simulation summary binary

Joakim Wallmark

Simulation study

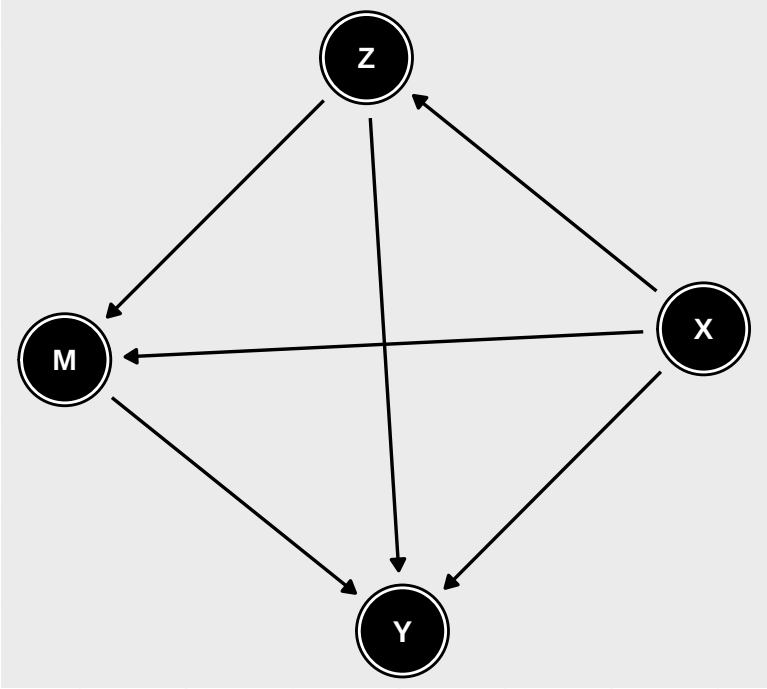
Simulation scenarios

Simulations of bias were done with model misspecification and confounding in separate comparable scenarios. 10000 iterations (1/10th of the linear case due to computation time) with a sample size of 2000 were used for each scenario.

Table 1: Variables used for simulations

variable	type	true model
X(additional covariate)	continuous	$X \sim gamma(8, 4.5)$
Z(exposure)	binary	$Z = I(Z* > 0)$ where $Z* \sim U_0 + U_1 X + N(0,1)$
M(mediator)	binary	$M = I(M* > 0)$ where $M* \sim \beta_0 + \beta_1 Z + \beta_2 X + N(0,1)$
Y(outcome)	binary	$Y = I(Y* > 0)$ where $Y* \sim \theta_0 + \theta_1 Z + \theta_2 M + \theta_3 ZM + \theta_4 X + N(0,1)$

DAG showing data generating process



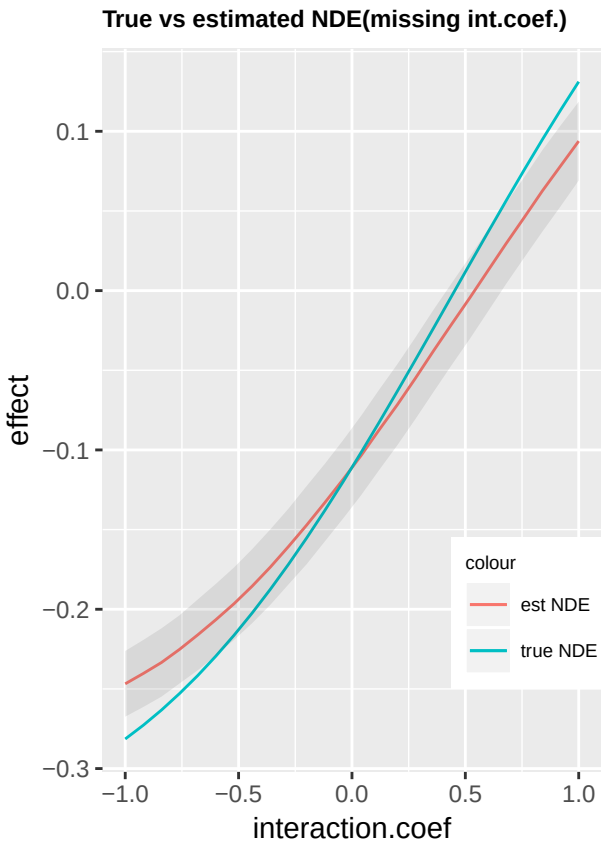
Simulation scenarios (model misspecification)

