First steps into CMAverse

Cécile Proust-Lima (Inserm, Univ. Bordeaux)

October 15, 2025

Contents

This is an introduction to the R package CMAverse. See the website with various vignettes: https://bs1125.github.io/CMAverse/

1 Installation of the R package

It is for now on github only. For the installation, install remotes R package if necessary and run:

```r
library(remotes)
install_github("BS1125/CMAverse")
```

```r
library(CMAverse)
```

2 Working dataset

We consider the setting of a binary treatment, 2 confounders at baseline, a first intermediate binary variable, a second intermediate continuous mediator, and a continuous outcome.

Here is an example of such setting:

```
 - Exposure = smoking status
 - Confounders = age, sex (male, ref female)
 - First mediator/confounder = fat indicator (e.g., BMI, body fat percentage, trunk fat percentage, wais
 - Second mediator = Inflammation (measured by C-reactive protein (CRP))
```

```
 - Outcome = cardiovascular health, Framingham Risk Score (FRS) (could have been blood pressure too).
```

There are some research questions about the role of inflammation in the relationship between smoking and cardiovascular health (Chen et al, 2025) or the role of BMI in the association between smoking and cardiovascular health (Lin et al., 2017).

CAUTION: The working data are simulated, and the effects found do not reflect AT ALL actual epidemiological results! I completely made up the models and data.

```r
set.seed(1)
n <- 100
Age <- rnorm(n, mean = 5, sd = 0.5)
Sex <- rbinom(n, 1, 0.5)
pa <- exp(0.2 - 0.1*Age + 0.1*Sex)/(1 + exp(0.2 - 0.1*Age + 0.1*Sex))
Smoking <- rbinom(n, 1, pa)
pm <- exp(0.7 - 0.5*Smoking + 0.2*Age + 0.5*Sex)/ (0.7 + exp(1 - 0.5*Smoking + 0.2*Age + 0.5*Sex))
Fat <- rbinom(n, 1, pm)
CRPtrue <- rnorm(n, 2 + 1.1*Smoking +1.5*Fat + 0.5*Age + 1*Sex, 1)
CRP <- rnorm(n,CRPtrue,1)
FRS <- rnorm(n, mean = 0.5 + 0.8*Smoking + 0.5*Fat + 0.6*CRPtrue + 0.3*Smoking*Fat + 0.5*Smoking*CRPtru
dataSim <- data.frame(ID=1:n,Age, Sex, Smoking,Fat,CRP,FRS)
#save(dataSim, file="dataSim_for_Session3.Rdata")
```

Here is a summary of the data:

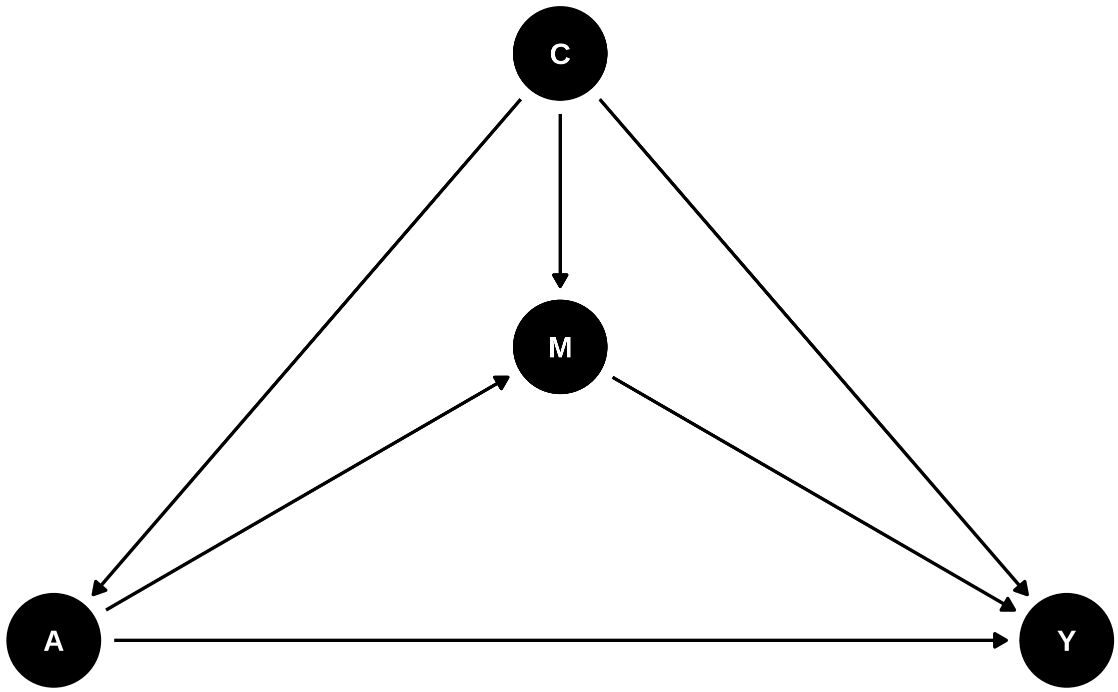
```r
summary(dataSim)
```

```
##        ID              Age             Sex           Smoking          Fat
##  Min.   :  1.00   Min.   :3.893   Min.   :0.00   Min.   :0.00   Min.   :0.00
##  1st Qu.: 25.75   1st Qu.:4.753   1st Qu.:0.00   1st Qu.:0.00   1st Qu.:0.00
##  Median : 50.50   Median :5.057   Median :0.00   Median :0.00   Median :1.00
##  Mean   : 50.50   Mean   :5.054   Mean   :0.38   Mean   :0.39   Mean   :0.65
##  3rd Qu.: 75.25   3rd Qu.:5.346   3rd Qu.:1.00   3rd Qu.:1.00   3rd Qu.:1.00
##  Max.   :100.00   Max.   :6.201   Max.   :1.00   Max.   :1.00   Max.   :1.00
##       CRP              FRS
##  Min.   : 1.997   Min.   : 1.516
##  1st Qu.: 5.227   1st Qu.: 4.564
##  Median : 6.603   Median : 5.996
##  Mean   : 6.347   Mean   : 7.008
##  3rd Qu.: 7.415   3rd Qu.: 9.403
##  Max.   :10.183   Max.   :15.720
```

3 Study of Smoking - CRP - FRS (neglecting Fat)

3.1 The DAG

```r
cmdag(outcome = "FRS", exposure = "Smoking", mediator = c("CRP"),
      basec = c("Age", "Sex"))
```



A (exposure): Smoking
M (mediator): CRP
Y (outcome): FRS
C (confounders not affected by the exposure): Age, Sex

We could use classical regression tools to compute the causal effects as shown in the slides.

3.2 Step by step using glm and posterior computations

Estimation of the two regression models

```
estRegY <- glm(formula = FRS ~ Smoking + CRP + Smoking * CRP + Age + Sex, family = gaussian(), data = da
estRegM <- glm(formula = CRP ~ Smoking + Age + Sex, family = gaussian(), data = dataSim)
summary(estRegY)
```

```
##
## Call:
## glm(formula = FRS ~ Smoking + CRP + Smoking * CRP + Age + Sex,
##     family = gaussian(), data = dataSim)
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.486723   1.502944   1.655  0.10135
## Smoking     1.760841   1.244489   1.415  0.16040
## CRP         0.308937   0.100363   3.078  0.00273 **
## Age         0.004353   0.295562   0.015  0.98828
## Sex         1.397417   0.269938   5.177 1.28e-06 ***
## Smoking:CRP 0.471868   0.179238   2.633  0.00990 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1.671607)
##
```

```
##     Null deviance: 1038.31  on 99  degrees of freedom
## Residual deviance:  157.13  on 94  degrees of freedom
## AIC: 342.98
##
## Number of Fisher Scoring iterations: 2
```

```
summary(estRegM)
```

```
##
## Call:
## glm(formula = CRP ~ Smoking + Age + Sex, family = gaussian(),
##     data = dataSim)
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.4716     1.7944   1.377   0.1716
## Smoking       1.3549     0.3225   4.201 5.95e-05 ***
## Age           0.6270     0.3516   1.783   0.0777 .
## Sex           0.4689     0.3240   1.447   0.1510
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 2.466087)
##
##     Null deviance: 294.04  on 99  degrees of freedom
## Residual deviance: 236.74  on 96  degrees of freedom
## AIC: 379.97
##
## Number of Fisher Scoring iterations: 2
```

Posterior computation of the regression-based natural effects:

```
theta <- coef(estRegY)
beta <- coef(estRegM)
Eage <- mean(dataSim$Age)
Esex <- mean(dataSim$Sex)
a <- 1
astar <- 0
m <- 5
CDE <- (theta[2]+ theta[6]*m)*(a-astar)
NDE <- theta[2] + theta[6] * (beta[1] + beta[2] * astar + beta[3]*Eage + beta[4]*Esex)*(a - astar)
NIE <- theta[3]*beta[2] + theta[6]*beta[2]*a
TE <- NDE + NIE
data.frame(CDE, NDE, NIE, TE)
```

```
##              CDE      NDE      NIE      TE
## Smoking 4.120182 4.506677 1.057903 5.56458
```

