while if accelerated failure time model is specified, the procedure lifereg is employed.
The SAS macro is case-sensitive and the options specified should be given in lower-case letters, unless otherwise specified.

A table with direct and indirect effects together with total effects and proportion mediated follows. The total effect is computed as the sum of the natural direct effect and the natural indirect effect when the outcome is continuous and the product of the natural direct and indirect effects in the other cases. The proportion mediated can be defined as the ratio of the natural indirect effect over the total effect when the outcome is continuous; the proportion mediated on outcome difference scale is given in the other cases using a transformation of the ratio scale (VanderWeele and Vansteelandt, 2010). The effects are reported for the mean level of the covariate C. The table contains p-values, and confidence intervals for each effect.

The reduced output is the default option. The table will just display controlled direct effect, natural direct effect, natural indirect effect, total effect and proportion mediated. When the option *output=full* is used, both conditional effects and effects evaluated at the mean covariate levels are shown. When output equal to full is chosen as an option, the investigator must enter fixed values for the covariates C at which to compute conditional effects. The macro statement is then as follows:

```
%mediation(data=,yvar=,avar=,mvar=,cvar=,a0=,a1=,m=,yreg=,mreg=,interaction=,cens=,output=,c=)
run;
```

When these commands are added, in addition to controlled direct effect, and the natural direct and indirect effect described above, two other effects are displayed. The natural direct and indirect effects we have been considering are sometimes called the "pure" natural direct effect and the "total" natural direct effect (Robins, Greenland, 1992). We can also consider the "total" natural direct effect and the "pure" natural indirect effect. The total natural direct effect expresses how much the outcome would change on average if the exposure changed from level $a^* = 0$ to level $a = 1$, but the mediator for each individual was fixed at the natural level which would have taken at exposure level $a = 1$. The pure natural indirect effect expresses how much the outcome would change on average if the exposure were controlled at level $a^* = 0$ but the mediator were changed from the natural level it would take if $a^* = 0$ to the level that would have taken at exposure level $a = 1$. These effects are also reported if the user selects $output = full$. For an in-depth description of these causal effects the interested reader can refer to VanderWeele and Vansteelandt (2009).

Finally, the investigator has the option of implementing mediation analysis when data arise from a case-control design, provided the outcome in the population is rare. The formulas for the effects remain the same, however the mediator regression will be run only for controls, in order to minimize the bias due to the design, and since with a rare outcome Y the controls will approximate the distribution of M in the population. To do so the option *casecontrol=true* can be used. In this case the macro statement changes to

```
%mediation(data=,yvar=,avar=,mvar=,cvar=,a0=,a1=,m=,yreg=,mreg=,interaction=,cens=,casecontrol=)
run;
```

Finally, the investigator can choose whether to obtain standard errors and confidence intervals via the delta method or a bootstrapping technique. The default is the delta method. To use bootstrapping the option boot true can be given. In this case the macro will compute 1,000 bootstrap samples from which causal effects are obtained along with their standard errors (SE) and percentile confidence intervals. If the investigator wishes to use a higher number of bootstrap samples, instead of "true" he or she inputs the number of bootstrap samples desired (e.g., boot = 5000 would estimate standard errors and confidence intervals using 5,000 bootstrap samples). When using the bootstrap the macro statement changes to

```
%mediation(data=,yvar=,avar=,mvar=,cvar=,a0=,a1=,m=,yreg=,mreg=,interaction=,cens=,boot=)
run;
```

Note that if the investigator wants to add a categorical variable as covariate, this must be recoded as a series of indicator variables.

3 Example

We present in this section an example of using the mediation macro when the outcome is failure time. The example is to be considered for illustration purposes only. Suppose interest lies in the effects of socio-economic position (SEP) on survival of colorectal cancer patients. SEP is measured by percentage of people living below the poverty line in the patients' county of residence. In this example, stage at diagnosis, measured as advanced vs non-advanced is investigated as a potential mediator of the relationship between residing in poor counties and the survival outcome. Data for Surveillance Epidemiology End Results (SEER) linked to American Community Survey (ACS) data for patients diagnosed in 1992-2005 and followed up to 2010 is employed to address this question.

In the present analysis we allow for interaction between the SEP measure and stage at diagnosis. Furthermore we adjust for potential confounders: gender, age at diagnosis, year at diagnosis and state of residence. Since the outcome is failure time we define a censoring variable taking value 1 if the individual is censored or value 0 if the event is observed.

