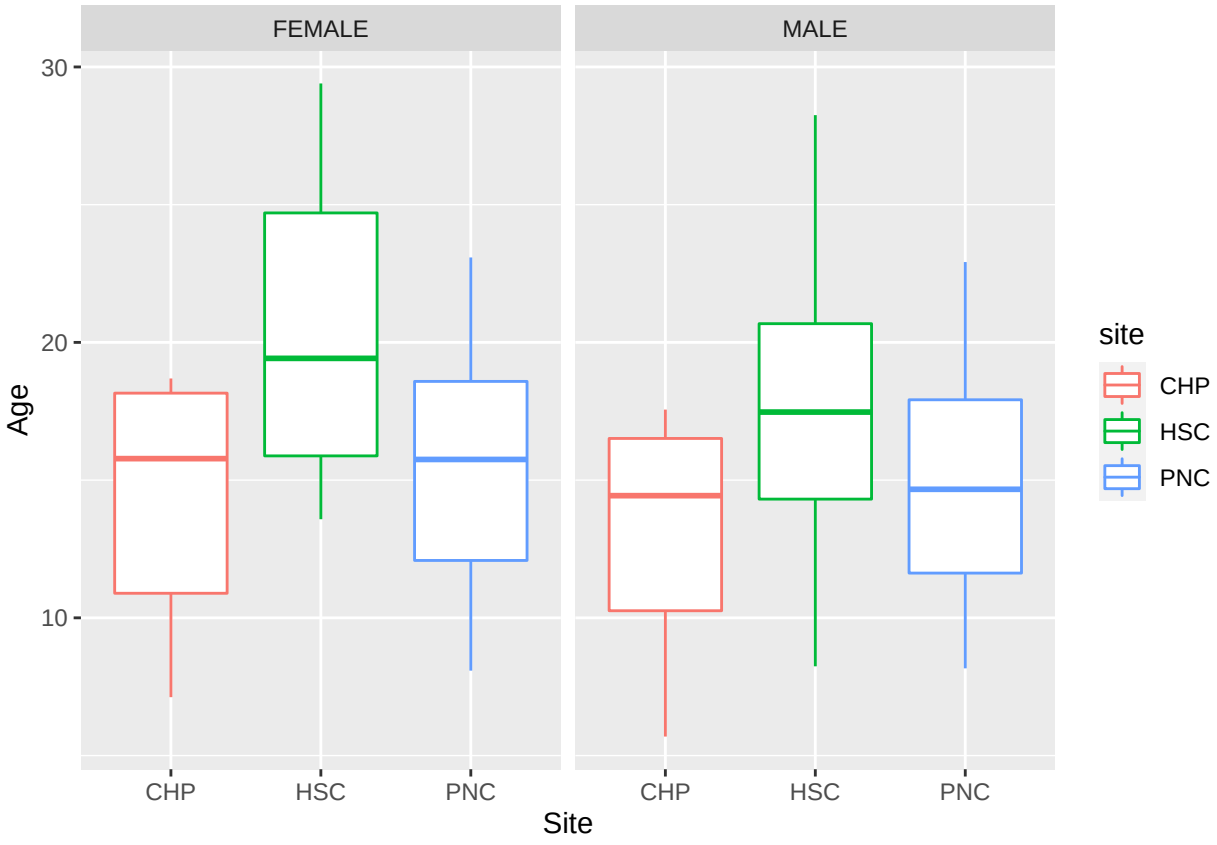
MSKIDS: Exploratory Analysis and ComBat Harmonization with Healthy Controls" (Part 4: Adjusting for ICV, interactions and non-linear effects)

Virgilio Gonzenbach

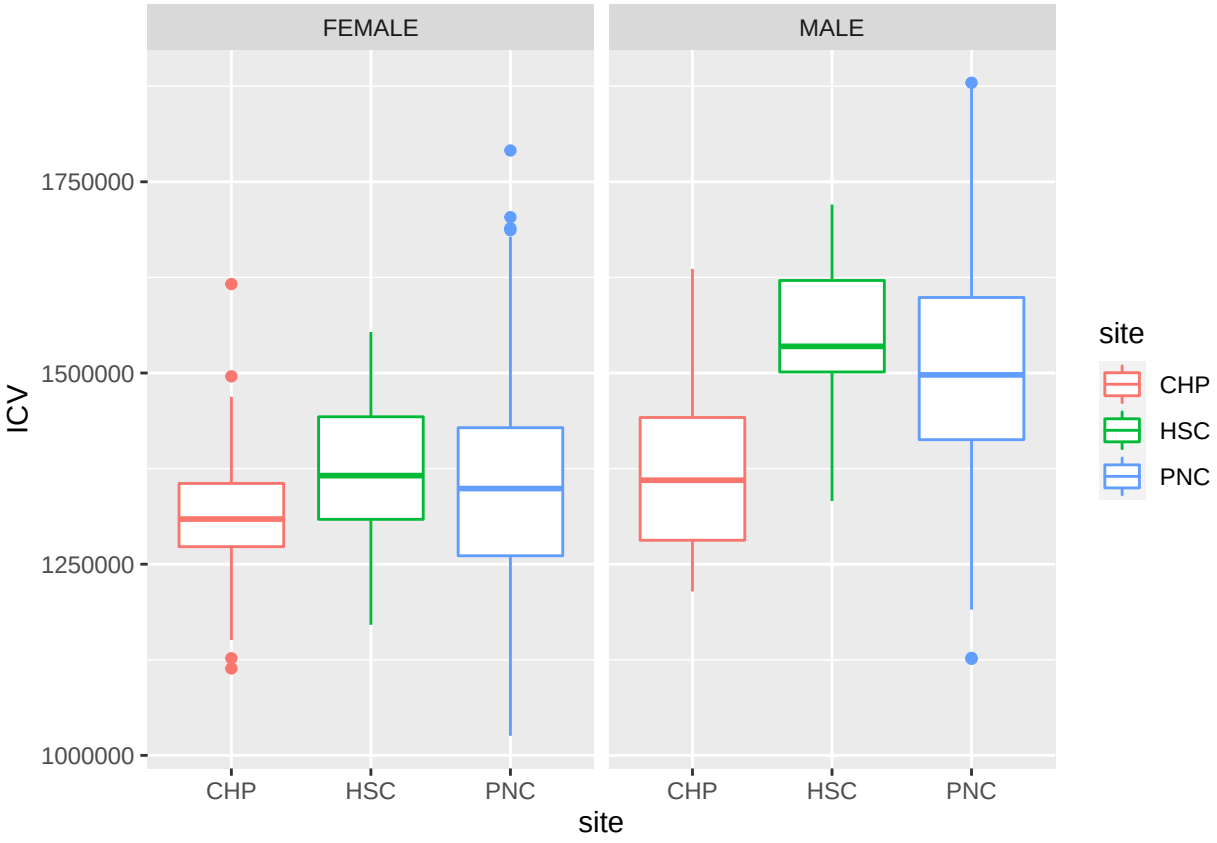
11/20/2020

Age



Females older than males in HSC.

ICV




```
## Analysis of Variance Table
##
## Model 1: X702 ~ sex + age
## Model 2: X702 ~ sex + age + batch
##   Res.Df        RSS Df  Sum of Sq      F  Pr(>F)
## 1   1276 2.0134e+13
## 2   1274 1.9938e+13  2 1.9531e+11 6.2397 0.00201 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

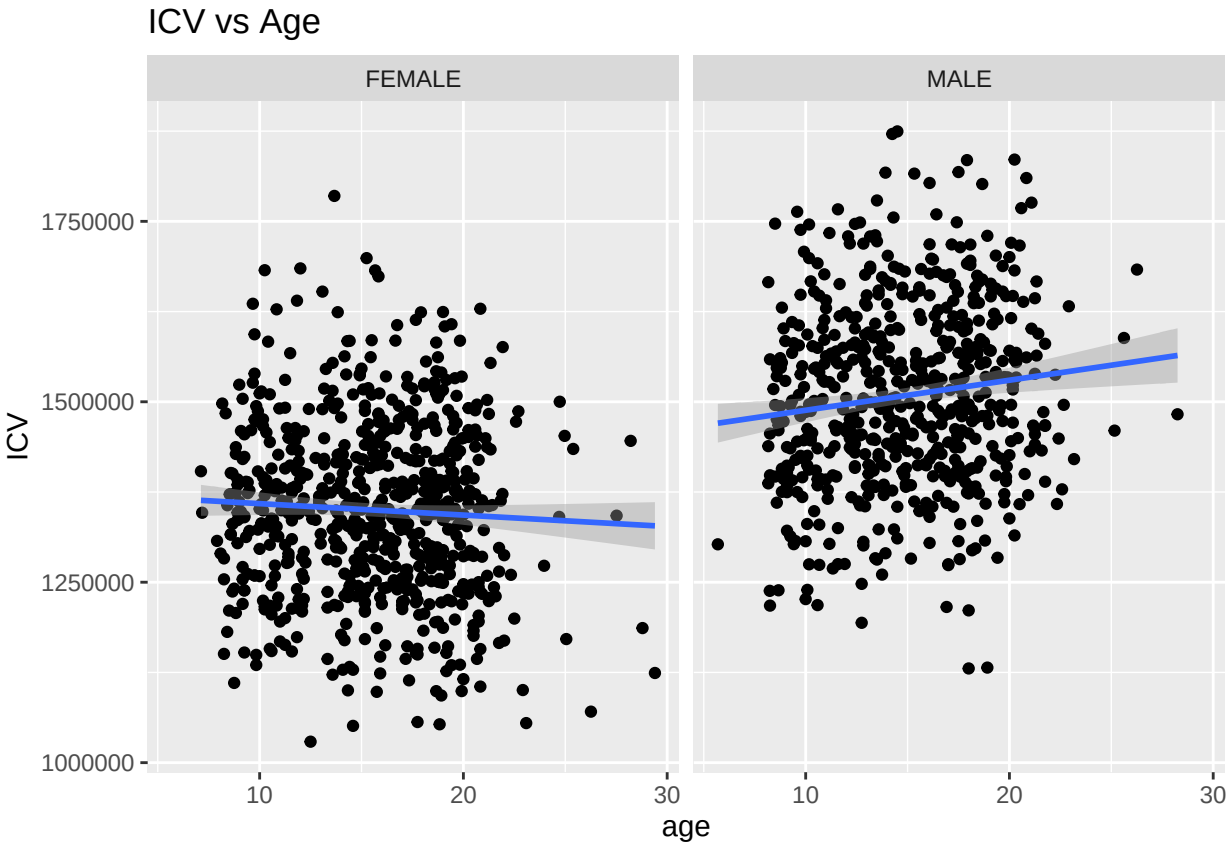
Harmonized

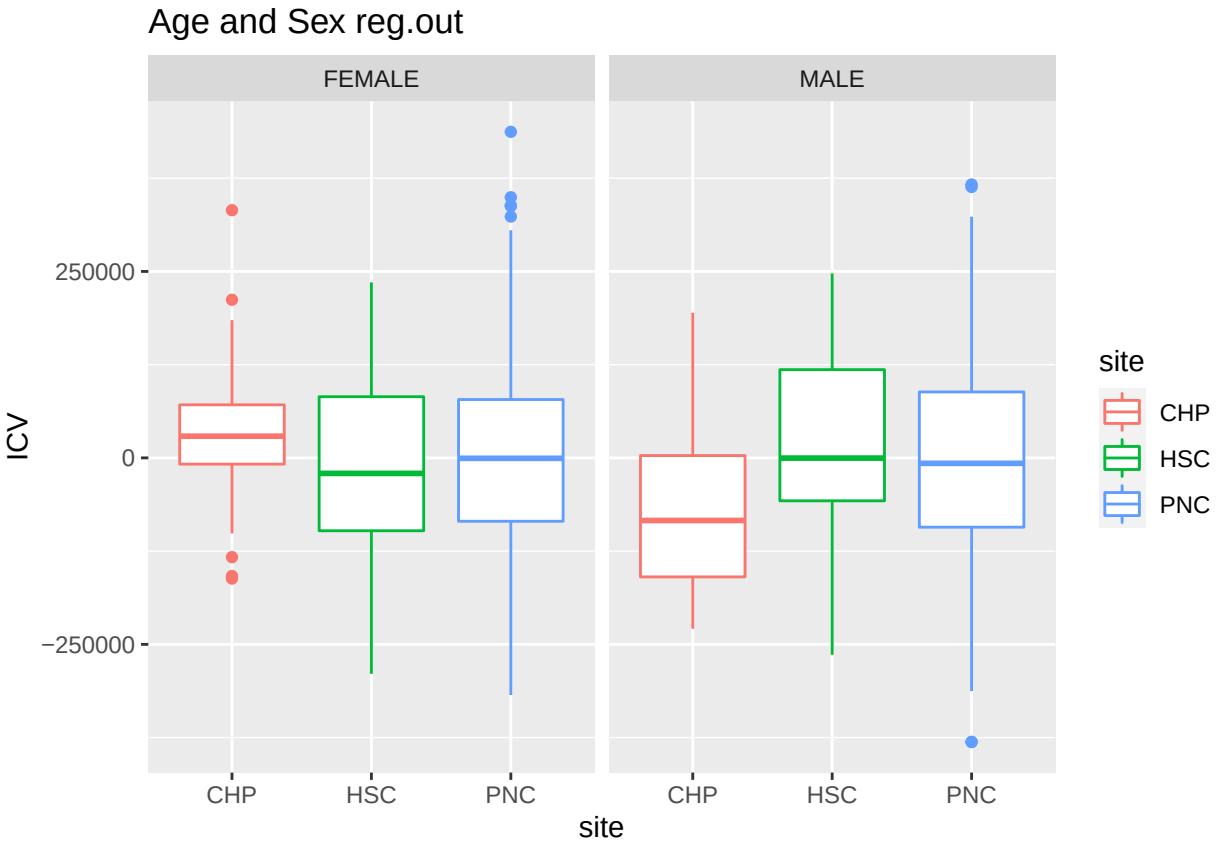
Interaction

