Appendix 1: Technical details

Proof of Theorem 1

The proof of Theorem 1 will be given in the more general setting of K distinct event types. Define therefore $\epsilon_t \in \{1, ..., K\}$ as the event type of person i and likewise for the counterfactual variables. Assume that the rate for the event type k is well modeled by the Aalen additive hazard model; that is assume that the rate satisfies

$$\begin{aligned}\lim_{dt \to 0} P(T \in]t, t+dt], \epsilon = k \mid T \geq t, X = x, Z = z, M = m)/dt \\ = \lambda_0^k(t) + \lambda_1^k(t)x + \lambda_2^k(t)z + \lambda_3^k(t)m, \qquad (4)\end{aligned}$$

where $\lambda_j^k(t)$ are potentially time dependent coefficient functions and X and Z can take vector values. Note that by a simple conditioning argument the all course rate satisfies

$$\begin{aligned}\lim_{dt \to 0} P(T \in]t, t+dt] \mid T \geq t, X = x, Z = z, M = m)/dt \\ = g_0(t) + g_1(t)x + g_2(t)z + g_3(t)m,\end{aligned}$$

for functions $g_j(t) = \sum_{k=1}^{K} \lambda_j(t)$.

In the case of multiple event types assumptions (A.1) to (A.5) read

(A.1) No unmeasured confounding of the Exposure-Outcome relation; that is $X \perp (T^{x,m}, \epsilon^{x,m})' \mid Z$.

(A.2) No unmeasured confounding of the Mediator-Outcome relation; that is $M \perp (T^{x,m}, \epsilon^{x,m})' \mid X, Z$.

(A.3) No unmeasured confounding of the Exposure-Mediator relation; that is $X \perp M^x \mid Z$.

(A.4) $M^{x^*} \perp (T^{x,m}, \epsilon^{x,m})' \mid Z$

(A.5) $M^X = M$ and $(T^{X,M}, \epsilon^{X,M})' = (T, \epsilon)'$.

Finally with $\gamma^k(t; x, M^x)$ denoting the counterfactual rate of event type k Theorem 1 can be restated as

Theorem 1 (generalized to multiple event types) Under assumptions (A.1)-(A.5) it holds that the total causal effect of changing the exposure from x^* to x measured on the rate difference scale can be expressed as

$$
\begin{aligned}
\underbrace{\gamma^k(t; x, M^x) - \gamma^k(t; x^*, M^{x^*})}_{\text{TE}(t)} &= \gamma^k(t; x, M^x) - \gamma^k(t; x^*, M^x) \\
&\quad + \gamma^k(t; x^*, M^x) - \gamma^k(t; x^*, M^{x^*}) \\
&= \underbrace{\lambda_1^k(t)(x - x^*)}_{\text{DE}(t)} + \underbrace{\lambda_3^k(t)\alpha_1(x - x^*)}_{\text{IE}(t)},
\end{aligned}
$$

with TE, DE, and IE denoting total effect, natural direct effect, and natural indirect effect, respectively.

Note that the theorem decomposes the total effect into the sum of the total direct effect and the pure indirect effect.[14] The decomposition could also have been done as

$$\gamma^k(t; x, M^x) - \gamma^k(t; x, M^{x^*}) + \gamma^k(t; x, M^{x^*}) - \gamma^k(t; x^*, M^{x^*}),$$

corresponding to sum of the pure direct effect and the total indirect effect. In the absence of interactions pure and total effects are equal.

Proof. The main challenge of the proof is to express the causal rate for event type k, which is denoted $\gamma^k(t; x, M^{x^*})$, as a function of the parameters of the estimated models. For ease of exposition we will assume that the baseline covariates in Z are distributed on the finite set $\mathbf{Z}$. However, by replacing summation by integration at suitable instances the proof also covers other types of baseline variables.