After having saved the dataset and inserted macro script we run the following command

```
%mediation(data=dat,yvar=survival,avar=new_poverty,mvar=grade,cvar=race date_c sex age_c state_2 state_3 state_4

state_5 state_6 state_7 state_8 , a0=0,a1=0.3,m=0,yreg=survAFT_exp,mreg=logistic,interaction=true, cens=censor)
run;
```

The first output provided is the results of the outcome and mediator regressions. Then the direct and indirect effects follow. We give the reduced output which provides estimates for the controlled direct effect, the natural direct and indirect effect, total effect and percentage mediated. We fit the survival outcome with accelerated failure time model since the event is non rare by the end of follow-up in this population. By inspecting Kaplan-Meier survival curves by quartiles of the exposure, the assumption of proportional hazard appears satisfied and we therefore fit the AFT model assuming exponential distribution. The survival analysis yields a negative effect of poverty on survival, marginally significant adjusting for grade. A positive, significant interaction between tumor stage at diagnosis and poverty is detected. The logistic regression analysis shows that the SEP measure is positively associated with stage at diagnosis. We

The SAS System

The LIFEREG Procedure

Analysis of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	4.6888	0.0104	4.6684	4.7092	202593	<.0001
new_poverty	1	-0.1941	0.1110	-0.4117	0.0235	3.06	0.0804
grade	1	-1.9766	0.0068	-1.9900	-1.9633	84535.5	<.0001
int	1	-0.8043	0.1374	-1.0737	-0.5349	34.24	<.0001
RACE	1	-0.2530	0.0087	-0.2701	-0.2359	844.10	<.0001
date_c	1	0.0106	0.0010	0.0086	0.0125	111.33	<.0001
SEX	1	0.1715	0.0053	0.1611	0.1819	1040.26	<.0001
age_c	1	-0.0492	0.0002	-0.0497	-0.0487	40841.6	<.0001
state_2	1	0.1033	0.0431	0.0189	0.1877	5.75	0.0164
state_3	1	0.0168	0.0096	-0.0019	0.0356	3.09	0.0787
state_4	1	-0.0640	0.0093	-0.0821	-0.0458	47.67	<.0001
state_5	1	-0.0533	0.0146	-0.0819	-0.0248	13.41	0.0003
state_6	1	-0.0583	0.0150	-0.0876	-0.0290	15.20	<.0001
state_7	1	0.0219	0.0102	0.0019	0.0420	4.59	0.0322
state_8	1	0.0053	0.0075	-0.0094	0.0200	0.50	0.4808
Scale	0	1.0000	0.0000	1.0000	1.0000		
Weibull Shape	0	1.0000	0.0000	1.0000	1.0000		

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.5229	0.0213	5098.1448	<.0001
new_poverty	1	0.8926	0.1884	22.4510	<.0001
RACE	1	0.2881	0.0170	286.4012	<.0001
date_c	1	-0.0116	0.00209	30.9004	<.0001
SEX	1	-0.00647	0.0109	0.3504	0.5539
age_c	1	-0.00754	0.000417	327.0952	<.0001
state_2	1	0.0806	0.0802	1.0103	0.3148
state_3	1	0.0801	0.0201	15.8235	<.0001
state_4	1	0.1100	0.0192	32.7443	<.0001
state_5	1	0.0443	0.0303	2.1451	0.1430
state_6	1	-0.0257	0.0313	0.6728	0.4121
state_7	1	0.0353	0.0213	2.7512	0.0972
state_8	1	0.1267	0.0156	66.3337	<.0001

Figure 1: Output of accelerated failure time outcome regression and mediator logistic regression

Obs	Effect	Estimate	p_value	_95__CI_lower	_95__CI_upper
1	cde	0.94344	0.59996	0.75896	1.17276
2	nde	0.93693	0.04717	0.87855	0.99919
3	nie	0.95126	0.00001	0.93017	0.97284
4	total effect	0.89127	0.00089	0.83278	0.95386

Obs	Effect	Estimate
1	proportion mediated	0.41995

Figure 2: Output of mediation analysis with causal effects estimated for a change in the exposure from 0 to 0.3 and at the mean level of the covariates

estimate a total effect of 0.89 indicating that the mean survival time of individuals living in counties with 30% of the population living below the poverty level is 89% that of individuals living in counties with no individuals living below the poverty line. The controlled direct effect, controlling the mediator at level $m = 0$, reveals that, had we intervened setting stage at diagnosis to be non-advanced for all individuals, the mean survival time of individuals living in counties with 30% of the population living below the poverty level would be 94% that of individuals living in counties with no individuals living below the poverty line. Natural direct and natural indirect effects take values 93% and 95%, respectively and stage at diagnosis is found to mediate 42% of the effect of poverty on survival. The results of this example should be interpreted with caution as several biases might be present. First, an aggregate, rather than an individual-level measure of socio-economic position is employed, potentially introducing ecologic bias. Second, the no-unmeasured confounding assumptions are likely violated in this example.

