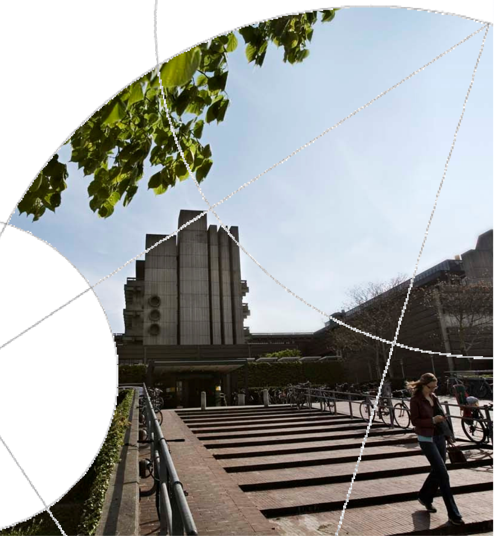
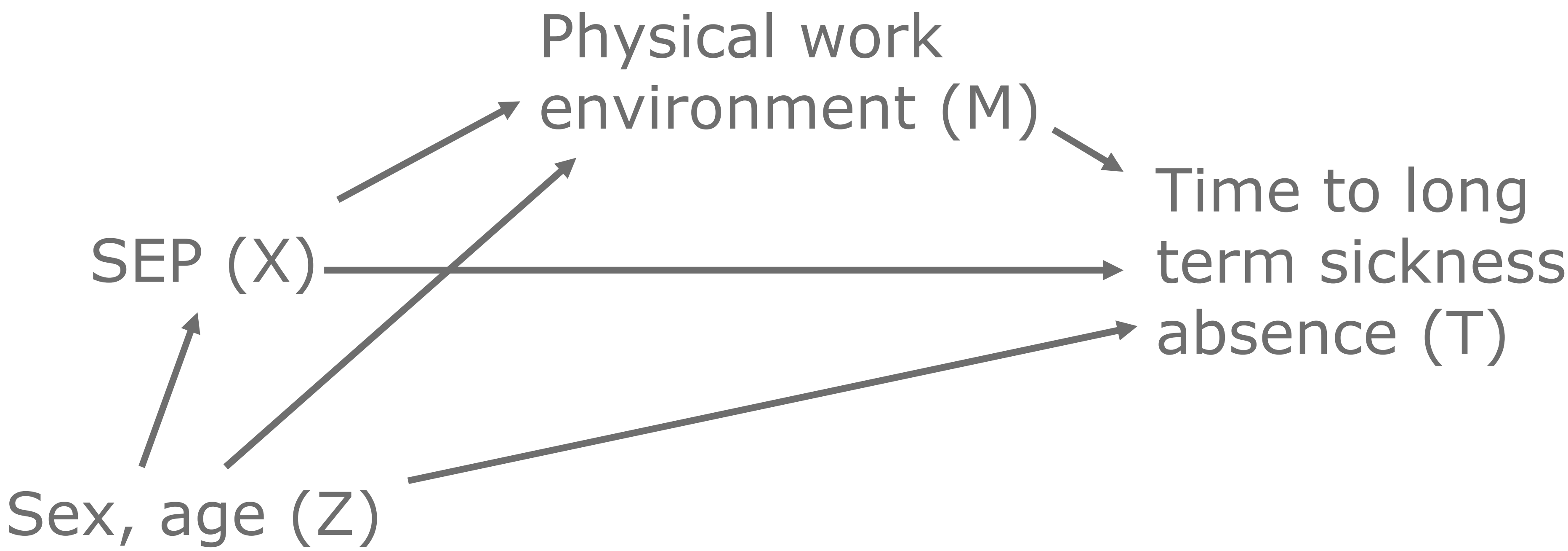
Mediation in a survival context

Theis Lange



An example: The current strategy

Christensen, Labriola, Lund & Kivimäki (2008) analyses the following problem:



An example: The current strategy (II)

Christensen et al. (08) uses a number of Cox models to asses how much of the effect of SEP is mediated through the physical work environment.