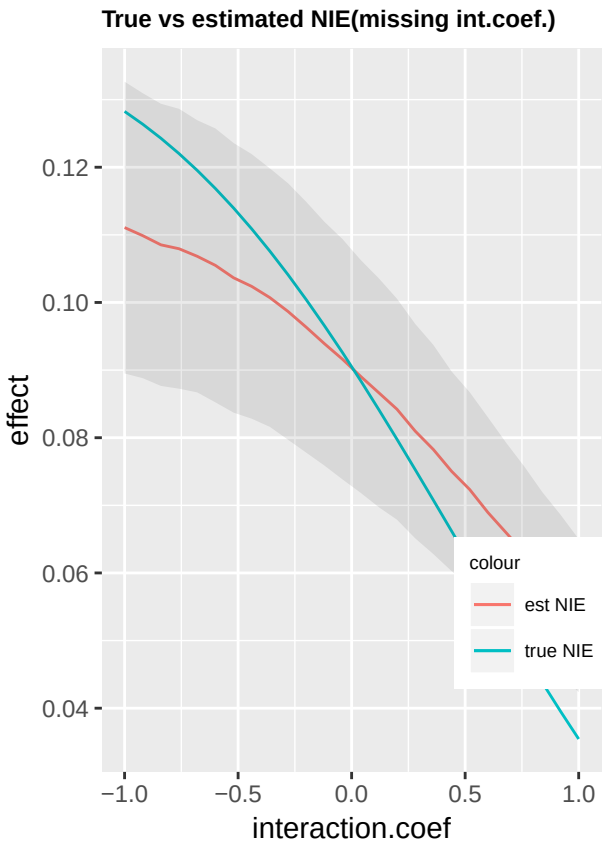
To get realistic parameter values, models were estimated using data from riksstroke from n=60353 patients. Parameter values used for simulations:

covariate.parameters = list(c(104, 8, 4.5)), true.exposure.coefs = c(I = -3.416096, X = 0.036231), true.mediator.coefs = c(I = 1, Z = -0.4, X = -0.008, ZX = 0), true.outcome.coefs = c(I = 4, Z = -0.4, M = -2, ZM = corr.coef[i], X = -0.03, ZX = 0, MX = 0, ZMX = 0),

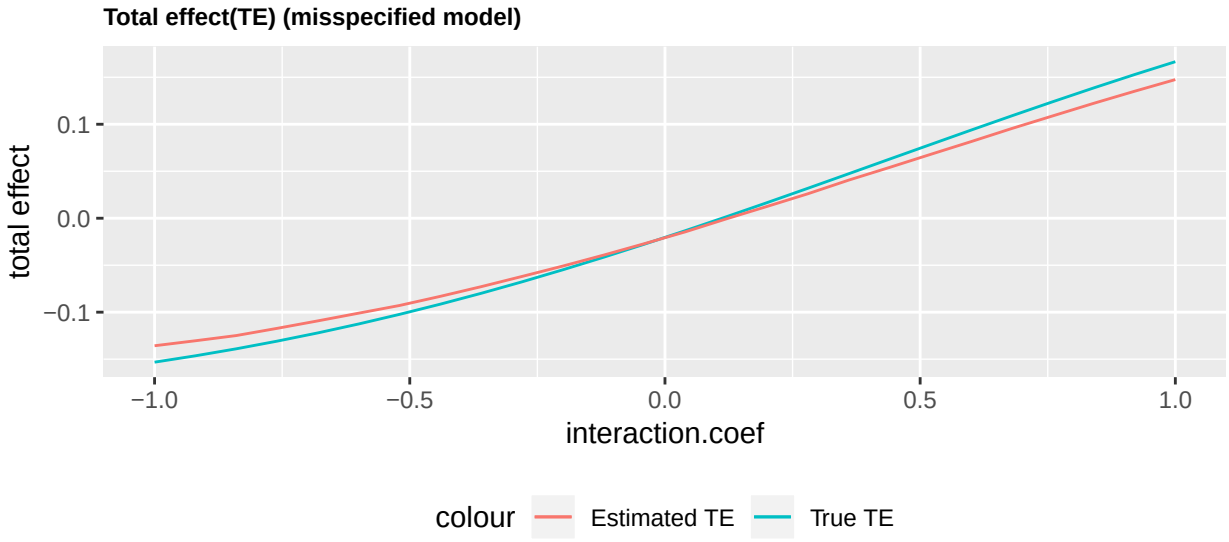
Estimated mediator model was set to the correct one. Estimated outcome model was misspecified without ZM interaction: $Y* \sim Z + M + X$.