3.3 Regression-based estimation using cmest function

We specify the model with cmest function using model="rb" (regression based). We have to specify the values for a and astar, and for the mediator (for CDE). To obtain the closed form estimates, we need to add estimation = "paramfunc" and inference = "delta" to obtain variance estimates using the Delta-Method.

```
estRB <- cmest(data = dataSim, model = "rb",
               outcome = "FRS",
               exposure = "Smoking",
```

```
               mediator = c("CRP"),
               basec = c("Age", "Sex"), EMint = TRUE,
               mreg = list("linear"), yreg = "linear",
               astar = 0, a = 1, mval = list(5),
               estimation = "paramfunc", inference = "delta")
summary(estRB)
```

```
## Causal Mediation Analysis
##
## # Outcome regression:
##
## Call:
## glm(formula = FRS ~ Smoking + CRP + Smoking * CRP + Age + Sex,
##     family = gaussian(), data = getCall(x$reg.output$yreg)$data,
##     weights = getCall(x$reg.output$yreg)$weights)
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.486723   1.502944   1.655  0.10135
## Smoking     1.760841   1.244489   1.415  0.16040
## CRP         0.308937   0.100363   3.078  0.00273 **
## Age         0.004353   0.295562   0.015  0.98828
## Sex         1.397417   0.269938   5.177 1.28e-06 ***
## Smoking:CRP 0.471868   0.179238   2.633  0.00990 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1.671607)
##
##     Null deviance: 1038.31  on 99  degrees of freedom
## Residual deviance:  157.13  on 94  degrees of freedom
## AIC: 342.98
##
## Number of Fisher Scoring iterations: 2
##
##
## # Mediator regressions:
##
## Call:
## glm(formula = CRP ~ Smoking + Age + Sex, family = gaussian(),
##     data = getCall(x$reg.output$mreg[[1L]])$data, weights = getCall(x$reg.output$mreg[[1L]])$weights
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.4716     1.7944   1.377   0.1716
## Smoking       1.3549     0.3225   4.201 5.95e-05 ***
## Age           0.6270     0.3516   1.783   0.0777 .
## Sex           0.4689     0.3240   1.447   0.1510
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 2.466087)
##
##     Null deviance: 294.04  on 99  degrees of freedom
```

```
## Residual deviance: 236.74  on 96  degrees of freedom
## AIC: 379.97
##
## Number of Fisher Scoring iterations: 2
##
##
## # Effect decomposition on the mean difference scale via the regression-based approach
##
## Closed-form parameter function estimation with
##  delta method standard errors, confidence intervals and p-values
##
##               Estimate Std.error  95% CIL 95% CIU    P.val
## cde           4.120182  0.426905 3.283464   4.957  < 2e-16 ***
## pnde          4.506677  0.347172 3.826233   5.187  < 2e-16 ***
## tnde          5.146006  0.321318 4.516235   5.776  < 2e-16 ***
## pnie          0.418574  0.168574 0.088176   0.749 0.013027 *
## tnie          1.057903  0.323662 0.423537   1.692 0.001081 **
## te            5.564580  0.336199 4.905641   6.224  < 2e-16 ***
## intref        0.386495  0.174831 0.043831   0.729 0.027059 *
## intmed        0.639329  0.286587 0.077629   1.201 0.025692 *
## cde(prop)     0.740430  0.067308 0.608510   0.872  < 2e-16 ***
## intref(prop) 0.069456  0.032104 0.006533   0.132 0.030507 *
## intmed(prop) 0.114893  0.049815 0.017257   0.213 0.021089 *
## pnie(prop)    0.075221  0.029079 0.018228   0.132 0.009687 **
## pm            0.190114  0.054018 0.084241   0.296 0.000432 ***
## int           0.184349  0.072113 0.043009   0.326 0.010577 *
## pe            0.259570  0.067308 0.127649   0.391 0.000115 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (cde: controlled direct effect; pnde: pure natural direct effect; tnde: total natural direct effect;
##
## Relevant variable values:
## $a
## [1] 1
##
## $astar
## [1] 0
##
## $mval
## $mval[[1]]
## [1] 5
##
##
## $basecval
## $basecval[[1]]
## [1] 5.054444
##
## $basecval[[2]]
## [1] 0.38
```

We find exactly the same effects as those computed step by step from glm estimates. Inded, the program performs exactly the same computations as above: two glms and derived closed-form solutions for the causal effects. With the add-in of correct standard errors thnaks to the Delta-Method.

There are many effects computed in the output. Here we focus on CDE, NDE (PNDE / TNDE), NIE (PNDE / TNDE), TE and PM (= TNIE / TE) only.

We can remove the others to not get confused with full=FALSE:

```
estRB2 <- cmest(data = dataSim, model = "rb", full=FALSE,
                outcome = "FRS",
                exposure = "Smoking",
                mediator = c("CRP"),
                basec = c("Age", "Sex"), EMint = TRUE,
                mreg = list("linear"), yreg = "linear",
                astar = 0, a = 1, mval = list(5),
                estimation = "paramfunc", inference = "delta")
summary(estRB2)
```

```
## Causal Mediation Analysis
##
## # Outcome regression:
##
## Call:
## glm(formula = FRS ~ Smoking + CRP + Smoking * CRP + Age + Sex,
##     family = gaussian(), data = getCall(x$reg.output$yreg)$data,
##     weights = getCall(x$reg.output$yreg)$weights)
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.486723   1.502944   1.655  0.10135
## Smoking     1.760841   1.244489   1.415  0.16040
## CRP         0.308937   0.100363   3.078  0.00273 **
## Age         0.004353   0.295562   0.015  0.98828
## Sex         1.397417   0.269938   5.177 1.28e-06 ***
## Smoking:CRP 0.471868   0.179238   2.633  0.00990 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1.671607)
##
##     Null deviance: 1038.31  on 99  degrees of freedom
## Residual deviance:  157.13  on 94  degrees of freedom
## AIC: 342.98
##
## Number of Fisher Scoring iterations: 2
##
##
## # Mediator regressions:
##
## Call:
## glm(formula = CRP ~ Smoking + Age + Sex, family = gaussian(),
##     data = getCall(x$reg.output$mreg[[1L]])$data, weights = getCall(x$reg.output$mreg[[1L]])$weights
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.4716     1.7944   1.377   0.1716
## Smoking       1.3549     0.3225   4.201 5.95e-05 ***
## Age           0.6270     0.3516   1.783   0.0777 .
```

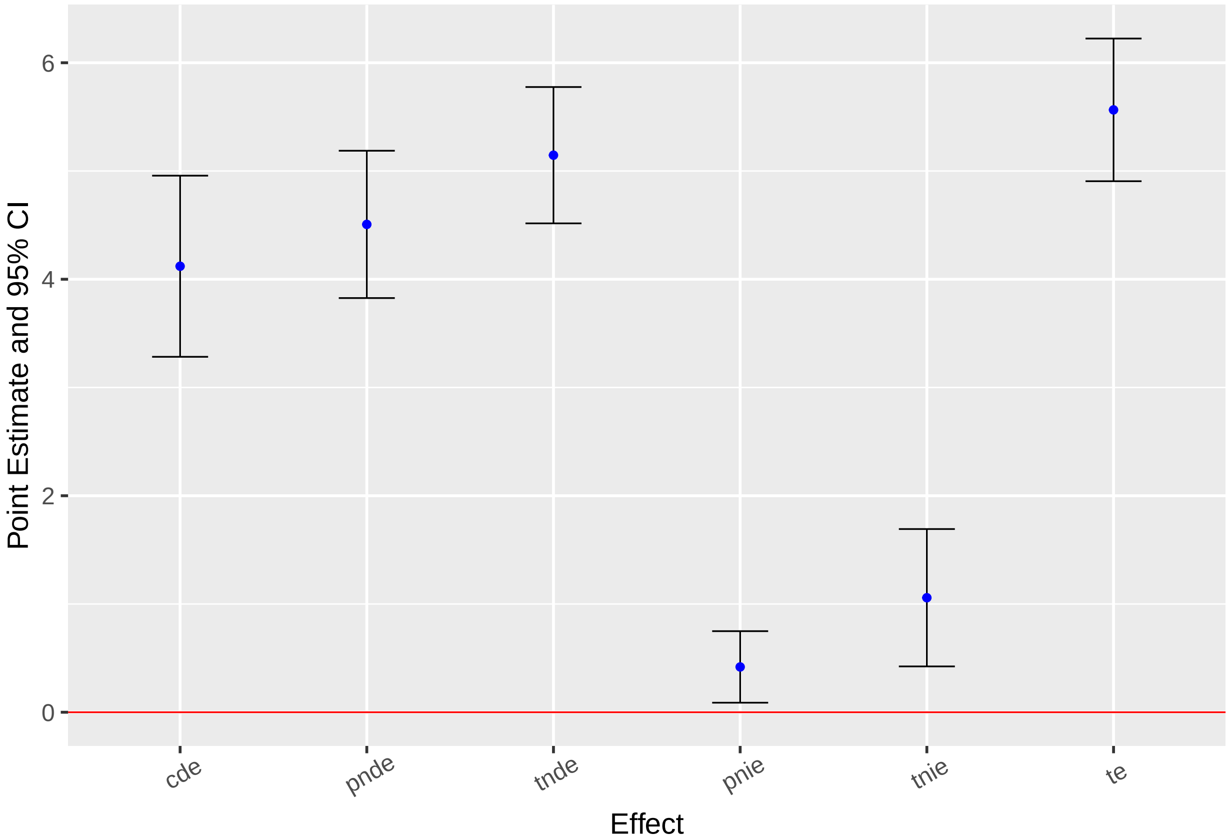
```
## Sex           0.4689     0.3240   1.447   0.1510
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 2.466087)
##
##     Null deviance: 294.04  on 99  degrees of freedom
## Residual deviance: 236.74  on 96  degrees of freedom
## AIC: 379.97
##
## Number of Fisher Scoring iterations: 2
##
##
## # Effect decomposition on the mean difference scale via the regression-based approach
##
## Closed-form parameter function estimation with
##  delta method standard errors, confidence intervals and p-values
##
##      Estimate Std.error 95% CIL 95% CIU   P.val
## cde   4.12018   0.42690 3.28346   4.957 < 2e-16 ***
## pnde  4.50668   0.34717 3.82623   5.187 < 2e-16 ***
## tnde  5.14601   0.32132 4.51623   5.776 < 2e-16 ***
## pnie  0.41857   0.16857 0.08818   0.749 0.01303 *
## tnie  1.05790   0.32366 0.42354   1.692 0.00108 **
## te    5.56458   0.33620 4.90564   6.224 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (cde: controlled direct effect; pnde: pure natural direct effect; tnde: total natural direct effect;
##
## Relevant variable values:
## $a
## [1] 1
##
## $astar
## [1] 0
##
## $mval
## $mval[[1]]
## [1] 5
##
##
## $basecval
## $basecval[[1]]
## [1] 5.054444
##
## $basecval[[2]]
## [1] 0.38
```