Initially rewrite the probability of an event of type k within the time interval $]t, t + dt]$ for the counterfactual variable $(T^{x,M^{x^*}}, \epsilon^{x,M^{x^*}})$ conditional on being at risk at time t as

$$\begin{aligned} &P(T^{x,M^{x^*}} \in dt, \epsilon^{x,M^{x^*}} = k \mid T^{x,M^{x^*}} \geq t) \\ &\quad = \sum_{z \in \mathbf{Z}} P(T^{x,M^{x^*}} \in dt, \epsilon^{x,M^{x^*}} = k \mid Z = z, T^{x,M^{x^*}} \geq t) P(Z = z \mid T^{x,M^{x^*}} \geq t). \end{aligned} \tag{5}$$

The first of the probabilities in the summation can be rewritten as

$$
\begin{aligned}
& P(T^{x,M^{x^*}} \in dt, \epsilon^{x,M^{x^*}} = k \mid Z = z, T^{x,M^{x^*}} \geq t) \\
= \; & \int_{m\in\mathbb{R}} P(T^{x,M^{x^*}} \in dt, \epsilon^{x,M^{x^*}} = k \mid M^{x^*} = m, Z = z, T^{x,M^{x^*}} \geq t) \\
& \quad \times P(M^{x^*} \in dm \mid Z = z, T^{x,M^{x^*}} \geq t) \\
\overset{(a)}{=} \; & \int_{m\in\mathbb{R}} P(T^{x,m} \in dt, \epsilon^{x,m} = k \mid Z = z, T^{x,m} \geq t) \\
& \quad \times P(M^{x^*} \in dm \mid Z = z, T^{x,M^{x^*}} \geq t) \\
\overset{(b)}{=} \; & \int_{m\in\mathbb{R}} P(T^{x,m} \in dt, \epsilon^{x,m} = k \mid Z = z, X = x, T^{x,m} \geq t) \\
& \quad \times \frac{P(M^{x^*} \in dm, T^{x,M^{x^*}} \geq t \mid Z = z)}{P(T^{x,M^{x^*}} \geq t \mid Z = z)} \\
= \; & \int_{m\in\mathbb{R}} P(T^{x,m} \in dt, \epsilon^{x,m} = k \mid Z = z, X = x, T^{x,m} \geq t) \\
& \quad \times \frac{P(T^{x,M^{x^*}} \geq t \mid M^{x^*} \in dm, Z = z) P(M^{x^*} \in dm \mid Z = z)}{P(T^{x,M^{x^*}} \geq t \mid Z = z)} \\
\overset{(c)}{=} \; & \int_{m\in\mathbb{R}} P(T^{x,m} \in dt, \epsilon^{x,m} = k \mid Z = z, M = m, X = x, T^{x,m} \geq t) \\
& \quad \times \frac{P(T^{x,m} \geq t \mid Z = z) P(M^{x^*} \in dm \mid Z = z)}{P(T^{x,M^{x^*}} \geq t \mid Z = z)}, \\
\overset{(d)}{=} \; & \int_{m\in\mathbb{R}} P(T \in dt, \epsilon = k \mid Z = z, M = m, X = x, T \geq t) \\
& \quad \times \frac{P(T^{x,m} \geq t \mid Z = z) P(M \in dm \mid Z = z, X = x^*)}{\int_{\tilde{m}\in\mathbb{R}} P(T^{x,\tilde{m}} \geq t \mid Z = z) P(M \in d\tilde{m} \mid Z = z, X = x^*)},
\end{aligned}
\tag{6}
$$

where equality (a) is due to assumption (A.4), equality (b) is due to assumption (A.1), equality (c) is by assumption (A.2) and (A.4), and equality (d) is due to assumption (A.3) and (A.5). Similar considerations combined with the expression for the survival function for the Aalen additive model, see Martinussen and Scheike (2006)[17] p. 146, yield that

$$
\begin{aligned}
P(T^{x,m} \geq t \mid Z = z) &= P(T \geq t \mid Z = z, X = x, M = m) \\
&= \exp\{-G_0(t) - G_1(t)x - G_2(t)z - G_3(t)m\},
\end{aligned}
$$