```
##
## Call:
## lm(formula = scale(X702) ~ sex + scale(age) + sex * scale(age) +
##     batch, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.66422 -0.60542 -0.01387  0.57157  2.92512
##
## Coefficients:
```

```
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                -0.4900883  0.1415265  -3.463 0.000552 ***
## sexMALE                     1.0807667  0.0475942  22.708  < 2e-16 ***
## scale(age)                 -0.0422095  0.0325500  -1.297 0.194949
## batchHSC-SIEMENSPRISMAFIT  0.0028628  0.1821631   0.016 0.987464
## batchPNC-SIEMENSTIMTRIO    0.0005652  0.1429621   0.004 0.996846
## sexMALE:scale(age)         0.1526571  0.0478278   3.192 0.001448 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8428 on 1273 degrees of freedom
## Multiple R-squared:  0.2924, Adjusted R-squared:  0.2897
## F-statistic: 105.2 on 5 and 1273 DF,  p-value: < 2.2e-16
```

```
## `geom_smooth()` using formula 'y ~ x'
```

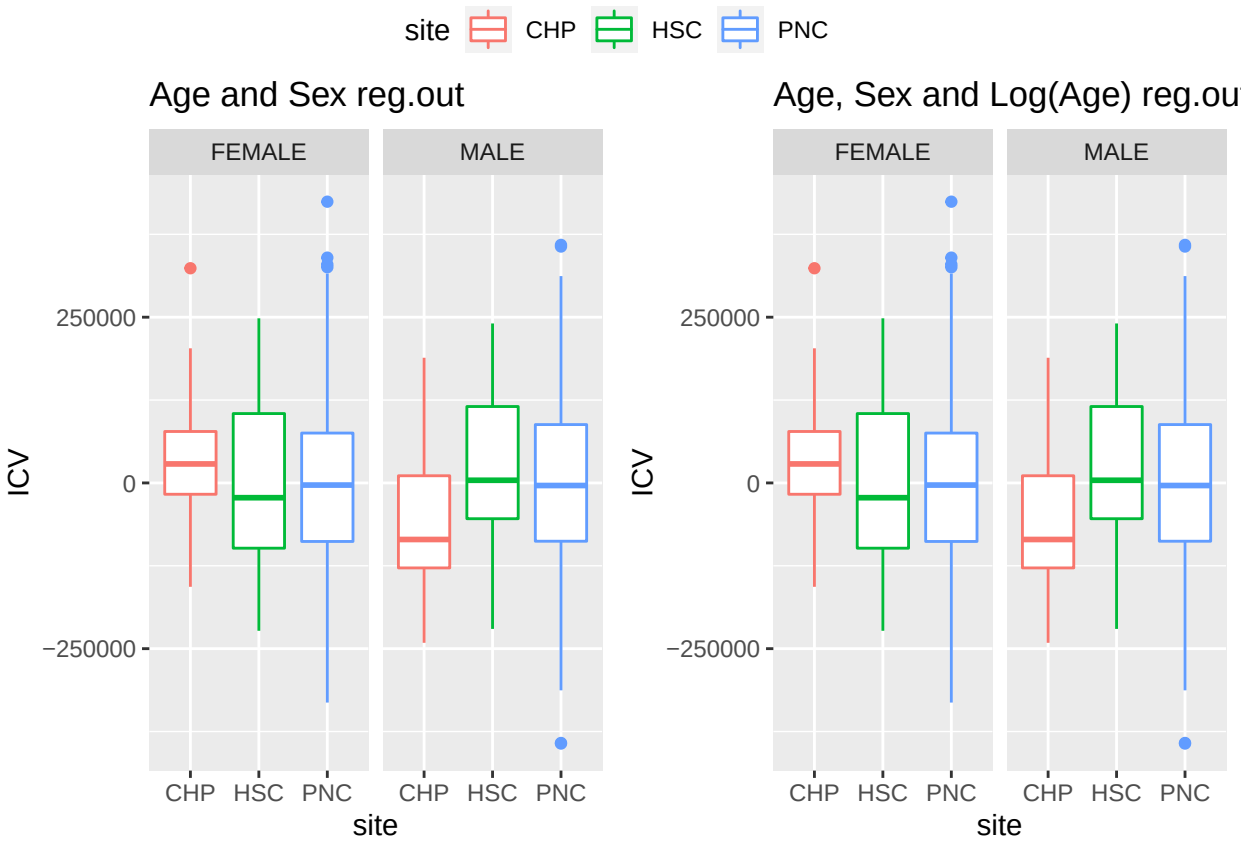




Non-linearity

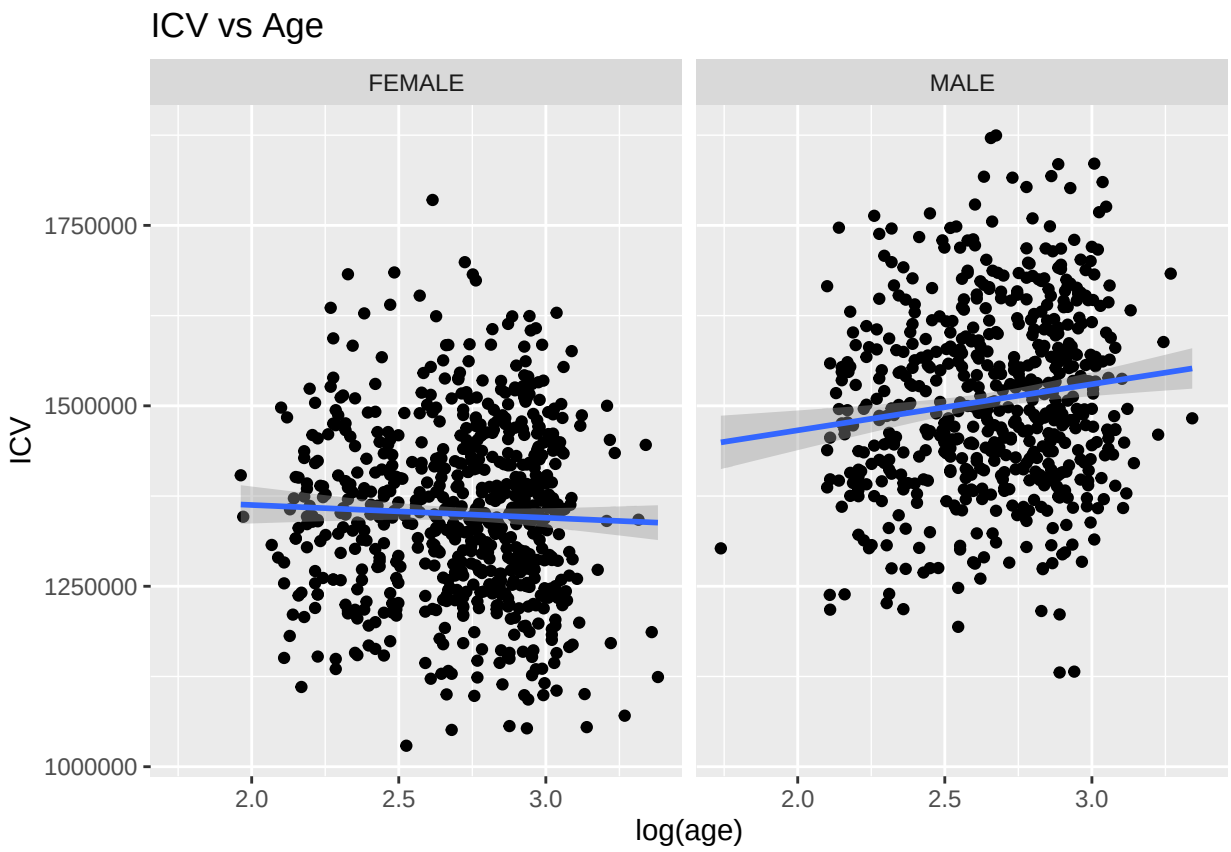
```
##
## Call:
## lm(formula = scale(X702) ~ sex + scale(age) + log(age) + scale(age^2) +
##     batch, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.59798 -0.60037 -0.02426  0.55575  2.89828
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                -4.83022    6.51606  -0.741    0.459
## sexMALE                     1.07560    0.04762  22.587   <2e-16 ***
## scale(age)                 -0.34175    1.28336  -0.266    0.790
## log(age)                    1.62230    2.42943   0.668    0.504
## scale(age^2)               -0.07157    0.65077  -0.110    0.912
## batchHSC-SIEMENSPRISMAFIT  0.03101    0.19167   0.162    0.871
## batchPNC-SIEMENSTIMTRIO   -0.03343    0.14476  -0.231    0.817
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8433 on 1272 degrees of freedom
## Multiple R-squared:  0.2922, Adjusted R-squared:  0.2888
## F-statistic: 87.51 on 6 and 1272 DF,  p-value: < 2.2e-16
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
##
## Call:
## lm(formula = scale(X702) ~ sex + scale(log(age)) + scale(age^2) +
##     batch, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.60160 -0.60208 -0.02381  0.55822  2.90307
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                -0.47074    0.14173  -3.321 0.000921 ***
## sexMALE                     1.07581    0.04760  22.603  < 2e-16 ***
## scale(log(age))             0.26605    0.08226   3.234 0.001251 **
## scale(age^2)               -0.24342    0.08374  -2.907 0.003712 **
## batchHSC-SIEMENSPRISMAFIT   0.04548    0.18375   0.247 0.804572
## batchPNC-SIEMENSTIMTRIO    -0.02797    0.14324  -0.195 0.845233
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.843 on 1273 degrees of freedom
## Multiple R-squared:  0.2921, Adjusted R-squared:  0.2894
## F-statistic: 105.1 on 5 and 1273 DF,  p-value: < 2.2e-16
```