We can plot the results:

```
ggcmest(estRB2) + ggplot2::theme(axis.text.x = ggplot2::element_text(angle = 30, vjust = 0.8))
```

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
```

```
## i The deprecated feature was likely used in the CMAverse package.
##   Please report the issue at <https://github.com/BS1125/CMAverse/issues>.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```



3.4 What about other techniques?

We could use alternative techniques to estimate these effects

3.4.1 With natural effect (NE) model

The dataset is expanded using imputation technique for the "missing" counterfactual. Here, no need for a mediator model.

```r
estNE <- cmest(data = dataSim, model = "ne", full=FALSE,
               outcome = "FRS",
               exposure = "Smoking",
               mediator = c("CRP"),
               basec = c("Age", "Sex"), EMint = TRUE,
               yreg = "linear",
               astar = 0, a = 1, mval = list(5))
```

```
##   |                                                                              |
```

```r
summary(estNE)
```

```
## Causal Mediation Analysis
##
```

```
## # Outcome regression:
##
## Call:
## glm(formula = FRS ~ Smoking + CRP + Smoking * CRP + Age + Sex,
##     family = gaussian(), data = getCall(x$reg.output$yreg)$data,
##     weights = getCall(x$reg.output$yreg)$weights)
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.486723   1.502944   1.655  0.10135
## Smoking     1.760841   1.244489   1.415  0.16040
## CRP         0.308937   0.100363   3.078  0.00273 **
## Age         0.004353   0.295562   0.015  0.98828
## Sex         1.397417   0.269938   5.177 1.28e-06 ***
## Smoking:CRP 0.471868   0.179238   2.633  0.00990 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1.671607)
##
##     Null deviance: 1038.31  on 99  degrees of freedom
## Residual deviance:  157.13  on 94  degrees of freedom
## AIC: 342.98
##
## Number of Fisher Scoring iterations: 2
##
##
##
## # Effect decomposition on the mean difference scale via the natural effect model
##
## Direct counterfactual imputation estimation with
##  bootstrap standard errors, percentile confidence intervals and p-values
##
##      Estimate Std.error 95% CIL 95% CIU  P.val
## cde    4.1202    0.3782  3.2133   4.683 <2e-16 ***
## pnde   4.5004    0.3302  3.8183   5.031 <2e-16 ***
## tnde   5.1558    0.3769  4.4286   5.835 <2e-16 ***
## pnie   0.4152    0.1612  0.1636   0.802 <2e-16 ***
## tnie   1.0705    0.4016  0.4850   2.094 <2e-16 ***
## te     5.5710    0.3793  4.8519   6.277 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (cde: controlled direct effect; pnde: pure natural direct effect; tnde: total natural direct effect;
##
## Relevant variable values:
## $a
## [1] 1
##
## $astar
## [1] 0
##
## $mval
## $mval[[1]]
```

```
## [1] 5
```

3.4.2 With the G-formula

We can use the G-formula.

```r
estGform <- cmest(data = dataSim, model = "gformula", full=FALSE,
                  outcome = "FRS",
                  exposure = "Smoking",
                  mediator = c("CRP"),
                  basec = c("Age", "Sex"), EMint = TRUE,
                  mreg = list("linear"), yreg = "linear",
                  astar = 0, a = 1, mval = list(5))
```

```
##   |                                                                              |
```

This method uses the same regression models as for rb technique. But the estimates are numerically computed with Monte-Carlo and bootstrap (with 200 samples by default):

```r
summary(estGform)
```

```
## Causal Mediation Analysis
##
## # Outcome regression:
##
## Call:
## glm(formula = FRS ~ Smoking + CRP + Smoking * CRP + Age + Sex,
##     family = gaussian(), data = getCall(x$reg.output$yreg)$data,
##     weights = getCall(x$reg.output$yreg)$weights)
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.486723   1.502944   1.655  0.10135
## Smoking     1.760841   1.244489   1.415  0.16040
## CRP         0.308937   0.100363   3.078  0.00273 **
## Age         0.004353   0.295562   0.015  0.98828
## Sex         1.397417   0.269938   5.177 1.28e-06 ***
## Smoking:CRP 0.471868   0.179238   2.633  0.00990 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1.671607)
##
##     Null deviance: 1038.31  on 99  degrees of freedom
## Residual deviance:  157.13  on 94  degrees of freedom
## AIC: 342.98
##
## Number of Fisher Scoring iterations: 2
##
##
## # Mediator regressions:
##
## Call:
## glm(formula = CRP ~ Smoking + Age + Sex, family = gaussian(),
##     data = getCall(x$reg.output$mreg[[1L]])$data, weights = getCall(x$reg.output$mreg[[1L]])$weights)
##
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.4716     1.7944   1.377   0.1716
## Smoking       1.3549     0.3225   4.201 5.95e-05 ***
## Age           0.6270     0.3516   1.783   0.0777 .
## Sex           0.4689     0.3240   1.447   0.1510
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 2.466087)
##
##     Null deviance: 294.04  on 99  degrees of freedom
## Residual deviance: 236.74  on 96  degrees of freedom
## AIC: 379.97
##
## Number of Fisher Scoring iterations: 2
##
##
## # Effect decomposition on the mean difference scale via the g-formula approach
##
## Direct counterfactual imputation estimation with
##  bootstrap standard errors, percentile confidence intervals and p-values
##
##      Estimate Std.error 95% CIL 95% CIU  P.val
## cde    4.1202    0.3975  3.1562   4.808 <2e-16 ***
## pnde   4.4495    0.3256  3.7676   5.047 <2e-16 ***
## tnde   5.0889    0.3373  4.5812   5.843 <2e-16 ***
## pnie   0.4186    0.1464  0.1467   0.703 <2e-16 ***
## tnie   1.0579    0.3574  0.4652   1.860 <2e-16 ***
## te     5.5074    0.3443  4.9531   6.295 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (cde: controlled direct effect; pnde: pure natural direct effect; tnde: total natural direct effect;
##
## Relevant variable values:
## $a
## [1] 1
##
## $astar
## [1] 0
##
## $mval
## $mval[[1]]
## [1] 5
```

In this specific setting, the call with model = "rb", and inference="bootstrap" is equivalent to model = "gformula". This is not always the case: the g-formula also handles exposure-affected confounders for M-Y relation.

This can be checked by setting the seed for the Bootstrap:

```
set.seed(1)
estRBBoot <- cmest(data = dataSim, model = "rb", full=FALSE,
               outcome = "FRS",
```

```r
                exposure = "Smoking",
                mediator = c("CRP"),
                basec = c("Age", "Sex"), EMint = TRUE,
                mreg = list("linear"), yreg = "linear",
                astar = 0, a = 1, mval = list(1), inference = "bootstrap")
```

```
##   |                                                                              |
```

```r
set.seed(1)
estGformBoot <- cmest(data = dataSim, model = "gformula", full=FALSE,
                outcome = "FRS",
                exposure = "Smoking",
                mediator = c("CRP"),
                basec = c("Age", "Sex"), EMint = TRUE,
                mreg = list("linear"), yreg = "linear",
                astar = 0, a = 1, mval = list(1))
```

```
##   |                                                                              |
```

```r
Compar <- cbind(estRBBoot$effect.pe, estRBBoot$effect.se, estGformBoot$effect.pe, estGformBoot$effect.se
colnames(Compar) <- c("rb-boot","SE rb-boot","Gform","SE Gform")
Compar
```