where $G_j(t) = \int_0^t g_j(s)ds$. Hence by (4) and the bounded convergence theorem it holds that

$$
\begin{aligned}
&\lim_{dt \to 0} P(T^{x,M^{x^*}} \in dt, \epsilon^{x,M^{x^*}} = k \mid T^{x,M^{x^*}} \geq t)/dt \\
&= \sum_{z \in \mathbf{Z}} p_z \big\{ E[\exp\{-G_0(t) - G_1(t)x - G_2(t)z - G_3(t)M\} \mid X = x^*, Z = z] \big\}^{-1} \\
&\qquad \times E\Big[(\lambda_0^k(t) + \lambda_1^k(t)x + \lambda_2^k(t)z + \lambda_3^k(t)M) \\
&\qquad\qquad \exp\{-G_0(t) - G_1(t)x - G_2(t)z - G_3(t)M\} \mid X = x^*, Z = z \Big] \\
&= \lambda_0^k(t) + \lambda_1^k(t)x \\
&\qquad + \sum_{z \in \mathbf{Z}} \big\{ \lambda_2^k(t) z p_z + \frac{E[\lambda_3^k(t) M \exp\{-G_3(t)M\} \mid X = x^*, Z = z]}{E[\exp\{-G_3(t)M\} \mid X = x^*, Z = z]} p_z \big\} \qquad (7)
\end{aligned}
$$

where $p_z = P(Z = z \mid T^{x,M^{x^*}} \geq t)$. For a random variable $U \sim N(\mu, \omega^2)$ it holds by properties of the characteristic function of a normal random variable[23] that

$$
\frac{E[U \exp\{aU\}]}{E[\exp\{aU\}]} = \mu + a\omega^2.
$$

Since M is conditionally normal this results implies that (7) can be rewritten as

$$\begin{aligned}
&\lim_{dt\to 0} P(T^{x,M^{x^*}} \in dt, \epsilon^{x,M^{x^*}} = k \mid T^{x,M^{x^*}} \geq t)/dt \\
&= \lambda_0^k(t) + \lambda_1^k(t)x + \sum_{z\in\mathbf{Z}} \left\{\lambda_2^k(t)zp_z + \lambda_3^k(t)(\alpha_0 + \alpha_1 x^* + \alpha_2 z - G_3^k(t)\sigma^2)p_z\right\} \\
&= \lambda_0^k(t) + \lambda_1^k(t)x + \lambda_3^k(t)(\alpha_0 + \alpha_1 x^* - G_3^k(t)\sigma^2) + \sum_{z\in\mathbf{Z}} \left\{\lambda_2^k(t)zp_z + \lambda_3^k(t)\alpha_2 z\right\}.
\end{aligned} \tag{8}$$

In summary it has been established that the counterfactual rate for event type k can be expressed as

$$\gamma^k(t; x, M^{x^*}) = \lambda_0^k(t)+\lambda_1^k(t)x+\lambda_3^k(t)(\alpha_0+\alpha_1 x^*-G_3^k(t)\sigma^2)+\sum_{z\in\mathbf{Z}} \left\{\lambda_2^k(t)+\lambda_3^k(t)\alpha_2\right\}zp_z. \tag{9}$$

Hence on the rate difference scale the total causal effect of changing the exposure from x^* to x can be decomposed as

$$\begin{aligned}
\gamma^k(t; x, M^x) - \gamma^k(t; x^*, M^{x^*}) &= \gamma^k(t; x, M^x) - \gamma^k(t; x^*, M^x) \\
&\quad +(\gamma^k(t; x^*, M^x) - \gamma^k(t; x^*, M^{x^*})) \\
&= \lambda_1^k(t)(x - x^*) + \lambda_3^k(t)\alpha_1(x - x^*).
\end{aligned}$$

□

Cumulative measures of mediation

Following the tradition for Aalen models we also report a cumulative measures of mediation. From Theorem 1 the following corollary can easily be derived

Corollary 1. Under assumptions (A.1)-(A.5) it holds that the counterfactual rate difference at time t can be expressed as

$$\begin{aligned}
&\int_0^t \gamma(s; x, M^x) - \gamma(s; x^*, M^{x^*}) ds \\
&\qquad = \quad (x - x^*) \int_0^t \lambda_1(s) ds + \alpha_1 (x - x^*) \int_0^t \lambda_3(s) ds,
\end{aligned}$$

where the first term measures the natural cumulative direct effect and the second term measures the natural cumulative indirect effect.