4 Estimators for direct and indirect when the outcome is failure time and the mediator is binary

In this section we derive estimators of direct and indirect causal effects when the outcome is failure time modeled using either Cox proportional hazard model or accelerated failure time models (AFT) and the binary mediator is modeled using logistic regression. The results coincide with the ones obtained for direct and indirect effects when outcome and mediator are binary as derived in Valeri and VanderWeele (2013). The causal effects have hazard ratio interpretation if the Cox model is employed and mean survival ratio interpretation under the AFT model. Notably, the derivations under the Cox model assume that the failure event is rare while this assumption is not required if AFT is employed.

Models

Let M be a binary mediator following a logistic model, A be an exposure and $\boldsymbol{C}$ be additional covariates. Assume that the outcome T is a failure time variable following a cox-proportional hazard model or an accelerated failure time model. We can define the mediator regression as

$$logit\{P(M = 1|A, C)\} = \beta_0 + \beta_1 a + \beta_2' c \tag{5}$$

We define the outcome model either as

$$\lambda_T(t|a, m, c) = \lambda_T(t|0, 0, 0)e^{\gamma_1 a + \gamma_2 m + \gamma_3 am + \gamma_4' c} \tag{6}$$

or as

$$log(T) = \theta_0 + \theta_1 a + \theta_2 m + \theta_3 am + \theta_4' c + \nu\epsilon \tag{7}$$

where ϵ follows the extreme value distribution and ν is a shape parameter taking value $\nu = 1$ if the exponential distribution is assumed and allowed to take different values if the Weibull distribution is assumed.

Cox-proportional hazard model for the outcome

Consider $\lambda_{T_{aM_{a^*}}}(t|c) = \frac{f_{T_{aM_{a^*}}}(t|c)}{S_{T_{aM_{a^*}}}(t|c)}$

where,

$f_{T_{aM_{a^*}}}(t|c) = \int f_{T_{am}}(t|M_{a^*} = m, c)dP_{M_{a^*}}(m|c)$

$= \int f_{T_{am}}(t|c)dP_{M_{a^*}}(m|c)$ by assumption (iv)

$= \int f_T(t|a, m, c)dP_M(m|a^*, c)$ by assumptions $(i) - (iii)$

$= \int \lambda_T(t|0, 0, 0)e^{\gamma_1 a+\gamma_2 m+\gamma_3 am+\gamma_4' c} exp\{-\Lambda_T(t|0, 0, 0)e^{\gamma_1 a+\gamma_2 m+\gamma_3 am+\gamma_4' c}\}dP_M(m|a^*, c)$

where $\Lambda_T(t|0, 0, 0) = \int_0^t \lambda_T(t|0, 0, 0)dt$, and

$S_{T_{aM_{a^*}}}(t|c) = \int exp\{-\Lambda_T(t|0, 0, 0)e^{\gamma_1 a+\gamma_2 m+\gamma_3 am+\gamma_4' c}\}dP_M(m|a^*, c)$.

Thus,

$$\lambda_{T_{aM_{a^*}}}(t|c) = \frac{\int \lambda_T(t|0,0,0)e^{\gamma_1 a+\gamma_2 m+\gamma_3 am+\gamma_4' c} exp\{-\Lambda_T(t|0,0,0)e^{\gamma_1 a+\gamma_2 m+\gamma_3 am+\gamma_4' c}\}dP_M(m|a^*,c)}{\int exp\{-\Lambda_T(t|0,0,0)e^{\gamma_1 a+\gamma_2 m+\gamma_3 am+\gamma_4' c}\}dP_M(m|a^*,c)}$$

$$\sim \lambda_T(t|0, 0, 0)e^{\gamma_1 a+\gamma_4' c} \int e^{(\gamma_2+\gamma_3 a)m} dP_M(m|a^*, c) = \lambda_T(t|0, 0, 0)e^{\gamma_1 a+\gamma_4' c} E[e^{(\gamma_2+\gamma_3 a)M}]$$

The approximation holds if we assume rare outcome so that $\Lambda_T(t|0, 0, 0) \sim 0$ which implies $exp\{-\Lambda_T(\cdot)e^{(\cdot)}\} \sim 1$.