```
## `geom_smooth()` using formula 'y ~ x'
```



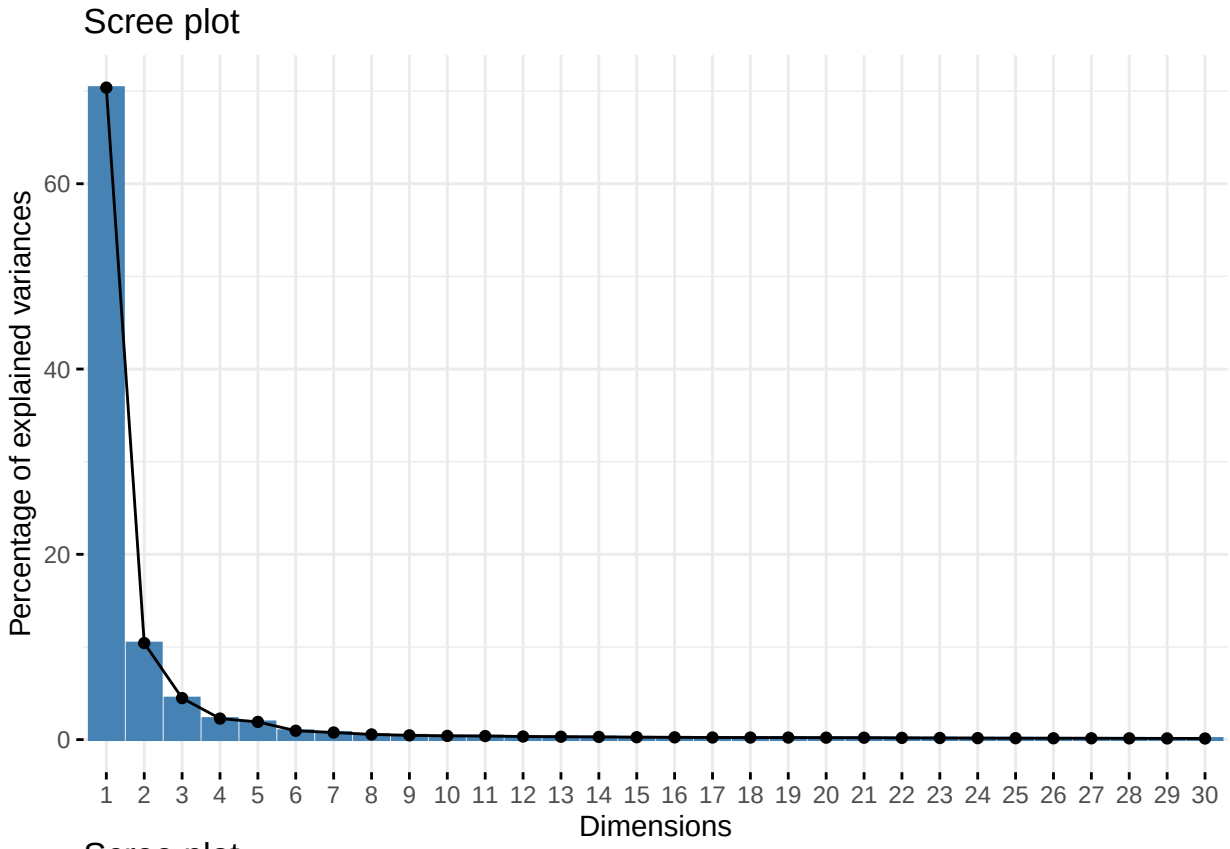

```
## `geom_smooth()` using formula 'y ~ x'
```

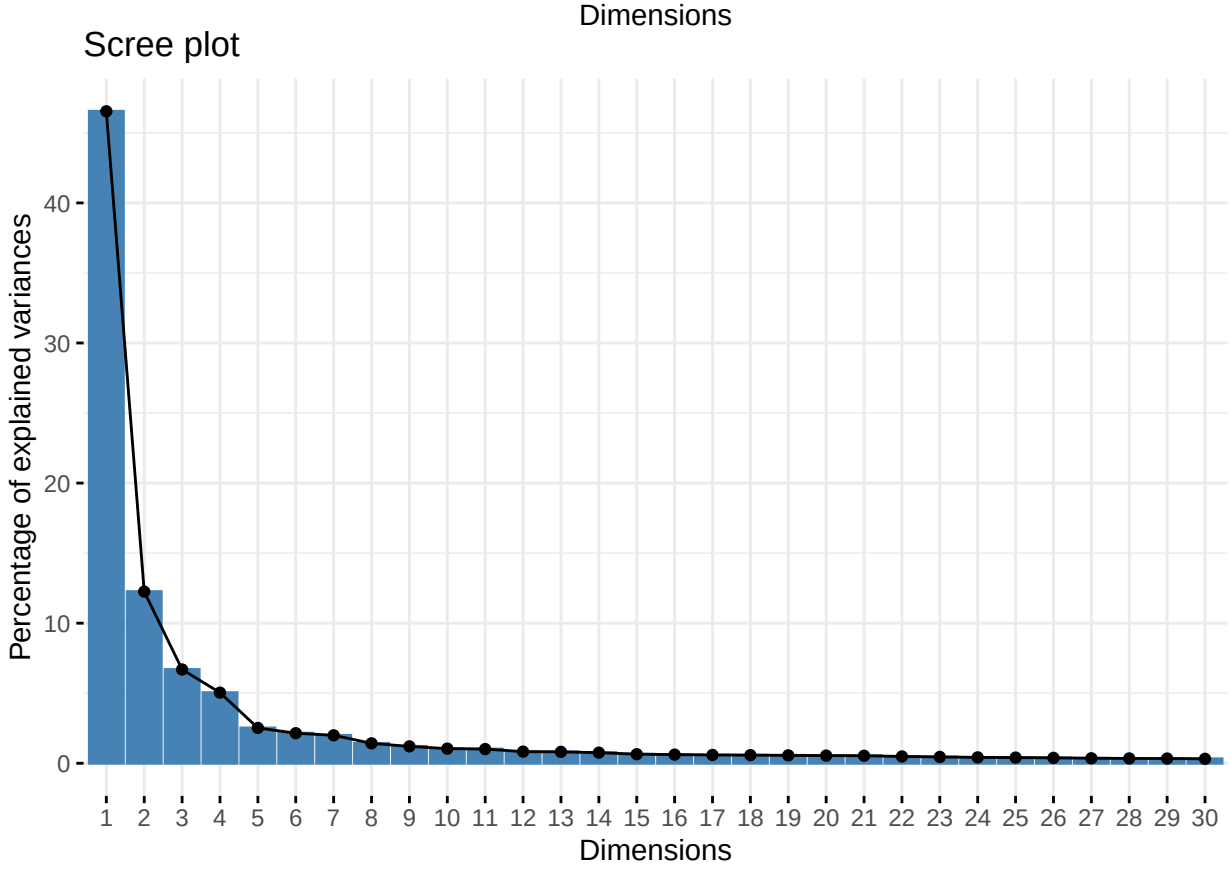
Number of significant ANOVA after accounting for sites

For each of the 145 ROI volumes, two models were used:
- Model 1: region ~ sex + age + ICV - Model 2: region ~ sex + age + ICV + batch