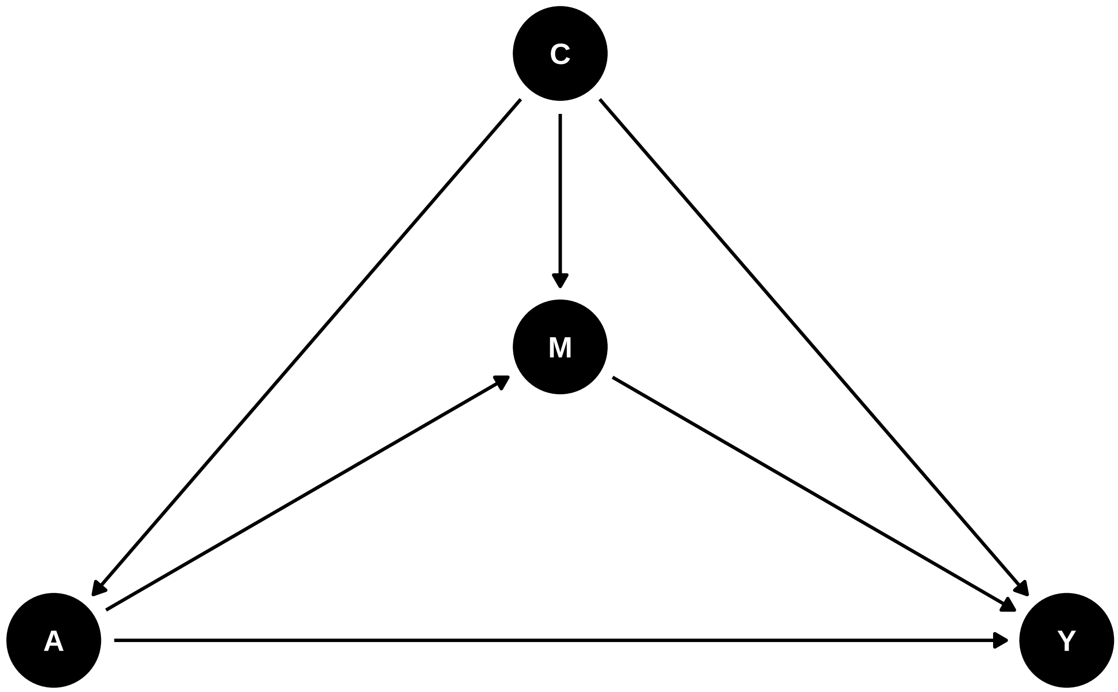
```
##        rb-boot SE rb-boot     Gform  SE Gform
## cde  2.2327096  1.0670433 2.2327096 1.0670433
## pnde 4.5873637  0.3539009 4.5873637 0.3539009
## tnde 5.2266924  0.3819970 5.2266924 0.3819970
## pnie 0.4185745  0.1649222 0.4185745 0.1649222
## tnie 1.0579031  0.3901685 1.0579031 0.3901685
## te   5.6452668  0.3757896 5.6452668 0.3757896
```

4 Study of Smoking - Fat - CRP - FRS

Let's now consider we have two intermediate mediators with Fat impacting CRP level. We can consider two settings:

```
 - the joint mediating effect of M1 and M2
```

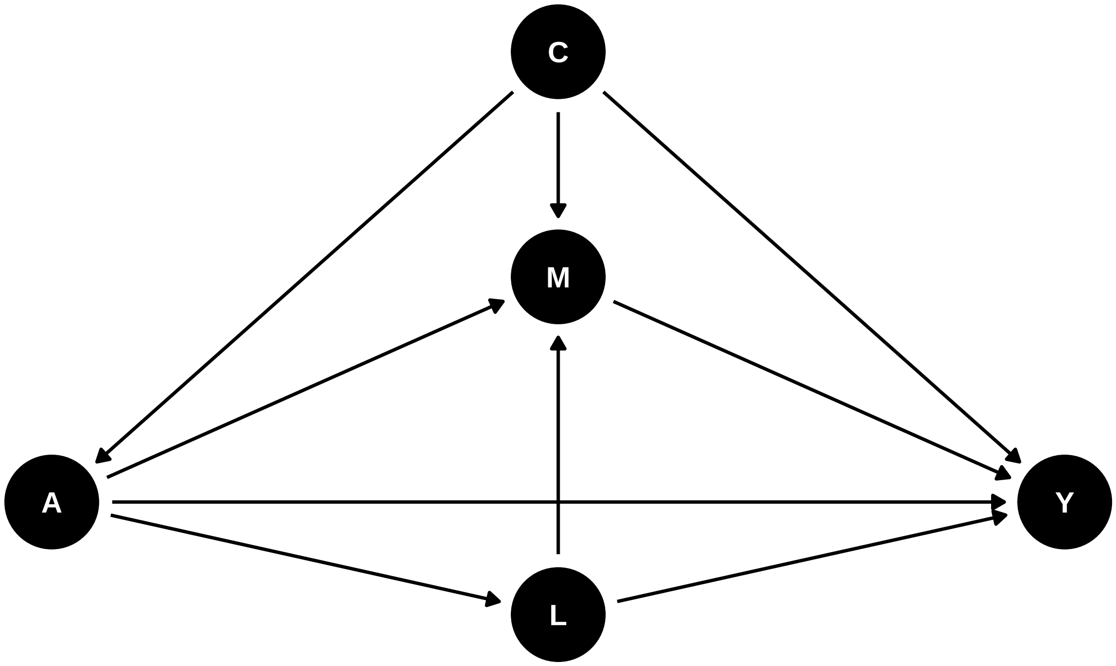
```r
cmdag(outcome = "FRS", exposure = "Smoking", mediator = c("Fat","CRP"),
      basec = c("Age", "Sex"))
```



A (exposure): Smoking
M (mediator): Fat, CRP
Y (outcome): FRS
C (confounders not affected by the exposure): Age, Sex

```
- the focus on M2 considering M1 as a confounder
```

```r
cmdag(outcome = "FRS", exposure = "Smoking", mediator = c("CRP"), postc = c("Fat"),
      basec = c("Age", "Sex"))
```



A (exposure): Smoking
M (mediator): CRP
Y (outcome): FRS
C (confounders not affected by the exposure): Age, Sex
L (confounders affected by the exposure): Fat

In this setting, we have Fat which is a confounder of CRP - FRS that is affected by Smoking. Only G-formula will be possible here.

4.1 Joint mediating effect of Fat and CRP

The function cmest handles multiple mediators. We can use the regression-based technique here with the Bootstrap technique for the uncertainty.

```
estJointRB <- cmest(data = dataSim, model = "rb", outcome = "FRS", exposure = "Smoking", full=FALSE,
mediator = c("Fat", "CRP"), basec = c("Age", "Sex"), EMint = TRUE,
mreg = list("logistic", "linear"), yreg = "linear",
astar = 0, a = 1, mval = list(0,5))
```

```
##   |                                                                                   |
```

```
summary(estJointRB)
```

```
## Causal Mediation Analysis
##
## # Outcome regression:
##
## Call:
## glm(formula = FRS ~ Smoking + Fat + CRP + Smoking * Fat + Smoking *
##     CRP + Age + Sex, family = gaussian(), data = getCall(x$reg.output$yreg)$data,
##     weights = getCall(x$reg.output$yreg)$weights)
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.53284    1.38140   1.834  0.06996 .
## Smoking      2.43377    1.15972   2.099  0.03859 *
```

```
## Fat          1.10905    0.35454   3.128  0.00236 **
## CRP          0.18238    0.10082   1.809  0.07373 .
## Age         -0.01084    0.27176  -0.040  0.96826
## Sex          1.45709    0.24780   5.880 6.54e-08 ***
## Smoking:Fat  0.26611    0.56036   0.475  0.63600
## Smoking:CRP  0.38763    0.18189   2.131  0.03574 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1.404391)
##
##     Null deviance: 1038.3  on 99  degrees of freedom
## Residual deviance:  129.2  on 92  degrees of freedom
## AIC: 327.41
##
## Number of Fisher Scoring iterations: 2
##
##
## # Mediator regressions:
##
## Call:
## glm(formula = Fat ~ Smoking + Age + Sex, family = binomial(),
##     data = getCall(x$reg.output$mreg[[1L]])$data, weights = getCall(x$reg.output$mreg[[1L]])$weights
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.03811    2.40822  -0.431    0.666
## Smoking     -0.24375    0.42940  -0.568    0.570
## Age          0.34278    0.47374   0.724    0.469
## Sex          0.06694    0.43553   0.154    0.878
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 129.49  on 99  degrees of freedom
## Residual deviance: 128.60  on 96  degrees of freedom
## AIC: 136.6
##
## Number of Fisher Scoring iterations: 4
##
##
##
## Call:
## glm(formula = CRP ~ Smoking + Age + Sex, family = gaussian(),
##     data = getCall(x$reg.output$mreg[[2L]])$data, weights = getCall(x$reg.output$mreg[[2L]])$weights
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.4716     1.7944   1.377   0.1716
## Smoking       1.3549     0.3225   4.201 5.95e-05 ***
## Age           0.6270     0.3516   1.783   0.0777 .
## Sex           0.4689     0.3240   1.447   0.1510
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## (Dispersion parameter for gaussian family taken to be 2.466087)
##
##     Null deviance: 294.04  on 99  degrees of freedom
## Residual deviance: 236.74  on 96  degrees of freedom
## AIC: 379.97
##
## Number of Fisher Scoring iterations: 2
##
##
## # Effect decomposition on the mean difference scale via the regression-based approach
##
## Direct counterfactual imputation estimation with
##  bootstrap standard errors, percentile confidence intervals and p-values
##
##      Estimate Std.error  95% CIL 95% CIU  P.val
## cde   4.37189   0.44246  3.47629   5.170 <2e-16 ***
## pnde  4.85608   0.39521  4.15439   5.616 <2e-16 ***
## tnde  5.37861   0.33826  4.81959   6.132 <2e-16 ***
## pnie  0.23601   0.20706 -0.21865   0.551   0.41
## tnie  0.75854   0.43797 -0.02102   1.632   0.08 .
## te    5.61462   0.35916  4.96036   6.395 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (cde: controlled direct effect; pnde: pure natural direct effect; tnde: total natural direct effect;
##
## Relevant variable values:
## $a
## [1] 1
##
## $astar
## [1] 0
##
## $mval
## $mval[[1]]
## [1] 0
##
## $mval[[2]]
## [1] 5
```

Here, we do not consider the impact of Fat on CRP. We directly look at the joint effect.