Confidence intervals with time dependent effects

Let θ_1 denote the collection of parameters from the ordinary regression of the mediator on the exposure and baseline covariates and $\theta_2(t)$ denotes the collection of parameter functions of the Aalen model. In order to discuss estimation uncertainty define the cumulative coefficient functions as $\Theta_2(t) = \int_0^t \theta_2(s) ds$. Under mild regularity conditions, see e.g. Condition 5.1 of Martinussen and Scheike (2006),[17] it holds that $\hat{\theta}_1$ is asymptotically normally distributed and for any t it holds that $\hat{\Theta}_2(t)$ is also asymptotically normally distributed. In addition the two vectors of estimates are uncorrelated. The covariance matrices of the two vectors of estimates are available as output from standard statistical software. Hence the cumulative direct effect at time t is asymptotically normal whilst the cumulative indirect effect at time t is asymptotically distributed as the product of two uncorrelated normal random variables. Confidence bands and tests involving cumulative indirect effect can either be computed using the delta rule or better yet by simulating a large number of realizations of the two uncorrelated random variables.

eAPPENDIX REFERENCE

23. Patel JK, Campbell B. *Handbook of the Normal Distribution*. New York: Marcel Dekker; 1996.

Appendix 2: Implementation Guidelines

Through the package "timereg" the statistical software package R has efficient tools for estimating Aalen models. As R is less used among epidemiologists this appendix provides a detailed explanation of how the results presented in the paper were computed.

The data frame `ses_dat_m` contains the following variables

`T2`	Event or censor time	`event`	LSA event yes/no
`SES`	SES in 5 levels	`logphys`	Logarithm of the physical index
`coha`	Cohabitation yes/no	`child`	Children yes/no
`age`	Age at baseline		

for all male persons in the sample ($N = 3458$).

The regression results presented in Table 3 are obtained by the following commands

```
ols_m <- glm(logphys ~ age + COHA + child + factor(SES),
             data=ses_dat_m)
summary(ols_m)
```

The first step in computing the numbers of Table 4 is to fit an Aalen model, which allows all coefficients to be time dependent. This can be done by the commands

```
library(timereg)
```

```
aalen_m1 <- aalen(Surv(T2,event) ~ age + COHA + child + logphys
                 + factor(SES), data = ses_dat_m, robust=T )
```

The command `summary(aalen_m1)` tests the four hypothesis that none of the coefficients need to be time dependent. As all of the p-values are above 80% the model is simplified such that the baseline risk is the only time dependent component. Interaction terms can be included by arguments like `age*logphys`. However, none of the interaction terms are significant after accounting for the number of tests conducted and are therefore not included in the final model. The simplified model can be estimated by the following commands

```
aalen_m2 <- aalen(Surv(T2,event) ~ const(age) + const(COHA)
           + const(child) + const(logphys)
           + const(factor(SES)), data = ses_dat_m, robust=T )
```

Finally the content of Table 4 is obtained by the command `summary(aalen_m2)`. The estimates of the direct effect (DE), indirect effect (IE), total effect (TE), and the ratio of indirect effect to total effect can be obtained directly from Table 4 by employing Theorem 1. However, only the confidence interval for DE is immediately available. The other three are non-linear transformations of the triple $(\hat{\lambda}_1, \hat{\lambda}_3, \hat{\alpha}_3)$. This triplet is asymptotically normal with a mean equal to the corresponding true values of the parameters and a covariance matrix with the structure

$$\Omega = \begin{pmatrix} \omega_{11} & \omega_{12} & 0 \\ \omega_{12} & \omega_{22} & 0 \\ 0 & 0 & \omega_{33} \end{pmatrix}.$$