```
##    Raw Combat
##     63      0
```

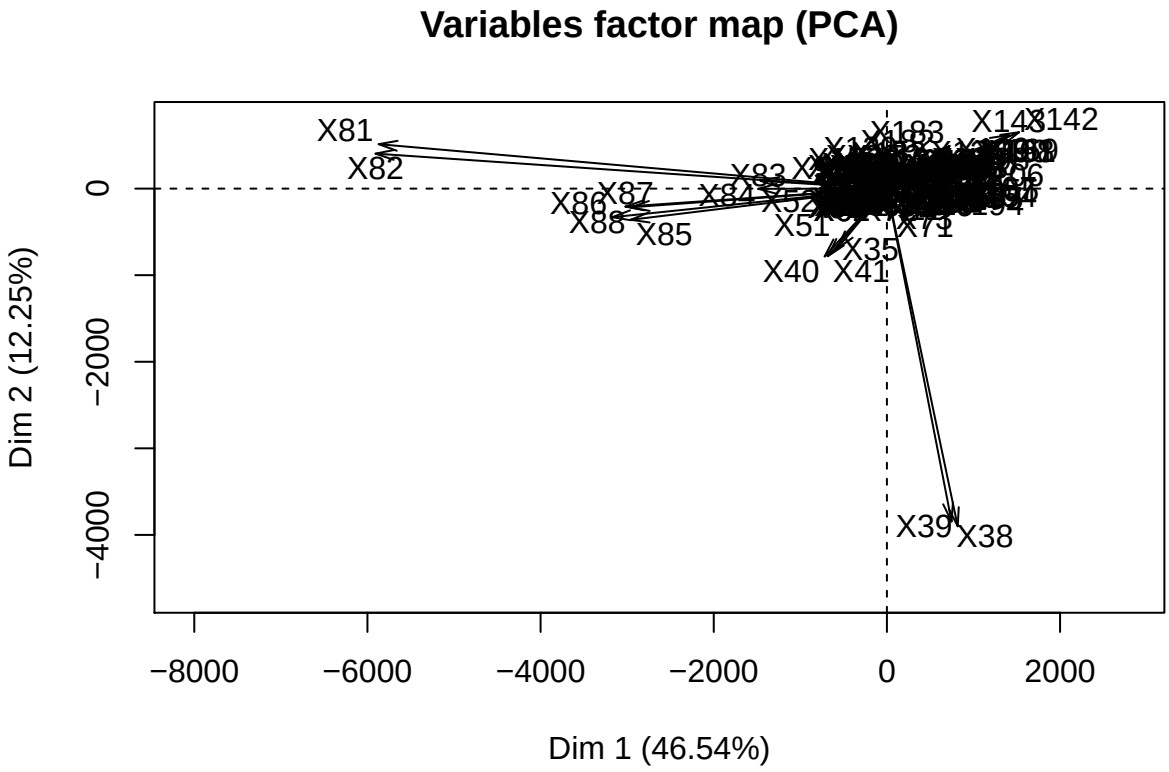
(Covariance) PCA





Components 1 and 2

- PC1: White matter
- PC2: Cerebellum

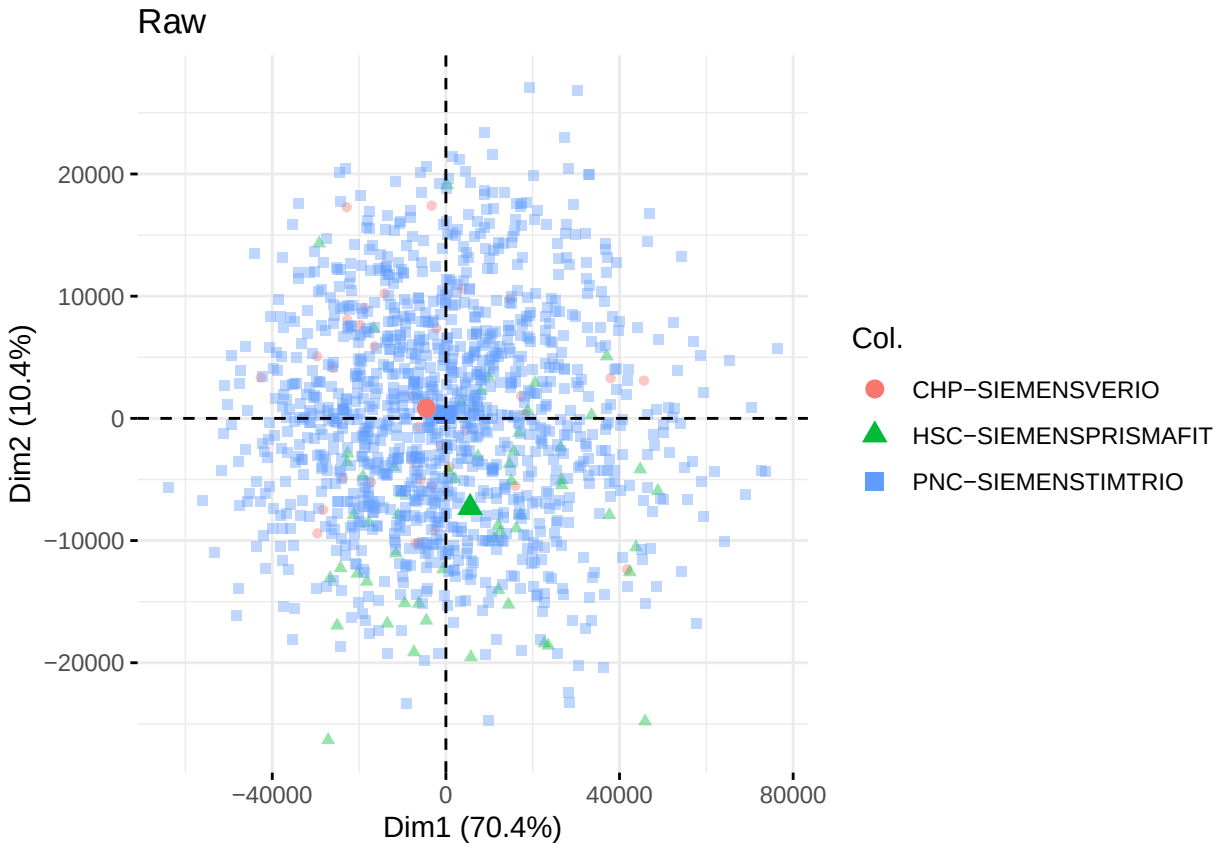


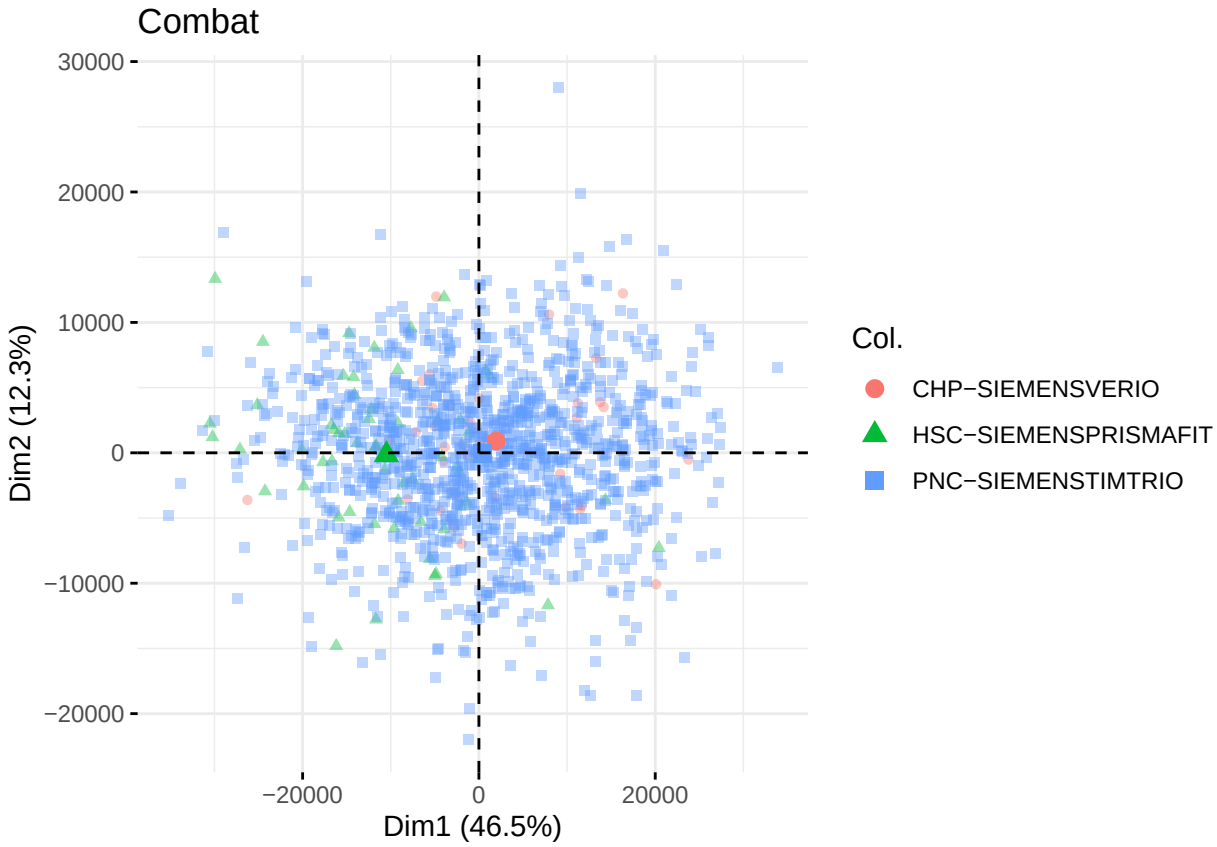

```
## [1] "Dim 1"
```

```
##       contr ROI_INDEX                         ROI_NAME HEMISPHERE TISSUE_SEG
## 1 25.972325        82            frontal lobe WM left       Left         WM
## 2 25.651379        81           frontal lobe WM right      Right         WM
## 3  7.495871        86           parietal lobe WM left       Left         WM
## 4  6.792212        87          temporal lobe WM right      Right         WM
## 5  6.519150        85          parietal lobe WM right      Right         WM
## 6  6.512898        88           temporal lobe WM left       Left         WM
## 7  1.736432       142 Right MFG   middle frontal gyrus      Right         GM
```