4.2 The stochastic/randomized analogues to NDE/NIE for CRP

The function cmest handles post-exposure confounders with the gformula method, and postc and postcreg arguments.

```r
estLGForm <- cmest(data = dataSim, model = "gformula", outcome = "FRS", exposure = "Smoking", full=FALSE
mediator = c("CRP"), basec = c("Age", "Sex"), EMint = TRUE, postc = "Fat", postcreg = list("logistic"),
mreg = list("linear"), yreg = "linear",
astar = 0, a = 1, mval = list(5))
```

```
##   |                                                                              |
```

The procedure now includes three regressions: one for FRS, one for CRP and one for Fat. It then estimates the direct and indirect effects under randomized intervention.

```r
summary(estLGForm)
```

```
## Causal Mediation Analysis
##
## # Outcome regression:
##
## Call:
## glm(formula = FRS ~ Smoking + CRP + Smoking * CRP + Age + Sex +
##     Fat, family = gaussian(), data = getCall(x$reg.output$yreg)$data,
##     weights = getCall(x$reg.output$yreg)$weights)
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.4841345  1.3718413   1.811   0.0734 .
## Smoking      2.3576955  1.1438134   2.061   0.0421 *
## CRP          0.1696740  0.0968004   1.753   0.0829 .
## Age         -0.0006594  0.2697825  -0.002   0.9981
## Sex          1.4550942  0.2467315   5.897 5.91e-08 ***
## Fat          1.2157788  0.2730525   4.453 2.36e-05 ***
## Smoking:CRP  0.4243397  0.1639503   2.588   0.0112 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1.392695)
##
##     Null deviance: 1038.31  on 99  degrees of freedom
## Residual deviance:  129.52  on 93  degrees of freedom
## AIC: 325.65
##
## Number of Fisher Scoring iterations: 2
##
##
## # Mediator regressions:
##
## Call:
## glm(formula = CRP ~ Smoking + Age + Sex + Fat, family = gaussian(),
##     data = getCall(x$reg.output$mreg[[1L]])$data, weights = getCall(x$reg.output$mreg[[1L]])$weights
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.1061     1.6474   1.278    0.204
## Smoking       1.4286     0.2962   4.824 5.36e-06 ***
## Age           0.5243     0.3232   1.622    0.108
## Sex           0.4493     0.2971   1.512    0.134
## Fat           1.3283     0.3032   4.381 3.05e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 2.073242)
##
##     Null deviance: 294.04  on 99  degrees of freedom
## Residual deviance: 196.96  on 95  degrees of freedom
## AIC: 363.57
##
```

```
## Number of Fisher Scoring iterations: 2
##
##
## # Regressions for mediator-outcome confounders affected by the exposure:
##
## Call:
## glm(formula = Fat ~ Smoking + Age + Sex, family = binomial(),
##     data = getCall(x$reg.output$postcreg[[1L]])$data, weights = getCall(x$reg.output$postcreg[[1L]])$
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.03811    2.40822  -0.431    0.666
## Smoking     -0.24375    0.42940  -0.568    0.570
## Age          0.34278    0.47374   0.724    0.469
## Sex          0.06694    0.43553   0.154    0.878
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 129.49  on 99  degrees of freedom
## Residual deviance: 128.60  on 96  degrees of freedom
## AIC: 136.6
##
## Number of Fisher Scoring iterations: 4
##
##
## # Effect decomposition on the mean difference scale via the g-formula approach
##
## Direct counterfactual imputation estimation with
##  bootstrap standard errors, percentile confidence intervals and p-values
##
##       Estimate Std.error 95% CIL 95% CIU  P.val
## cde    4.50371   0.36088 3.68214   5.044 <2e-16 ***
## rpnde  4.77814   0.28186 4.18304   5.240 <2e-16 ***
## rtnde  5.39562   0.31949 4.66761   5.981 <2e-16 ***
## rpnie  0.24690   0.12960 0.01358   0.509   0.03 *
## rtnie  0.86439   0.33777 0.29396   1.600 <2e-16 ***
## te     5.64252   0.32772 4.90975   6.186 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (cde: controlled direct effect; rpnde: randomized analogue of pure natural direct effect; rtnde: ranc
##
## Relevant variable values:
## $a
## [1] 1
##
## $astar
## [1] 0
##
## $mval
## $mval[[1]]
## [1] 5
```

5 Sensitivity analyses

These analyses assume that there is no unmeasured confounding and that the mediator is measured perfectly. The package includes the function cmsens which provides tools to assess two issues in the data: unmeasured confounding (sens="uc") and measurement error (sens="me").

5.1 Measurement error

The continuous mediator CRP is likely suffering from measurement error (this is indeed what I simulated). For continuous variables, cmsens implements two techniques, the regression calibration (for independent continuous variables) and the SIMEX approach. This is usable with regression technique and g-formula.

Let's go back to the first model (no post-exposure confounding). Here is with SIMEX considering masurement error with standard deviation of 0.5, 1, 1.5 in the regression based estimation:

```
me1simex <- cmsens(object = estRB2, sens = "me", MEmethod = "simex",
MEvariable = "CRP", MEvartype = "con", MEerror = c(0.5, 1,1.5))
```

And with the Gformula. For this, because of computation time, I reduce the simulations to 100 for SIMEX and consider only 1 for the standard deviation of the error:

```
me2simex <- cmsens(object = estGform, sens = "me", MEmethod = "simex",
MEvariable = "CRP", MEvartype = "con", MEerror = c(1), B=100)
```

```
##   |                                                                                |
```

```
summary(me1simex)
```

```
## Sensitivity Analysis For Measurement Error
##
## The variable measured with error: CRP
## Type of the variable measured with error: continuous
##
## # Measurement error 1:
## [1] 0.5
##
## ## Error-corrected regressions for measurement error 1:
##
## ### Outcome regression:
## Call:
## simexreg(reg = getCall(x$sens[[1L]]$reg.output$yreg)$reg, formula = FRS ~
##     Smoking + CRP + Smoking * CRP + Age + Sex, data = getCall(x$sens[[1L]]$reg.output$yreg)$data,
##     MEvariable = "CRP", MEvartype = "continuous", MEerror = 0.5,
##     variance = TRUE, lambda = c(0.5, 1, 1.5, 2), B = 200, weights = getCall(x$sens[[1L]]$reg.output$y
##
## Naive coefficient estimates:
## (Intercept)     Smoking         CRP         Age         Sex Smoking:CRP
## 2.486723376 1.760841426 0.308936513 0.004352802 1.397417130 0.471868177
##
## Naive var-cov estimates:
##             (Intercept)     Smoking          CRP           Age           Sex
## (Intercept)  2.25884167 -0.23388005 -0.027437735 -0.4080038569 -0.0153259552
## Smoking     -0.23388005  1.54875397  0.056866277 -0.0257507361  0.0170237498
## CRP         -0.02743773  0.05686628  0.010072718 -0.0059664924 -0.0025813822
## Age         -0.40800386 -0.02575074 -0.005966492  0.0873570602  0.0008017148
## Sex         -0.01532596  0.01702375 -0.002581382  0.0008017148  0.0728666187
## Smoking:CRP  0.03261408 -0.21696585 -0.009836705  0.0050268376 -0.0023823806
```

```
##              Smoking:CRP
## (Intercept)  0.032614078
## Smoking     -0.216965852
## CRP         -0.009836705
## Age          0.005026838
## Sex         -0.002382381
## Smoking:CRP  0.032126091
##
## Variable measured with error:
## CRP
## Measurement error:
## [1] 0.5
##
## Error-corrected results:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.364475   1.515855   1.560  0.12216
## Smoking      1.272754   1.382398   0.921  0.35957
## CRP          0.341858   0.107325   3.185  0.00196 **
## Age         -0.005822   0.299775  -0.019  0.98455
## Sex          1.358297   0.274063   4.956 3.17e-06 ***
## Smoking:CRP  0.536101   0.200718   2.671  0.00891 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## ### Mediator regressions:
## Call:
## simexreg(reg = getCall(x$sens[[1L]]$reg.output$mreg[[1L]])$reg,
##     formula = CRP ~ Smoking + Age + Sex, data = getCall(x$sens[[1L]]$reg.output$mreg[[1L]])$data,
##     MEvariable = "CRP", MEvartype = "continuous", MEerror = 0.5,
##     variance = TRUE, lambda = c(0.5, 1, 1.5, 2), B = 200, weights = getCall(x$sens[[1L]]$reg.output$m
##
## Naive coefficient estimates:
## (Intercept)     Smoking         Age         Sex
##   2.4716154   1.3548883   0.6270273   0.4689042
##
## Naive var-cov estimates:
##             (Intercept)      Smoking          Age          Sex
## (Intercept)  3.21994122 -0.054974094 -0.625591495 -0.031114398
## Smoking     -0.05497409  0.104004016  0.003245383 -0.005239678
## Age         -0.62559149  0.003245383  0.123620169 -0.001329962
## Sex         -0.03111440 -0.005239678 -0.001329962  0.104947611
##
## Variable measured with error:
## CRP
## Measurement error:
## [1] 0.5
##
## Error-corrected results:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.5398     1.8150   1.399   0.1649
## Smoking       1.3639     0.3261   4.183 6.37e-05 ***
## Age           0.6155     0.3534   1.741   0.0848 .
## Sex           0.4532     0.3236   1.401   0.1646
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## ## Error-corrected causal effects on the mean difference scale for measurement error 1:
##      Estimate Std.error 95% CIL 95% CIU   P.val
## cde    3.9533    0.4535  3.0644   4.842 < 2e-16 ***
## pnde   4.3945    0.3604  3.6881   5.101 < 2e-16 ***
## tnde   5.1257    0.3332  4.4727   5.779 < 2e-16 ***
## pnie   0.4663    0.1840  0.1057   0.827 0.01127 *
## tnie   1.1975    0.3690  0.4743   1.921 0.00117 **
## te     5.5919    0.3530  4.9002   6.284 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## ----------------------------------------------------------------
##
## # Measurement error 2:
## [1] 1
##
## ## Error-corrected regressions for measurement error 2:
##
## ### Outcome regression:
## Call:
## simexreg(reg = getCall(x$sens[[2L]]$reg.output$yreg)$reg, formula = FRS ~
##     Smoking + CRP + Smoking * CRP + Age + Sex, data = getCall(x$sens[[2L]]$reg.output$yreg)$data,
##     MEvariable = "CRP", MEvartype = "continuous", MEerror = 1,
##     variance = TRUE, lambda = c(0.5, 1, 1.5, 2), B = 200, weights = getCall(x$sens[[2L]]$reg.output$y
##
## Naive coefficient estimates:
## (Intercept)     Smoking         CRP         Age         Sex Smoking:CRP
## 2.486723376 1.760841426 0.308936513 0.004352802 1.397417130 0.471868177
##
## Naive var-cov estimates:
##             (Intercept)     Smoking          CRP           Age           Sex
## (Intercept)  2.25884167 -0.23388005 -0.027437735 -0.4080038569 -0.0153259552
## Smoking     -0.23388005  1.54875397  0.056866277 -0.0257507361  0.0170237498
## CRP         -0.02743773  0.05686628  0.010072718 -0.0059664924 -0.0025813822
## Age         -0.40800386 -0.02575074 -0.005966492  0.0873570602  0.0008017148
## Sex         -0.01532596  0.01702375 -0.002581382  0.0008017148  0.0728666187
## Smoking:CRP  0.03261408 -0.21696585 -0.009836705  0.0050268376 -0.0023823806
##              Smoking:CRP
## (Intercept)  0.032614078
## Smoking     -0.216965852
## CRP         -0.009836705
## Age          0.005026838
## Sex         -0.002382381
## Smoking:CRP  0.032126091
##
## Variable measured with error:
## CRP
## Measurement error:
## [1] 1
##
## Error-corrected results:
##             Estimate Std. Error t value Pr(>|t|)
```