Now, considering that $M \sim Be(p)$ with $p = \frac{exp(\beta_0+\beta_1 a+\beta_2' c)}{1+exp(\beta_0+\beta_1 a+\beta_2' c)}$, by the properties of the moment generating functions

$$E[e^{tM}] = 1 - p + pe^t$$

and we finally obtain

$$\lambda_{T_{aM_{a^*}}}(t|c) = \lambda_T(t|0, 0, 0)e^{\gamma_1 a+\gamma_4' c}(1 - \frac{exp(\beta_0 + \beta_1 a^* + \beta_2' c)}{1 + exp(\beta_0 + \beta_1 a^* + \beta_2' c)} + \frac{exp(\beta_0 + \beta_1 a^* + \beta_2' c + \gamma_2 + \gamma_3 a)}{1 + exp(\beta_0 + \beta_1 a^* + \beta_2' c)})$$

Define the conditional total hazard ratio by $\lambda_c^{TE} = \frac{\lambda_{T_{aM_a}}}{\lambda_{T_{a^*M_{a^*}}}}$. Likewise, we define the pure direct effect hazard ratio by $\lambda_c^{NDE} = \frac{\lambda_{T_{aM_{a^*}}}}{\lambda_{T_{a^*M_{a^*}}}}$ and the total indirect effect hazard ratio by $\lambda_c^{NIE} = \frac{\lambda_{T_{aM_a}}}{\lambda_{T_{aM_{a^*}}}}$. Then we have the following decomposition can be given:

$$\lambda_c^{TE} = \lambda_c^{NDE} \times \lambda_c^{NIE}$$

and

$$\lambda_c^{NDE} = \frac{\lambda_{T_{aM_{a^*}}}}{\lambda_{T_{a^*M_{a^*}}}} = e^{\gamma_1(a-a^*)} \frac{1 - \frac{exp(\beta_0+\beta_1 a^*+\beta_2' c)}{1+exp(\beta_0+\beta_1 a^*+\beta_2' c)} + \frac{exp(\beta_0+\beta_1 a^*+\beta_2' c+\gamma_2+\gamma_3 a)}{1+exp(\beta_0+\beta_1 a^*+\beta_2' c)}}{1 - \frac{exp(\beta_0+\beta_1 a^*+\beta_2' c)}{1+exp(\beta_0+\beta_1 a^*+\beta_2' c)} + \frac{exp(\beta_0+(\beta_1+\gamma_3) a^*+\beta_2' c+\gamma_2)}{1+exp(\beta_0+\beta_1 a^*+\beta_2' c)}}$$

$$\lambda_c^{NIE} = \frac{\lambda_{T_{aM_a}}}{\lambda_{T_{aM_{a^*}}}} = \frac{1 - \frac{exp(\beta_0+\beta_1 a+\beta_2' c)}{1+exp(\beta_0+\beta_1 a+\beta_2' c)} + \frac{exp(\beta_0+\beta_1 a+\beta_2' c+\gamma_2+\gamma_3 a)}{1+exp(\beta_0+\beta_1 a+\beta_2' c)}}{1 - \frac{exp(\beta_0+\beta_1 a^*+\beta_2' c)}{1+exp(\beta_0+\beta_1 a^*+\beta_2' c)} + \frac{exp(\beta_0+\beta_1 a^*+\gamma_3 a+\beta_2' c+\gamma_2)}{1+exp(\beta_0+\beta_1 a^*+\beta_2' c)}}$$

Accelerated failure time model for the outcome

Consider the counterfactual $T_{aM_{a^*}}$.