```
## [1] "Dim 2"
```

```
##       contr ROI_INDEX                         ROI_NAME HEMISPHERE TISSUE_SEG
## 1 43.019776        39         Left Cerebellum Exterior       Left         GM
## 2 41.814467        38        Right Cerebellum Exterior      Right         GM
## 3  1.734476        40    Right Cerebellum White Matter      Right         WM
## 4  1.721974        41     Left Cerebellum White Matter       Left         WM
## 5  1.366730        35                       Brain Stem       Both       NONE
## 6  1.199287       142 Right MFG   middle frontal gyrus      Right         GM
## 7  1.080577       143  Left MFG   middle frontal gyrus       Left         GM
```


Raw
20000
10000
0
−10000
−20000
Dim2 (10.4%)
−40000
0
40000
80000
Dim1 (70.4%)
Col.
CHP−SIEMENSVERIO
HSC−SIEMENSPRISMAFIT
PNC−SIEMENSTIMTRIO

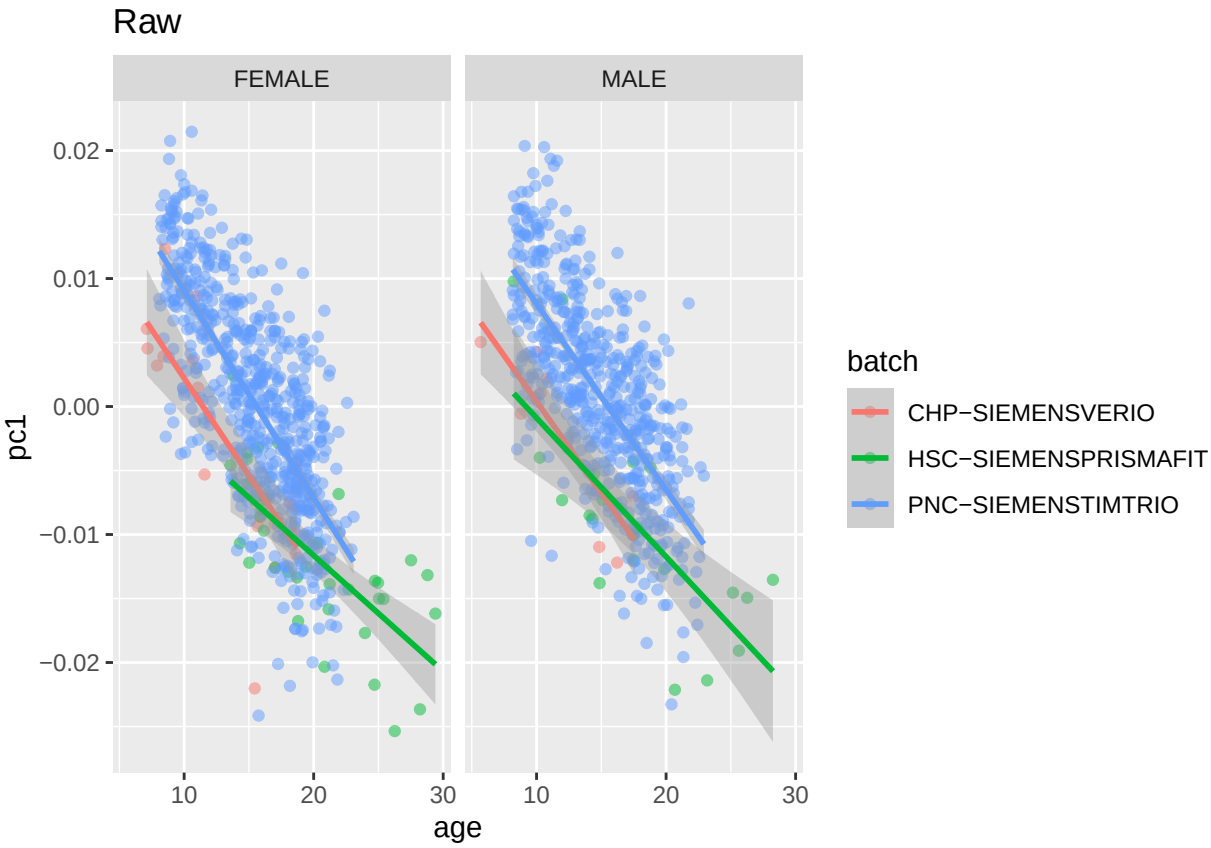

Combat
30000
20000
10000
0
−10000
−20000
Dim2 (12.3%)
−20000
0
20000
Dim1 (46.5%)
Col.
CHP−SIEMENSVERIO
HSC−SIEMENSPRISMAFIT
PNC−SIEMENSTIMTRIO

PC1

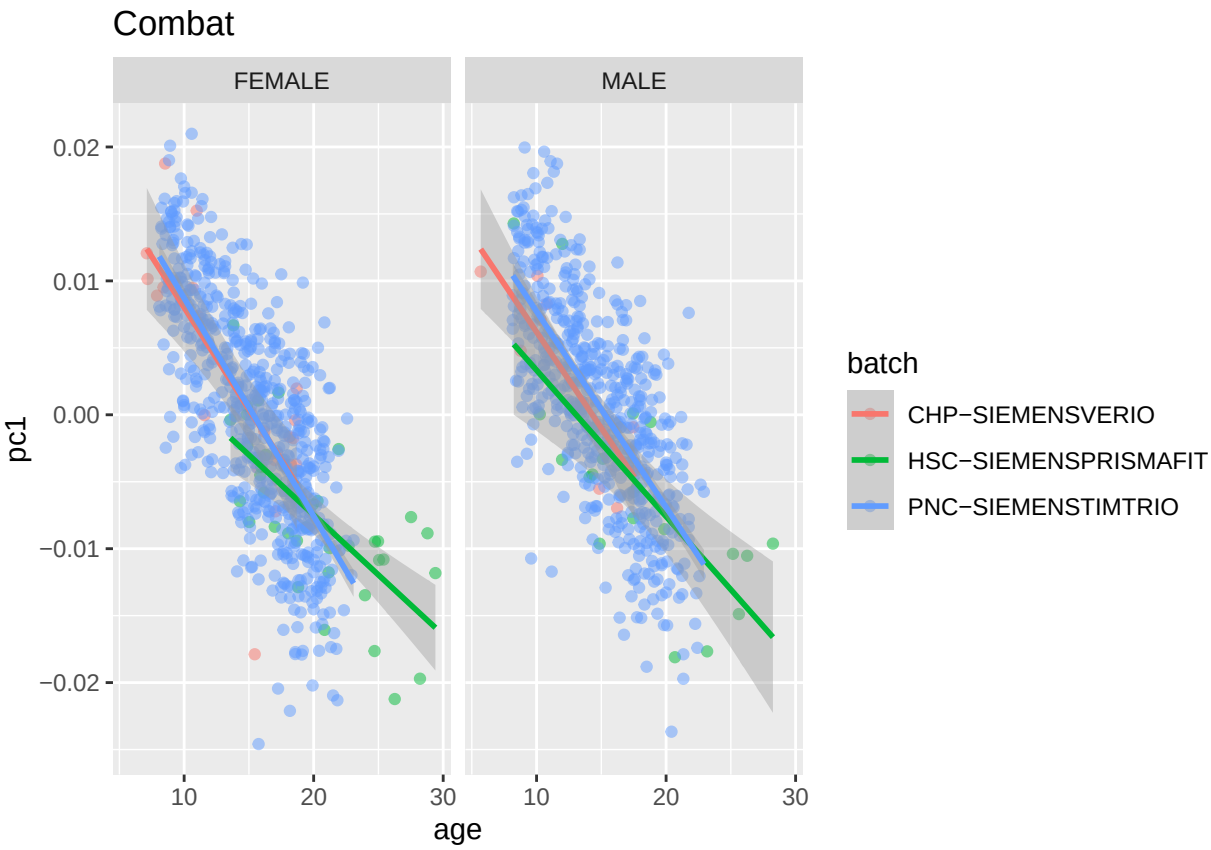
```
## [1] "RAW"
```

```
##
## Call:
## glm(formula = scale(pc1) ~ batch + scale(age) + sex, data = dat)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -2.86653  -0.44837   0.01187   0.47507   2.08436
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                -0.74183    0.11222  -6.610 5.63e-11 ***
## batchHSC-SIEMENSPRISMAFIT   0.16582    0.14423   1.150    0.250
## batchPNC-SIEMENSTIMTRIO     0.80613    0.11336   7.111 1.91e-12 ***
## scale(age)                 -0.70160    0.01928 -36.392  < 2e-16 ***
## sexMALE                    -0.02734    0.03773  -0.725    0.469
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.446651)
##
##     Null deviance: 1278.00  on 1278  degrees of freedom
## Residual deviance:  569.03  on 1274  degrees of freedom
## AIC: 2605.8
##
## Number of Fisher Scoring iterations: 2
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
## `geom_smooth()` using formula 'y ~ x'
```



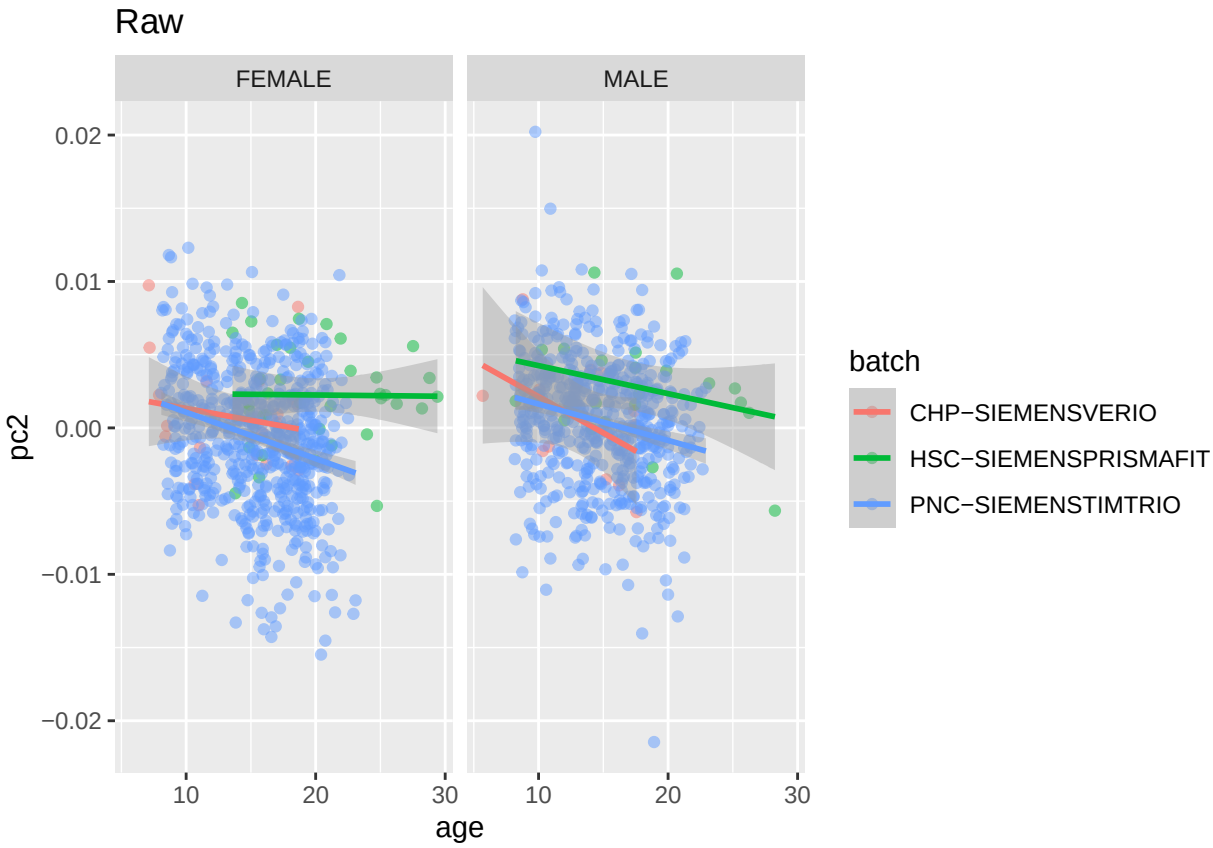