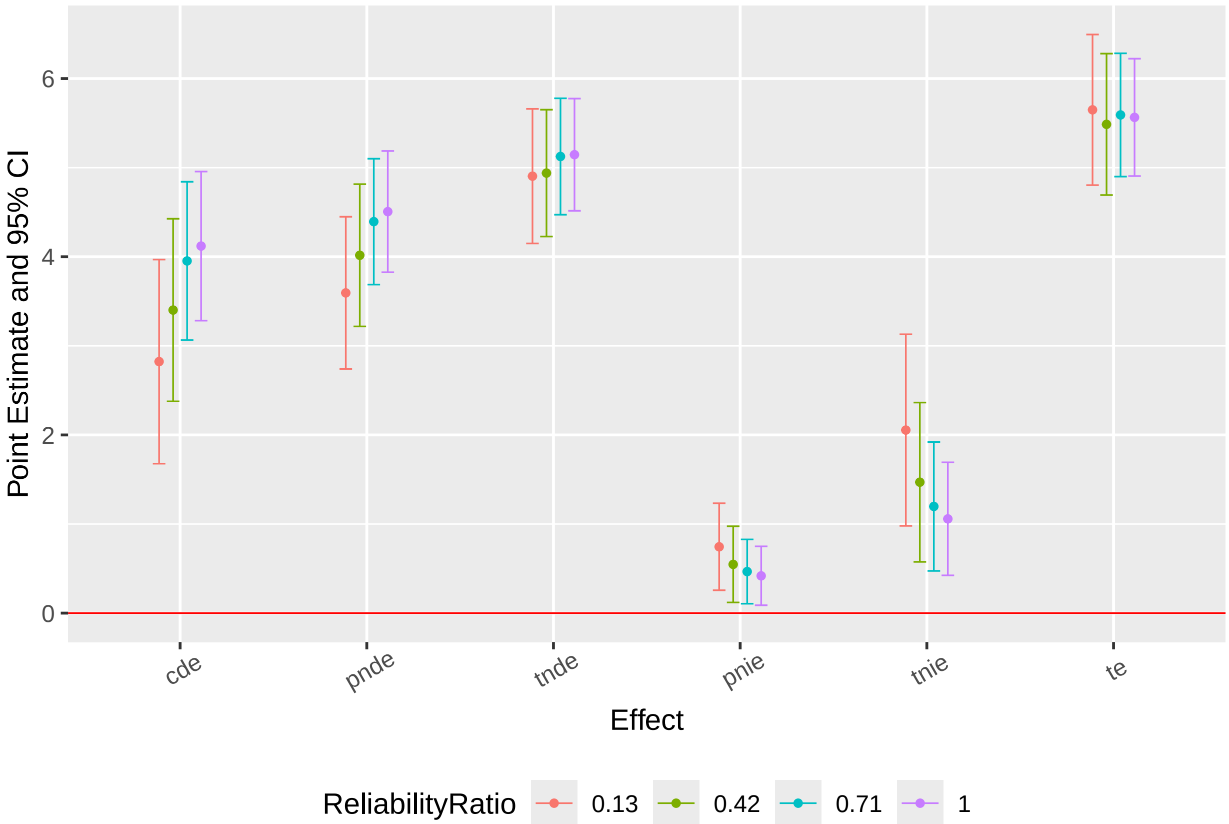
```
## (Intercept)  2.29881    1.51805   1.514  0.13330
## Smoking     -0.19130    1.62094  -0.118  0.90630
## CRP          0.42604    0.12986   3.281  0.00145 **
## Age         -0.08582    0.29725  -0.289  0.77342
## Sex          1.27866    0.26713   4.787 6.29e-06 ***
## Smoking:CRP  0.71863    0.23380   3.074  0.00277 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## ### Mediator regressions:
## Call:
## simexreg(reg = getCall(x$sens[[2L]]$reg.output$mreg[[1L]])$reg,
##     formula = CRP ~ Smoking + Age + Sex, data = getCall(x$sens[[2L]]$reg.output$mreg[[1L]])$data,
##     MEvariable = "CRP", MEvartype = "continuous", MEerror = 1,
##     variance = TRUE, lambda = c(0.5, 1, 1.5, 2), B = 200, weights = getCall(x$sens[[2L]]$reg.output$m
##
## Naive coefficient estimates:
## (Intercept)     Smoking         Age         Sex
##   2.4716154   1.3548883   0.6270273   0.4689042
##
## Naive var-cov estimates:
##             (Intercept)      Smoking          Age          Sex
## (Intercept)  3.21994122 -0.054974094 -0.625591495 -0.031114398
## Smoking     -0.05497409  0.104004016  0.003245383 -0.005239678
## Age         -0.62559149  0.003245383  0.123620169 -0.001329962
## Sex         -0.03111440 -0.005239678 -0.001329962  0.104947611
##
## Variable measured with error:
## CRP
## Measurement error:
## [1] 0.5
##
## Error-corrected results:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.5398     1.8150   1.399   0.1649
## Smoking       1.3639     0.3261   4.183 6.37e-05 ***
## Age           0.6155     0.3534   1.741   0.0848 .
## Sex           0.4532     0.3236   1.401   0.1646
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## ## Error-corrected causal effects on the mean difference scale for measurement error 2:
##      Estimate Std.error 95% CIL 95% CIU   P.val
## cde    3.4019    0.5230  2.3768   4.427 7.8e-11 ***
## pnde   4.0166    0.4073  3.2182   4.815 < 2e-16 ***
## tnde   4.9393    0.3632  4.2276   5.651 < 2e-16 ***
## pnie   0.5470    0.2182  0.1194   0.975 0.01216 *
## tnie   1.4698    0.4561  0.5758   2.364 0.00127 **
## te     5.4864    0.4050  4.6925   6.280 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## ----------------------------------------------------------------
##
```

```
## # Measurement error 3:
## [1] 1.5
##
## ## Error-corrected regressions for measurement error 3:
##
## ### Outcome regression:
## Call:
## simexreg(reg = getCall(x$sens[[3L]]$reg.output$yreg)$reg, formula = FRS ~
##     Smoking + CRP + Smoking * CRP + Age + Sex, data = getCall(x$sens[[3L]]$reg.output$yreg)$data,
##     MEvariable = "CRP", MEvartype = "continuous", MEerror = 1.5,
##     variance = TRUE, lambda = c(0.5, 1, 1.5, 2), B = 200, weights = getCall(x$sens[[3L]]$reg.output$y
##
## Naive coefficient estimates:
## (Intercept)     Smoking         CRP         Age         Sex Smoking:CRP
## 2.486723376 1.760841426 0.308936513 0.004352802 1.397417130 0.471868177
##
## Naive var-cov estimates:
##             (Intercept)     Smoking          CRP           Age           Sex
## (Intercept)  2.25884167 -0.23388005 -0.027437735 -0.4080038569 -0.0153259552
## Smoking     -0.23388005  1.54875397  0.056866277 -0.0257507361  0.0170237498
## CRP         -0.02743773  0.05686628  0.010072718 -0.0059664924 -0.0025813822
## Age         -0.40800386 -0.02575074 -0.005966492  0.0873570602  0.0008017148
## Sex         -0.01532596  0.01702375 -0.002581382  0.0008017148  0.0728666187
## Smoking:CRP  0.03261408 -0.21696585 -0.009836705  0.0050268376 -0.0023823806
##              Smoking:CRP
## (Intercept)  0.032614078
## Smoking     -0.216965852
## CRP         -0.009836705
## Age          0.005026838
## Sex         -0.002382381
## Smoking:CRP  0.032126091
##
## Variable measured with error:
## CRP
## Measurement error:
## [1] 1.5
##
## Error-corrected results:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.1992     1.4303   1.538 0.127502
## Smoking      -1.8165     1.8872  -0.963 0.338247
## CRP           0.5274     0.1380   3.823 0.000237 ***
## Age          -0.1756     0.2841  -0.618 0.538052
## Sex           1.1764     0.2650   4.439 2.46e-05 ***
## Smoking:CRP   0.9280     0.2727   3.403 0.000982 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## ### Mediator regressions:
## Call:
## simexreg(reg = getCall(x$sens[[3L]]$reg.output$mreg[[1L]])$reg,
##     formula = CRP ~ Smoking + Age + Sex, data = getCall(x$sens[[3L]]$reg.output$mreg[[1L]])$data,
##     MEvariable = "CRP", MEvartype = "continuous", MEerror = 1.5,
##     variance = TRUE, lambda = c(0.5, 1, 1.5, 2), B = 200, weights = getCall(x$sens[[3L]]$reg.output$m
```