$E[T_{aM_{a^*}}] = \int E[T_{am}|M_{a^*}, c] dP_{M_{a^*}}(m|c)$

$= \int E[T_{am}|c] dP_{M_{a^*}}(m|c)$ by assumption (iv)

$= \int E[T|a, m, c] dP_M(m|a^*, c)$ by assumptions $(i) - (iii)$

$= \int E[e^{\theta_0+\theta_1 a+\theta_2 m+\theta_3 am+\theta_4' c+\nu\epsilon}] dP_M(m|a^*, c)$

$= e^{\theta_0+\theta_1 a+\theta_4' c} E[e^{\nu\epsilon}] E[e^{(\theta_2+\theta_3 a)M}]$

$= e^{\theta_0+\theta_1 a+\theta_4' c} E[e^{\nu\epsilon}](1 - \frac{exp(\beta_0+\beta_1 a^*+\beta_2' c)}{1+exp(\beta_0+\beta_1 a^*+\beta_2' c)} + \frac{exp(\beta_0+\beta_1 a^*+\beta_2' c+\theta_2+\theta_3 a)}{1+exp(\beta_0+\beta_1 a^*+\beta_2' c)})$

If we consider the mean survival time conditional on the vector of covariates C as our measure of interest, the following decomposition can be given

$$E(T_{aM_a})\}/\{E(T_{a^*M_{a^*}}) = [\{E(T_{aM_{a^*}})\}/\{E(T_{a^*M_{a^*}})\}] \times [\{E(T_{aM_a})\}/\{E(T_{a^*M_{a^*}})\}]$$

where the total effect, given by $\{E(T_{aM_a})\}/\{E(T_{a^*M_{a^*}})\}$, decomposes in pure direct effect

$$\{E(T_{aM_{a^*}})\}/\{E(T_{a^*M_{a^*}})\} = e^{\theta_1(a-a^*)} \frac{1 - \frac{exp(\beta_0+\beta_1 a^*+\beta_2' c)}{1+exp(\beta_0+\beta_1 a^*+\beta_2' c)} + \frac{exp(\beta_0+\beta_1 a^*+\beta_2' c+\theta_2+\theta_3 a)}{1+exp(\beta_0+\beta_1 a^*+\beta_2' c)}}{1 - \frac{exp(\beta_0+\beta_1 a^*+\beta_2' c)}{1+exp(\beta_0+\beta_1 a^*+\beta_2' c)} + \frac{exp(\beta_0+(\beta_1+\theta_3) a^*+\beta_2' c+\theta_2)}{1+exp(\beta_0+\beta_1 a^*+\beta_2' c)}}$$

times total indirect effect

$$\{E(T_{aM_a})\}/\{E[(T_{aM_{a^*}})\} = \frac{1 - \frac{exp(\beta_0+\beta_1 a+\beta_2' c)}{1+exp(\beta_0+\beta_1 a+\beta_2' c)} + \frac{exp(\beta_0+\beta_1 a+\beta_2' c+\theta_2+\theta_3 a)}{1+exp(\beta_0+\beta_1 a+\beta_2' c)}}{1 - \frac{exp(\beta_0+\beta_1 a^*+\beta_2' c)}{1+exp(\beta_0+\beta_1 a^*+\beta_2' c)} + \frac{exp(\beta_0+\beta_1 a^*+\theta_3 a+\beta_2' c+\theta_2)}{1+exp(\beta_0+\beta_1 a^*+\beta_2' c)}}$$

These estimators coincide with the ones obtained when a rare binary outcome is modeled using logistic regression and a binary mediator is modeled using logistic regression (Valeri and VanderWeele, 2013). Note that when a survival outcome is modeled using Cox proportional hazard model the event is assumed to be rare. This assumption is not required if an accelerated failure time model is employed instead.

References

- Hafeman, D. M., VanderWeele, T. J. (2011). Alternative assumptions for the identification of direct and indirect effects. *Epidemiology*, 22:753-764.

- Imai, K., Keele, L., Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, 15:309-334.

- Pearl J. Direct and indirect effects. In Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence, 2001; 411-420, San Francisco, CA, USA, 2001. Morgan Kaufmann Publishers Inc.

- Shpitser, I., VanderWeele, T. J. (2011). A complete graphical criterion for the adjustment formula in mediation analysis. *International Journal of Biostatistics*, 7:1-24.

- Surveillance, Epidemiology, and End Results (SEER) Program Research Data (1973-2010), National Cancer Institute, DCCPS, Surveillance Research Program, Surveillance Systems Branch, released April 2013.

- Valeri, L. and VanderWeele, T.J. (2013). Mediation analysis allowing for exposure-mediator interaction and causal interpretation: SAS and SPSS macros. *Psychological Methods*, 18:137-150.

- VanderWeele, T. J., Vansteelandt S. (2009). Conceptual issues concerning mediation, interventions and composition. *Statistics and Its Interface*, 2:457-468.

- VanderWeele, T.J.(2011). Causal mediation analysis with survival data. *Epidemiology*, 22:582-585.