```
## [1] "COMBAT"

##
## Call:
## glm(formula = scale(pc1) ~ batch + scale(age) + sex, data = dat_combat)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -2.95890  -0.45886   0.01302   0.49080   2.14427
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                -0.05275    0.11498  -0.459    0.646
## batchHSC-SIEMENSPRISMAFIT -0.02594    0.14778  -0.176    0.861
## batchPNC-SIEMENSTIMTRIO    0.07025    0.11615   0.605    0.545
## scale(age)                 -0.72680    0.01975 -36.794   <2e-16 ***
## sexMALE                    -0.02427    0.03866  -0.628    0.530
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.4688875)
##
##     Null deviance: 1278.00  on 1278  degrees of freedom
## Residual deviance:  597.36  on 1274  degrees of freedom
## AIC: 2667.9
##
## Number of Fisher Scoring iterations: 2
```

PC2

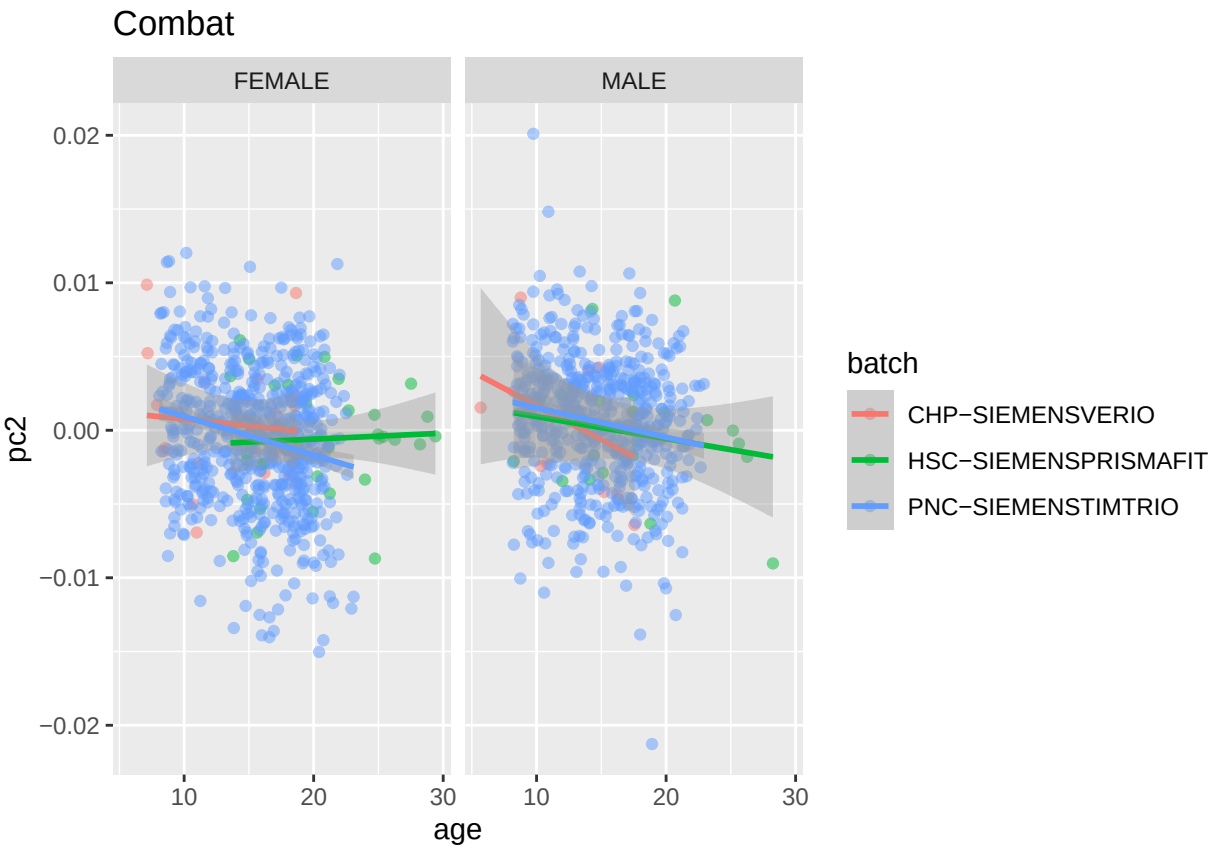
```
## [1] "RAW"
```

```
##
## Call:
## glm(formula = scale(pc2) ~ batch + scale(age) + sex, data = dat)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -4.5395  -0.6252   0.0682   0.6411   4.0426
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)               -0.007143   0.161632  -0.044 0.964756
## batchHSC-SIEMENSPRISMAFIT  0.705936   0.207732   3.398 0.000699 ***
## batchPNC-SIEMENSTIMTRIO   -0.118154   0.163270  -0.724 0.469400
## scale(age)                -0.229509   0.027767  -8.266 3.45e-16 ***
## sexMALE                    0.184021   0.054345   3.386 0.000730 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.926507)
##
##     Null deviance: 1278.0  on 1278  degrees of freedom
## Residual deviance: 1180.4  on 1274  degrees of freedom
## AIC: 3539
##
## Number of Fisher Scoring iterations: 2
```

```
## `geom_smooth()` using formula 'y ~ x'
```



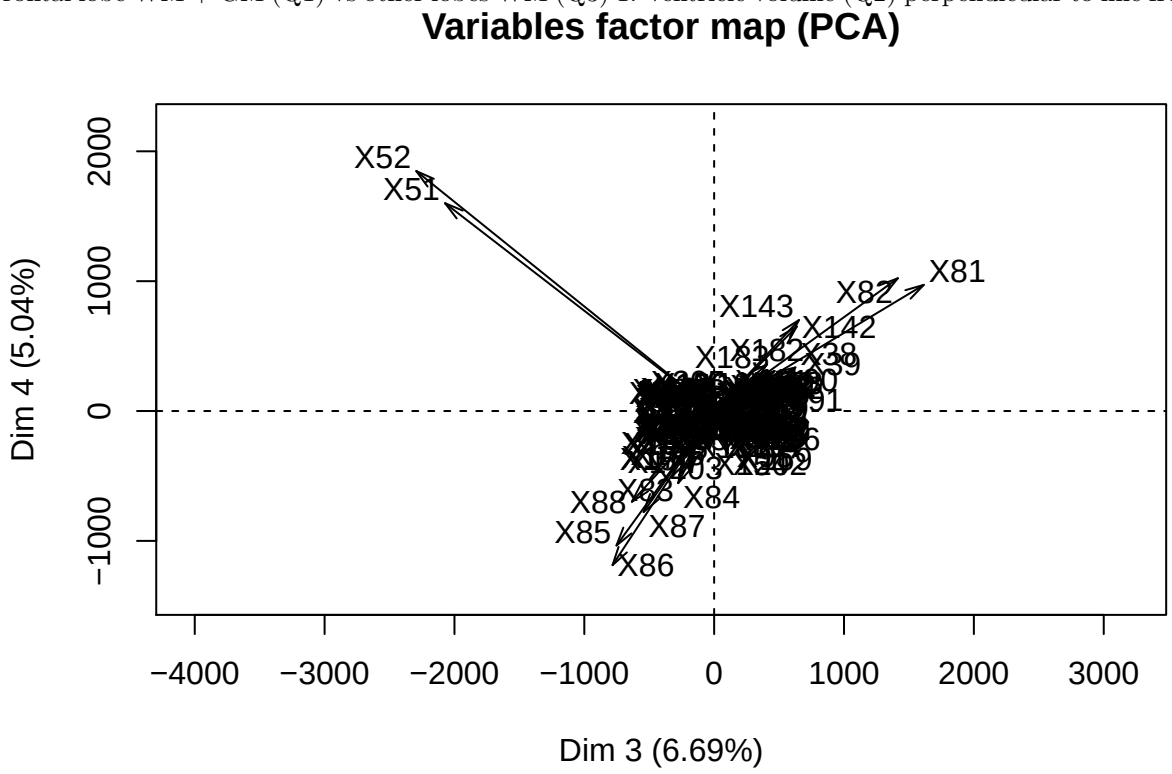

```
## `geom_smooth()` using formula 'y ~ x'
```




```
## [1] "COMBAT"