The diagonal components are available from Table 4 and ω_{12} (the covariance between $\hat{\lambda}_1$ and $\hat{\lambda}_3$) can be printed by the command `aalen_m$var.gamma`. This assymptotic result implies that confidence intervals for IE, TE, and IE/TE can be computed by simulation. The function below performs such a simulation.

```
CI_comp <- function(mean_lambda1 ,mean_lambda3,covar11,covar12,
                    covar22, mean_alpha , var_alpha, G=10^4)
{
    require(mvtnorm)
    Omega <- matrix(c(covar11,covar12,covar12,covar22),nrow=2)
    IE <- rep(0,G);TE <- rep(0,G);Q <- rep(0,G)
    for(i in 1:G)
    {
       lambda <- rmvnorm(1, mean = c(mean_lambda1,mean_lambda2),
                         sigma = Omega)
       alpha <- rnorm(1,mean=mean_alpha, sd=sqrt(var_alpha))
       IE[i] <- lambda[2] * alpha
```

```
        TE[i] <- IE[i] + lambda[1]
        Q[i] <- IE[i]/TE[i]
    }
    print("IE:")
    print(mean(IE))
    print(quantile(IE,c(0.025, 0.975)))
    print("TE:")
    print(mean(TE))
    print(quantile(TE,c(0.025, 0.975)))
    print("Q:")
    print(mean(Q))
    print(quantile(Q,c(0.025, 0.975)))
}
```

When the function has been loaded in R the confidence intervals can be computed by a command like

```
f_out(mean_lambda1=0.000561 ,mean_lambda2=0.000234,
      covar11=0.000197^2, covar12=-1.12*10^(-9), covar22=0.000054^2,
      mean_alpha=0.67, var_alpha=0.066^2)
```

Appendix 3: Bias and coverage rates

To asses bias and coverage rate of the stated confidence intervals we conduct a simulation study where we take the estimated parameters (including the unreported baseline hazard) as "true values" and simulate new data sets. The censoring mechanism is taken as in the real data set. On each simulated data set we compute all of the parameters discussed in the illustration and can thereby asses bias and coverage rates. The result of 20,000 Monte Carlo replications is presented in Table 6 below. For the total, direct and indirect effects the bias is consistently less than 2.5%. For the ratio between the total and indirect effects we compare the median of the simulated ratios with the true ratio as the bias cannot be defined for a ratio. We find that the discrepancy is less than 1 percentage point. For the total, direct and indirect effects the coverage rates for the confidence intervals are found to be very close to the nominal level while the coverage rates for the ratios of indirect to total effect are slightly oversized.

Table 6: Relative bias and coverage rates (in parentheses) both in percentage points based on 20,000 Monte Carlo simulations.

	DE	IE	TE	IE/TE*
Men I→II	-2.29 (94.9)	-2.55 (94.7)	-2.34 (94.8)	-0.11 (97.3)
Men I→III	-2.26 (94.8)	-2.50 (94.8)	-2.35 (94.6)	-0.13 (97.4)
Men I→IV	-1.62 (94.8)	-2.46 (94.8)	-2.01 (94.8)	-0.03 (95.3)
Men I→V	-1.69 (94.9)	-2.46 (94.8)	-2.01 (94.8)	-0.15 (95.0)
Women I→II	-2.19 (95.1)	-2.10 (94.5)	-2.17 (95.0)	-0.09 (97.0)
Women I→III	-1.96 (95.0)	-2.10 (94.5)	-2.01 (95.0)	-0.63 (97.1)
Women I→IV	-2.16 (95.1)	-2.07 (94.7)	-2.10 (95.2)	-0.90 (97.5)
Women I→V	-2.20 (94.6)	-2.07 (94.8)	-2.15 (94.7)	0.00 (96.2)

*the median of differences between simulated and true values.