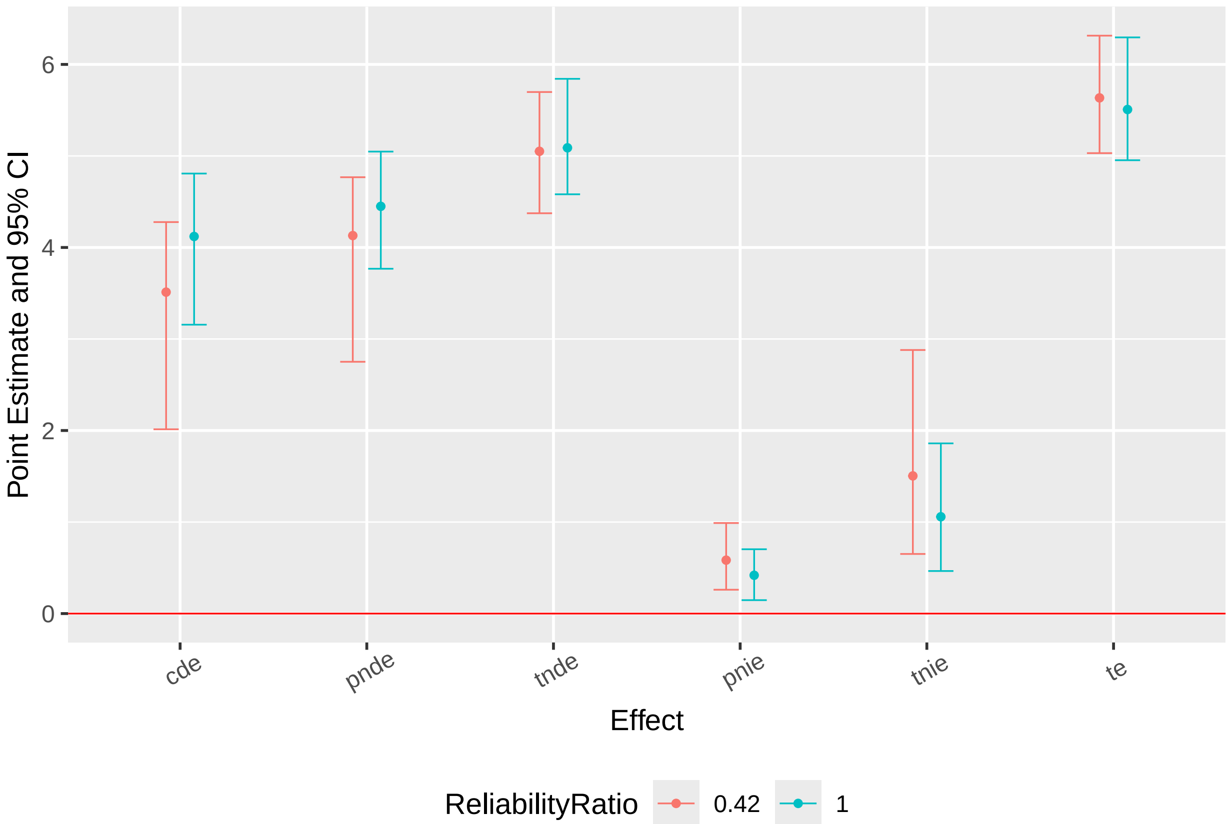
```
##
## Naive coefficient estimates:
## (Intercept)     Smoking         Age         Sex
##   2.4716154   1.3548883   0.6270273   0.4689042
##
## Naive var-cov estimates:
##             (Intercept)      Smoking          Age          Sex
## (Intercept)  3.21994122 -0.054974094 -0.625591495 -0.031114398
## Smoking     -0.05497409  0.104004016  0.003245383 -0.005239678
## Age         -0.62559149  0.003245383  0.123620169 -0.001329962
## Sex         -0.03111440 -0.005239678 -0.001329962  0.104947611
##
## Variable measured with error:
## CRP
## Measurement error:
## [1] 0.5
##
## Error-corrected results:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.5398     1.8150   1.399   0.1649
## Smoking       1.3639     0.3261   4.183 6.37e-05 ***
## Age           0.6155     0.3534   1.741   0.0848 .
## Sex           0.4532     0.3236   1.401   0.1646
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## ## Error-corrected causal effects on the mean difference scale for measurement error 3:
##      Estimate Std.error 95% CIL 95% CIU    P.val
## cde    2.8235    0.5844  1.6780   3.969 1.36e-06 ***
## pnde   3.5946    0.4362  2.7397   4.449 2.22e-16 ***
## tnde   4.9047    0.3852  4.1497   5.660  < 2e-16 ***
## pnie   0.7446    0.2491  0.2564   1.233  0.00280 **
## tnie   2.0547    0.5485  0.9797   3.130  0.00018 ***
## te     5.6493    0.4314  4.8038   6.495  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## ----------------------------------------------------------------
##
## (cde: controlled direct effect; pnde: pure natural direct effect; tnde: total natural direct effect;
##
## Relevant variable values:
## $a
## [1] 1
##
## $astar
## [1] 0
##
## $mval
## $mval[[1]]
## [1] 5
##
##
## $basecval
```

```
## $basecval[[1]]
## [1] 5.054444
##
## $basecval[[2]]
## [1] 0.38
```

```
ggcmsens(me1simex) +
ggplot2::theme(axis.text.x = ggplot2::element_text(angle = 30, vjust = 0.8))
```




```
ggcmsens(me2simex) +
ggplot2::theme(axis.text.x = ggplot2::element_text(angle = 30, vjust = 0.8))
```



The total effect doesn't change much but the indirect effect does. There is an attenuation of the effect of the mediator in not taking into account the measurement error.