##
## Call:
## glm(formula = scale(pc2) ~ batch + scale(age) + sex, data = dat_combat)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -4.6064  -0.6336   0.0573   0.6559   4.0709
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                -0.05878    0.16434  -0.358 0.720671
## batchHSC-SIEMENSPRISMAFIT   0.07840    0.21121   0.371 0.710568
## batchPNC-SIEMENSTIMTRIO    -0.03190    0.16601  -0.192 0.847668
## scale(age)                 -0.18659    0.02823  -6.609 5.67e-11 ***
## sexMALE                     0.18439    0.05526   3.337 0.000871 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.9578326)
##
##     Null deviance: 1278.0  on 1278  degrees of freedom
## Residual deviance: 1220.3  on 1274  degrees of freedom
## AIC: 3581.5
##
## Number of Fisher Scoring iterations: 2
```

Components 3 and 4

Plane describes:
1. Frontal lobe WM + GM (Q1) vs other lobes WM (Q3) 1. ventricle volume (Q2) perpendicular to line from



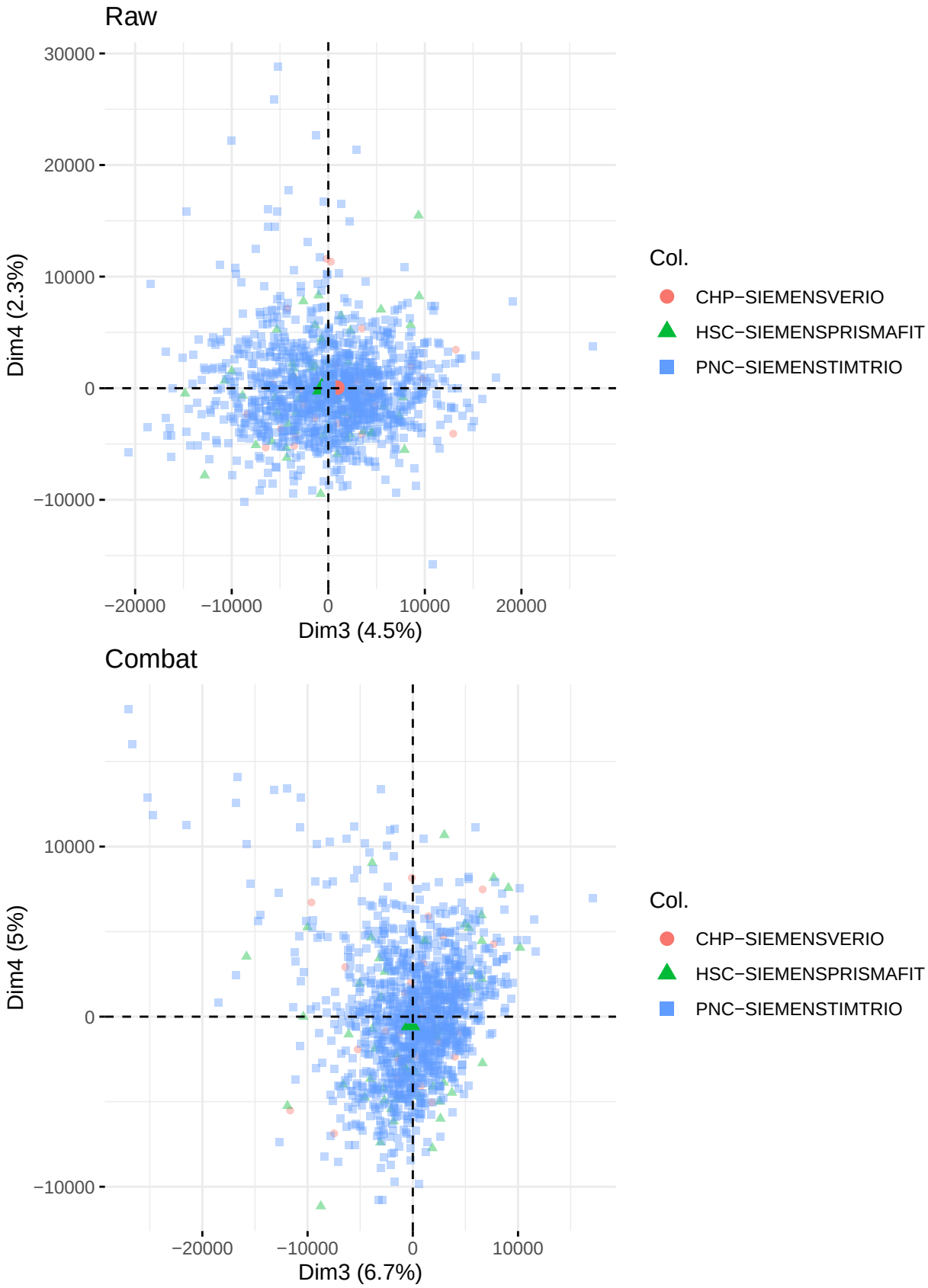
(1)

```
## [1] "Dim 3"
```

```
##       contr ROI_INDEX                ROI_NAME HEMISPHERE TISSUE_SEG
## 1 27.237835        52   Left Lateral Ventricle       Left         VN
## 2 22.227643        51  Right Lateral Ventricle      Right         VN
## 3 13.491099        81    frontal lobe WM right      Right         WM
## 4 10.361772        82     frontal lobe WM left       Left         WM
## 5  3.170683        86    parietal lobe WM left       Left         WM
## 6  2.928781        85   parietal lobe WM right      Right         WM
## 7  2.227252        39 Left Cerebellum Exterior       Left         GM
```

```
## [1] "Dim 4"
```

```
##       contr ROI_INDEX                ROI_NAME HEMISPHERE TISSUE_SEG
## 1 23.468697        52  Left Lateral Ventricle       Left         VN
## 2 17.605789        51 Right Lateral Ventricle      Right         VN
## 3  9.682115        86   parietal lobe WM left       Left         WM
## 4  7.370174        85  parietal lobe WM right      Right         WM
## 5  7.185209        82    frontal lobe WM left       Left         WM
## 6  6.486172        81   frontal lobe WM right      Right         WM
## 7  4.160532        87  temporal lobe WM right      Right         WM
```

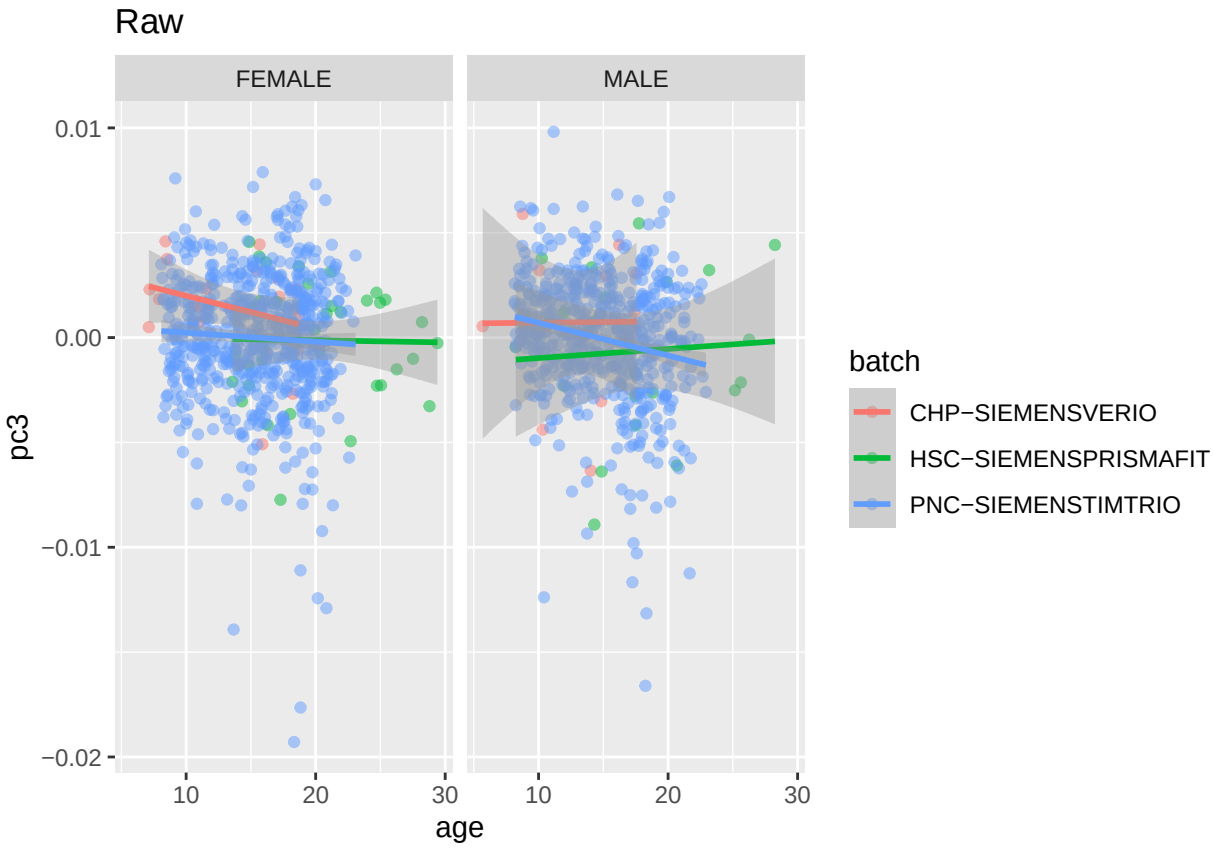

Raw
30000
20000
10000
0
−10000
Dim4 (2.3%)
−20000
−10000
0
10000
20000
Dim3 (4.5%)
Col.
CHP−SIEMENSVERIO
HSC−SIEMENSPRISMAFIT
PNC−SIEMENSTIMTRIO
Combat
10000
0
−10000
Dim4 (5%)
−20000
−10000
0
10000
Dim3 (6.7%)
Col.
CHP−SIEMENSVERIO
HSC−SIEMENSPRISMAFIT
PNC−SIEMENSTIMTRIO

PC3

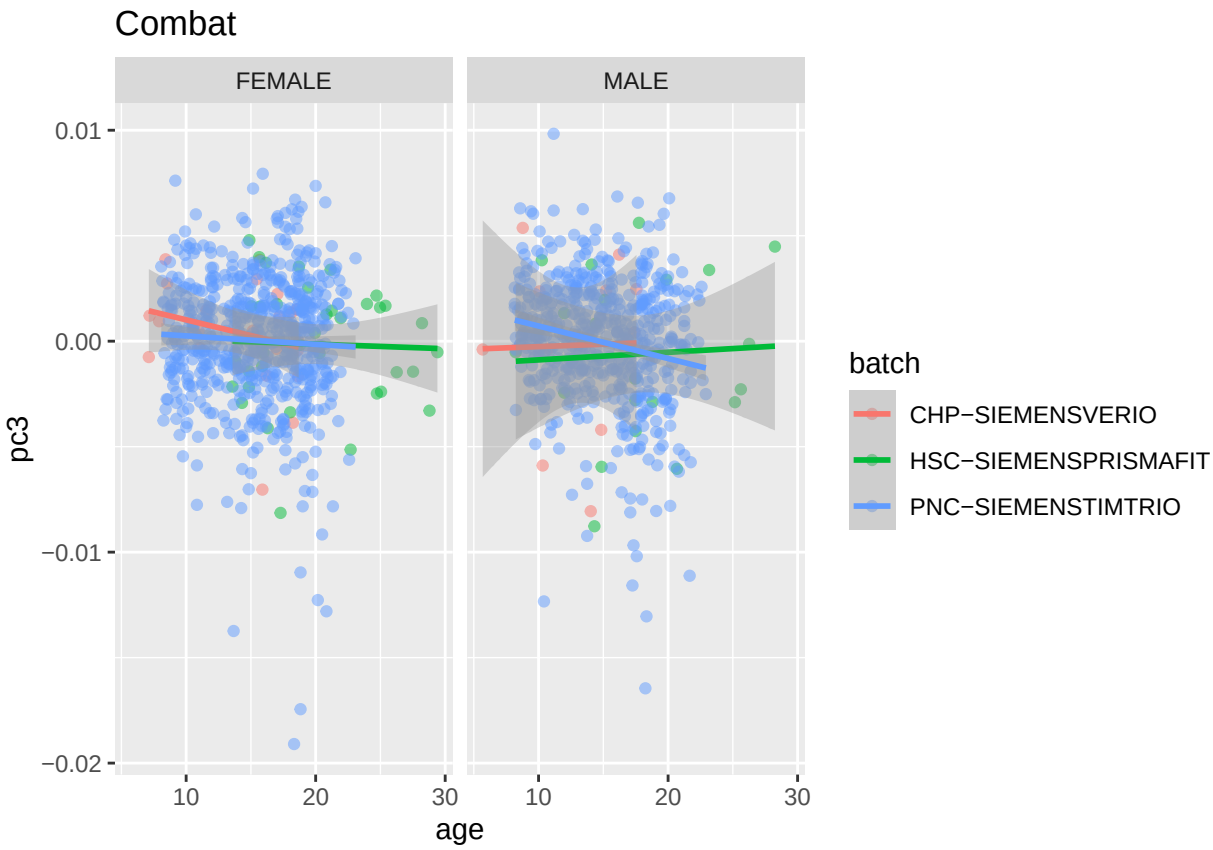
```
## [1] "RAW"
```