	Men		
Model	**RR**	**(95% CI)**	**Change**
Adjusted for age and family status			
I Executive managers/academics	1.00	–	
II Middle managers/3–4 years further education	2.04	(0.93 to 4.47)	
III Other white-collar workers	2.67	(1.39 to 5.13)	
IV Skilled blue-collar workers	3.56	(1.85 to 6.85)	
V Semiskilled and unskilled workers	4.22	(2.23 to 7.97)	
SEP as a linear term	1.37	(1.21 to 1.55)	
Adjusted for age, family status, health behaviour and physical work environment			
I Executive managers/academics	1.00	–	
II Middle managers/3–4 years further education	1.59	(0.72 to 3.53)	0.45 (22%)
III Other white-collar workers	1.85	(0.94 to 3.64)	0.82 (31%)
IV Skilled blue-collar workers	1.58	(0.75 to 3.33)	1.98 (56%)
V Semiskilled and unskilled workers	1.83	(0.90 to 3.73)	2.39 (57%)
SEP as a linear term	1.10	(0.95 to 1.28)	0.27 (20%)

An example: The current strategy (II)

Christensen et al. (08) uses a number of Cox models to asses how much of the effect of SEP is mediated through the physical work environment.

Problems with this approach:

- What is the (causal) interpretation of the drop in HRs?
- What are the confidence bands for the stated drop of e.g. the 57% for SEP group V?
- The different Cox models are not mathematically consistent.

How is the problem tackled with other outcomes?

Define counterfactual variables

T(m,x) Outcome of interest (binary for now) when exposure is set to x and mediated set to m

M(x) Mediator when exposure is set to x.

T(M(x),x*) Outcome of interest when exposure is set to x* and mediated set to the value it would take if exposure had been x.

And effects:

Natural indirect effect is defined as: E[T(M(x),x) - T(M(x*),x)]

Natural direct effect is defined as: E[T(M(x),x) - T(M(x),x*)]

How to do this in a survival context…

The most important step is to replace the Cox model by the Aalen additive hazard model; that is assume that the rate of the observed outcome satisfies the following:

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

where $\lambda_j^k(t)$ are potentially time dependent coefficient functions.

The Aalen additive hazard model is just as flexible as the Cox model, but instead of estimating HR we are estimating hazard differences.

Currently not implemented in SAS, but R has well developed tools.

How to do this in a survival context… (II)

Mediator can be of any type, but the formulas simplifies somewhat when assuming conditional normality of the mediator; that is assume

$$M = \alpha_0 + \alpha_1 x + \alpha_2 z + e,$$

where e is a normally distributed error term with variance σ^2.

Estimation:

Under standard assumptions of “no unmeasured confounding of any relations” the models for both T and M can be estimated.

How to do this in a survival context… (III)

Denote the causal rate when the exposure is set to x and the mediator to m by

$$\gamma^k(t;x,m) = \lim_{dt\to 0} P(T^{x,m} \in]t, t+dt], \epsilon^{x,m} = k \mid T^{x,m} \geq t)/dt.$$

Under two more technical assumptions on the counterfactuals (consistency and the Pearl (01) identifiability assumption) the total effect of changing the exposure from x* to x can be written as

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

DE(t) = $\gamma^k(t;x,M^x) - \gamma^k(t;x^*,M^x)$

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

How to do this in a survival context… (IV)

Notes:

- The formulas closely mirrors the similar formulas for a normal outcome.
- The simplicity hinges on both normality of the mediator and absence of interaction. However, both can easily be incorporated at the price of a bit more cumbersome calculations.
- As everything is estimated within an internally consistent mathematical setup derived quantities e.g DE/TE along with there confidence interval can be computed straightforward.
- A similar analysis could be conducted on the cumulative incidence functions if these are better suited for the problem at hand.

Back to our example (Christensen et al. (08))

- As expected non of the effects in this problem are time dependent.

- The table below illustrates how the technique can be used to the analysis in Christensen et al. (08).

	DE ($*10^{-4}$) (95% CI)	IE ($*10^{-4}$) (95% CI)	TE ($*10^{-4}$) (95% CI)	IE/TE (95% CI)
I -> II	…	…	…	…
I -> III	…	…	…	…
I -> IV	…	…	…	…
I -> V	1.9 (1.7; 2.1)	2.0 (1.8; 2.2)	3.9 (3.6; 4.2)	51% (45%; 56%)