In the model with fat as a post-exposure confounder (I run only 100 simulations and just for a variance of 1 to reduce the time):

```
meLsimex <- cmsens(object = estLGForm , sens = "me", MEmethod = "simex",
MEvariable = "CRP", MEvartype = "con", MEerror = c(1), B = 100)
```

```
##   |                                                                          |
```

```
summary(meLsimex)
```

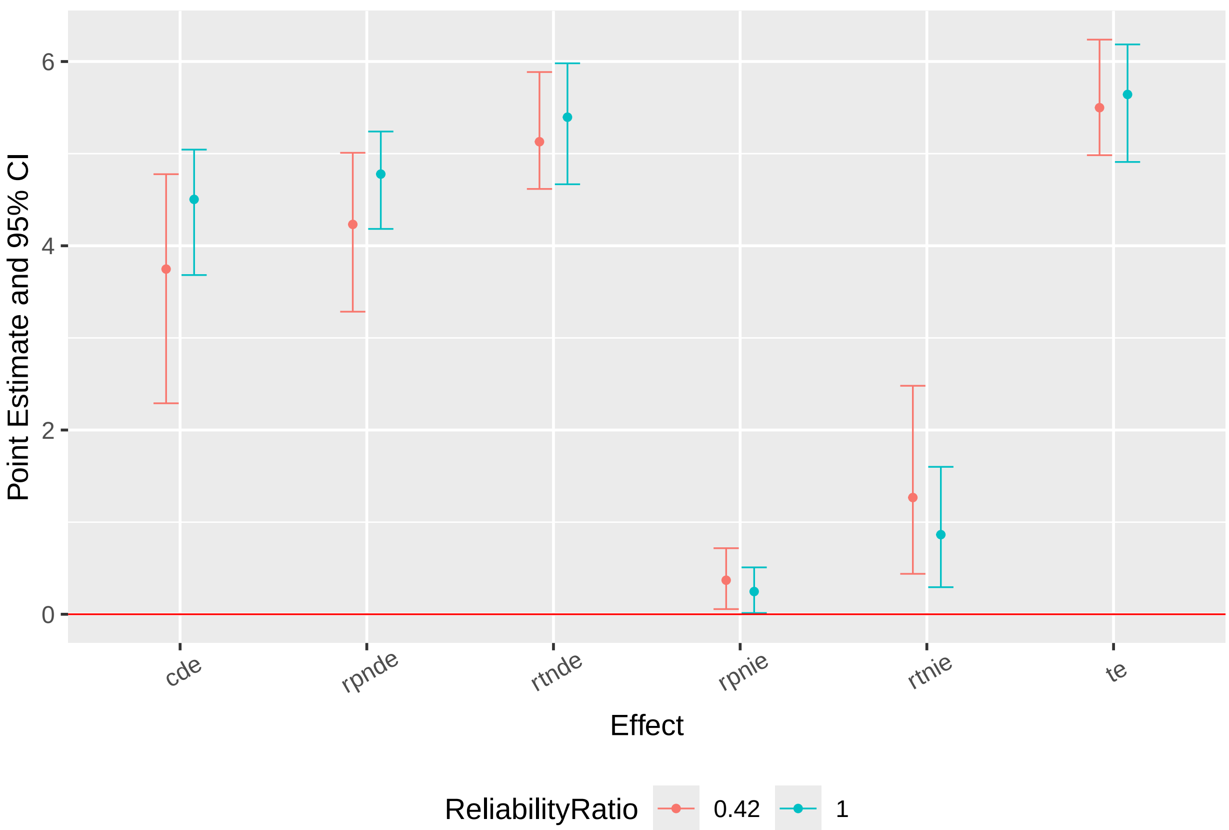
```
## Sensitivity Analysis For Measurement Error
##
## The variable measured with error: CRP
## Type of the variable measured with error: continuous
##
## # Measurement error 1:
## [1] 1
##
## ## Error-corrected regressions for measurement error 1:
##
## ### Outcome regression:
## Call:
## simexreg(reg = getCall(x$sens[[1L]]$reg.output$yreg)$reg, formula = FRS ~
##     Smoking + CRP + Smoking * CRP + Age + Sex + Fat, data = getCall(x$sens[[1L]]$reg.output$yreg)$da
##     MEvariable = "CRP", MEvartype = "continuous", MEerror = 1,
##     variance = TRUE, lambda = c(0.5, 1, 1.5, 2), B = 100, weights = getCall(x$sens[[1L]]$reg.output$
##
```

```
## Naive coefficient estimates:
##  (Intercept)      Smoking          CRP          Age          Sex          Fat
##  2.484134525  2.357695476  0.169674013 -0.000659403  1.455094230  1.215778843
##  Smoking:CRP
##  0.424339689
##
## Naive var-cov estimates:
##              (Intercept)     Smoking          CRP           Age           Sex
## (Intercept)  1.8819484253 -0.19493450 -0.022841492 -0.3399267018 -0.0127763116
## Smoking     -0.1949345002  1.30830917  0.043185373 -0.0216050558  0.0159197123
## CRP         -0.0228414920  0.04318537  0.009370313 -0.0049357593 -0.0025558263
## Age         -0.3399267018 -0.02160506 -0.004935759  0.0727825756  0.0006533646
## Sex         -0.0127763116  0.01591971 -0.002555826  0.0006533646  0.0608764338
## Fat         -0.0001587613  0.03660208 -0.008540274 -0.0003073735  0.0035370486
## Smoking:CRP  0.0271785410 -0.18219542 -0.007861560  0.0042001129 -0.0021231482
##                       Fat  Smoking:CRP
## (Intercept) -0.0001587613  0.027178541
## Smoking      0.0366020792 -0.182195424
## CRP         -0.0085402739 -0.007861560
## Age         -0.0003073735  0.004200113
## Sex          0.0035370486 -0.002123148
## Fat          0.0745576468 -0.002914685
## Smoking:CRP -0.0029146849  0.026879713
##
## Variable measured with error:
## CRP
## Measurement error:
## [1] 1
##
## Error-corrected results:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.40921    1.35008   1.784  0.07760 .
## Smoking      0.49983    1.53417   0.326  0.74531
## CRP          0.27123    0.12469   2.175  0.03215 *
## Age         -0.05578    0.26862  -0.208  0.83597
## Sex          1.34723    0.24429   5.515 3.12e-07 ***
## Fat          0.91898    0.29198   3.147  0.00221 **
## Smoking:CRP  0.65866    0.21830   3.017  0.00329 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## ### Mediator regressions:
## Call:
## simexreg(reg = getCall(x$sens[[1L]]$reg.output$mreg[[1L]])$reg,
##     formula = CRP ~ Smoking + Age + Sex + Fat, data = getCall(x$sens[[1L]]$reg.output$mreg[[1L]])$da
##     MEvariable = "CRP", MEvartype = "continuous", MEerror = 1,
##     variance = TRUE, lambda = c(0.5, 1, 1.5, 2), B = 100, weights = getCall(x$sens[[1L]]$reg.output$
##
## Naive coefficient estimates:
## (Intercept)     Smoking         Age         Sex         Fat
##   2.1061342   1.4285989   0.5243007   0.4492681   1.3283412
##
## Naive var-cov estimates:
##             (Intercept)      Smoking          Age          Sex          Fat
```

```
## (Intercept)  2.71396875 -0.047620601 -0.523979037 -0.025783940 -0.025298249
## Smoking     -0.04762060  0.087719416  0.002333822 -0.004480425  0.005102181
## Age         -0.52397904  0.002333822  0.104477511 -0.001012989 -0.007110640
## Sex         -0.02578394 -0.004480425 -0.001012989  0.088249666 -0.001359188
## Fat         -0.02529825  0.005102181 -0.007110640 -0.001359188  0.091946488
##
## Variable measured with error:
## CRP
## Measurement error:
## [1] 1
##
## Error-corrected results:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.3305     1.1147   2.091   0.0392 *
## Smoking       1.4259     0.3105   4.592 1.35e-05 ***
## Age           0.5034     0.2315   2.175   0.0321 *
## Sex           0.3752     0.3346   1.121   0.2650
## Fat           1.2646     0.2635   4.800 5.90e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## ### Regressions for mediator-outcome confounders affected by the exposure:
##
## Call:
## glm(formula = Fat ~ Smoking + Age + Sex, family = binomial(),
##     data = getCall(x$sens[[1L]]$reg.output$postcreg[[1L]])$data,
##     weights = getCall(x$sens[[1L]]$reg.output$postcreg[[1L]])$weights)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.03811    2.40822  -0.431    0.666
## Smoking     -0.24375    0.42940  -0.568    0.570
## Age          0.34278    0.47374   0.724    0.469
## Sex          0.06694    0.43553   0.154    0.878
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 129.49  on 99  degrees of freedom
## Residual deviance: 128.60  on 96  degrees of freedom
## AIC: 136.6
##
## Number of Fisher Scoring iterations: 4
##
##
##
## ## Error-corrected causal effects on the mean difference scale for measurement error 1:
##       Estimate Std.error 95% CIL 95% CIU  P.val
## cde    3.74718   0.66154 2.29035   4.777 <2e-16 ***
## rpnde  4.23195   0.46706 3.28469   5.009 <2e-16 ***
## rtnde  5.12948   0.33184 4.61684   5.886 <2e-16 ***
## rpnie  0.36960   0.17947 0.05593   0.717   0.02 *
## rtnie  1.26713   0.55444 0.43897   2.480 <2e-16 ***
## te     5.49907   0.33831 4.98321   6.238 <2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## ----------------------------------------------------------------
##
## (cde: controlled direct effect; rpnde: randomized analogue of pure natural direct effect; rtnde: ranc
##
## Relevant variable values:
## $a
## [1] 1
##
## $astar
## [1] 0
##
## $mval
## $mval[[1]]
## [1] 5
```

```
ggcmsens(meLsimex) +
ggplot2::theme(axis.text.x = ggplot2::element_text(angle = 30, vjust = 0.8))
```



6 References:

Chen, Zhuangzhuang, Jun Li, Hao Rao, et al. 2025. «Fat Distribution, Inflammatory Mechanisms, and Cardiovascular Disease Risk: Mediation Analysis Based on the Framingham Risk Score». BMC Cardiovascular Disorders 25 (1): 664. https://doi.org/10.1186/s12872-025-05135-3.

Lin, Sheng-Hsuan, Jessica Young, Roger Logan, Eric J. Tchetgen Tchetgen, et Tyler J. VanderWeele. 2017. «Parametric Mediational G-Formula Approach to Mediation Analysis with Time-Varying Exposures, Mediators, and Confounders». Epidemiology (Cambridge, Mass.) 28 (2): 266-74. https://doi.org/10.1097/

EDE.0000000000000609.