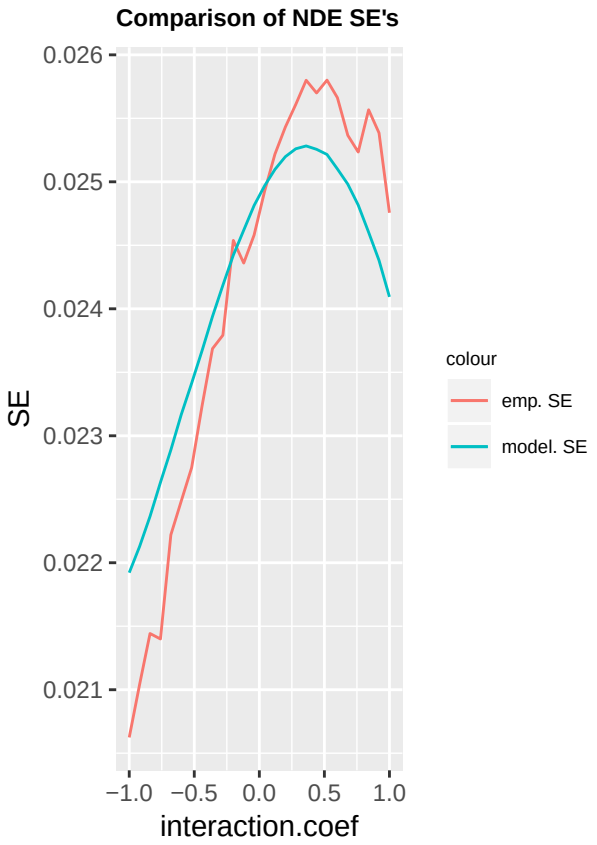
Results

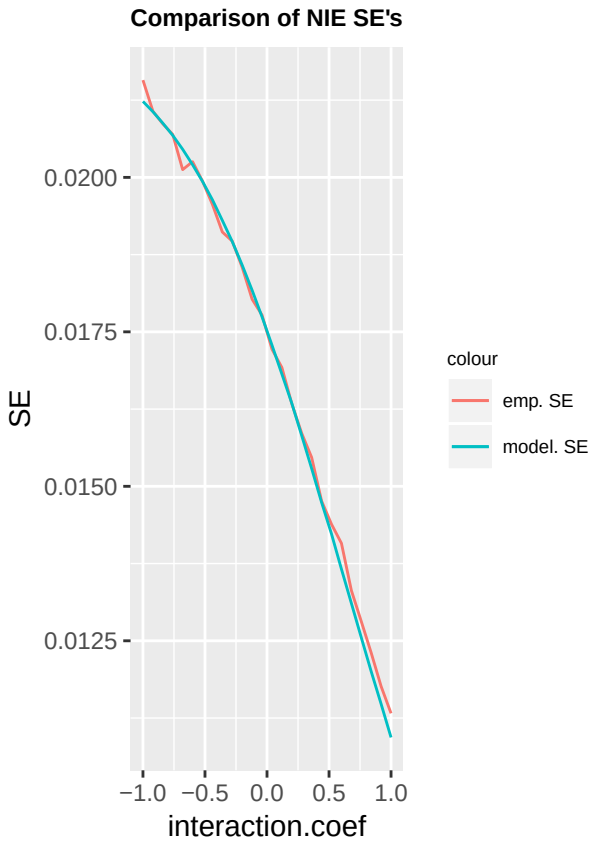

True vs estimated NDE(missing int.coef.)
effect
interaction.coef
0.1
0.0
−0.1
−0.2
−0.3
−1.0
−0.5
0.0
0.5
1.0
colour
est NDE
true NDE

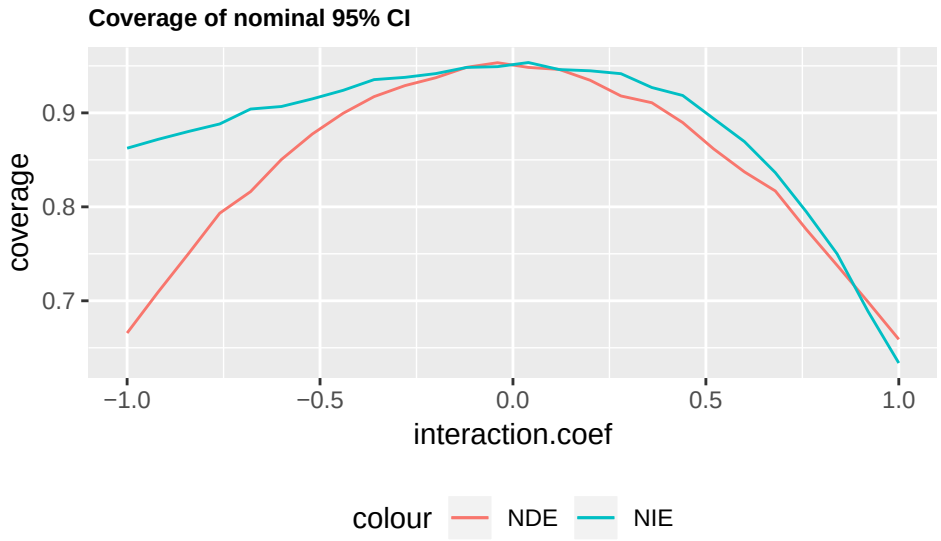

True vs estimated NIE(missing int.coef.)
effect
interaction.coef
0.12
0.10
0.08
0.06
0.04
−1.0
−0.5
0.0
0.5
1.0
colour
est NIE
true NIE

Total effect is biased. This differs from the situation where we used linear models:








Coverage of nominal 95% CI
coverage
0.9
0.8
0.7
−1.0
−0.5
0.0
0.5
1.0
interaction.coef
colour
NDE
NIE