```
##
## Call:
## glm(formula = scale(pc3) ~ batch + scale(age) + sex, data = dat)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -6.2306  -0.5002   0.0580   0.6444   3.1264
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.34482    0.16682   2.067 0.038939 *
## batchHSC-SIEMENSPRISMAFIT -0.31886    0.21440  -1.487 0.137219
## batchPNC-SIEMENSTIMTRIO   -0.33459    0.16851  -1.986 0.047298 *
## scale(age)                -0.11115    0.02866  -3.878 0.000111 ***
## sexMALE                   -0.04428    0.05609  -0.790 0.429955
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.9869864)
##
##     Null deviance: 1278.0  on 1278  degrees of freedom
## Residual deviance: 1257.4  on 1274  degrees of freedom
## AIC: 3619.9
##
## Number of Fisher Scoring iterations: 2
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
## `geom_smooth()` using formula 'y ~ x'
```



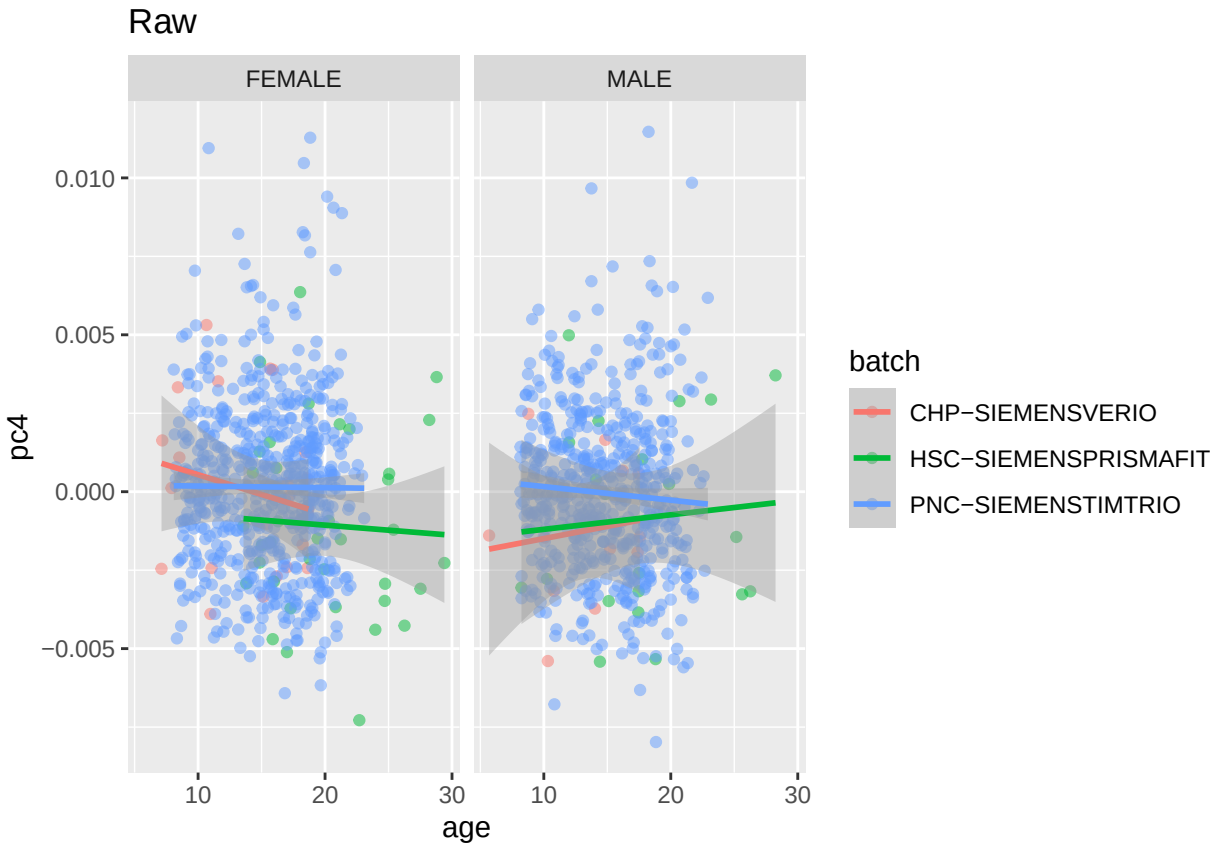

```
## [1] "COMBAT"

##
## Call:
## glm(formula = scale(pc3) ~ batch + scale(age) + sex, data = dat_combat)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -6.1891  -0.5008   0.0579   0.6458   3.1308
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.03966    0.16717   0.237 0.812526
## batchHSC-SIEMENSPRISMAFIT -0.01334    0.21485  -0.062 0.950499
## batchPNC-SIEMENSTIMTRIO   -0.01899    0.16886  -0.112 0.910465
## scale(age)                -0.10939    0.02872  -3.809 0.000146 ***
## sexMALE                   -0.04667    0.05621  -0.830 0.406551
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.9910885)
##
##     Null deviance: 1278.0  on 1278  degrees of freedom
## Residual deviance: 1262.6  on 1274  degrees of freedom
## AIC: 3625.2
##
## Number of Fisher Scoring iterations: 2
```

PC4

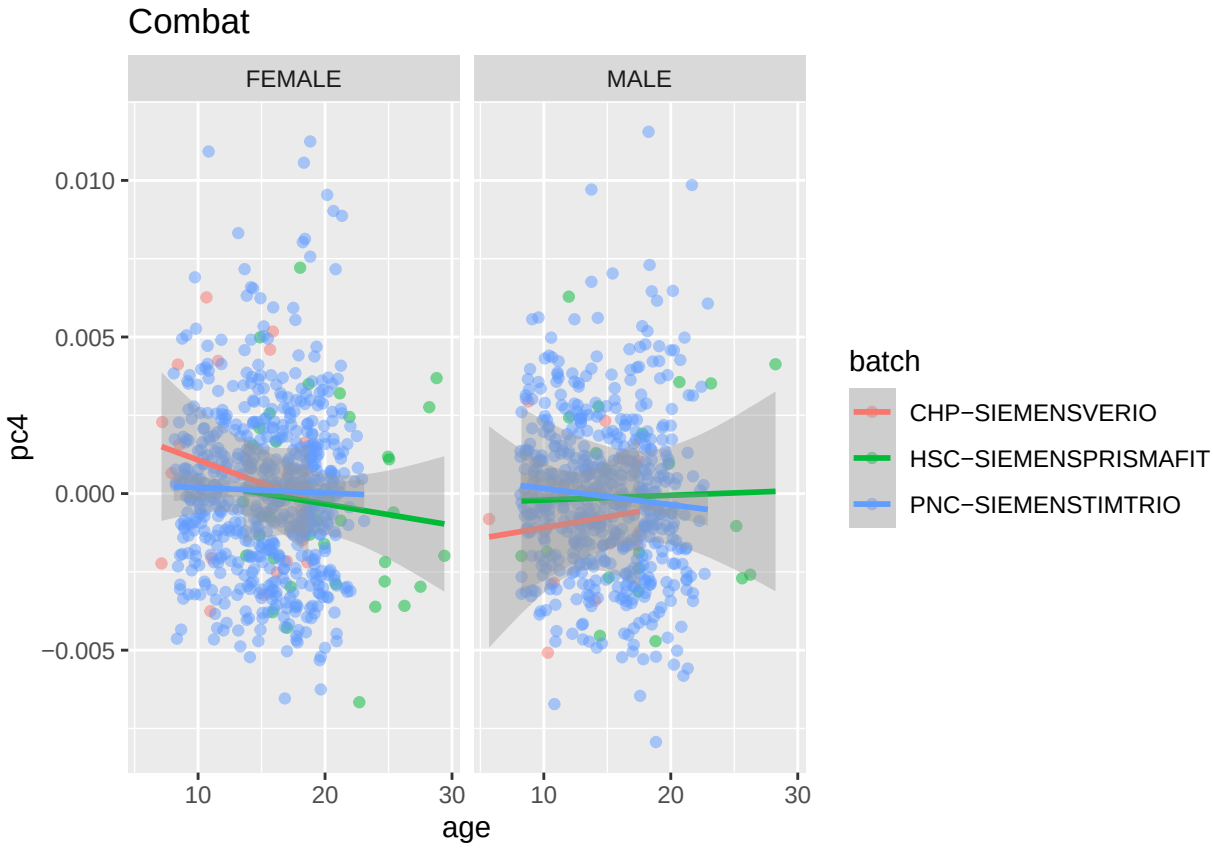
```
## [1] "RAW"
```

```
##
## Call:
## glm(formula = scale(pc4) ~ batch + scale(age) + sex, data = dat)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.9493  -0.6816  -0.0537   0.5775   4.3643
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                -0.13875    0.16735  -0.829    0.407
## batchHSC-SIEMENSPRISMAFIT -0.16793    0.21508  -0.781    0.435
## batchPNC-SIEMENSTIMTRIO    0.19916    0.16904   1.178    0.239
## scale(age)                 -0.03340    0.02875  -1.162    0.246
## sexMALE                    -0.08300    0.05627  -1.475    0.140
##
## (Dispersion parameter for gaussian family taken to be 0.9932038)
##
##     Null deviance: 1278.0  on 1278  degrees of freedom
## Residual deviance: 1265.3  on 1274  degrees of freedom
## AIC: 3627.9
##
## Number of Fisher Scoring iterations: 2
```

```
## `geom_smooth()` using formula 'y ~ x'
```

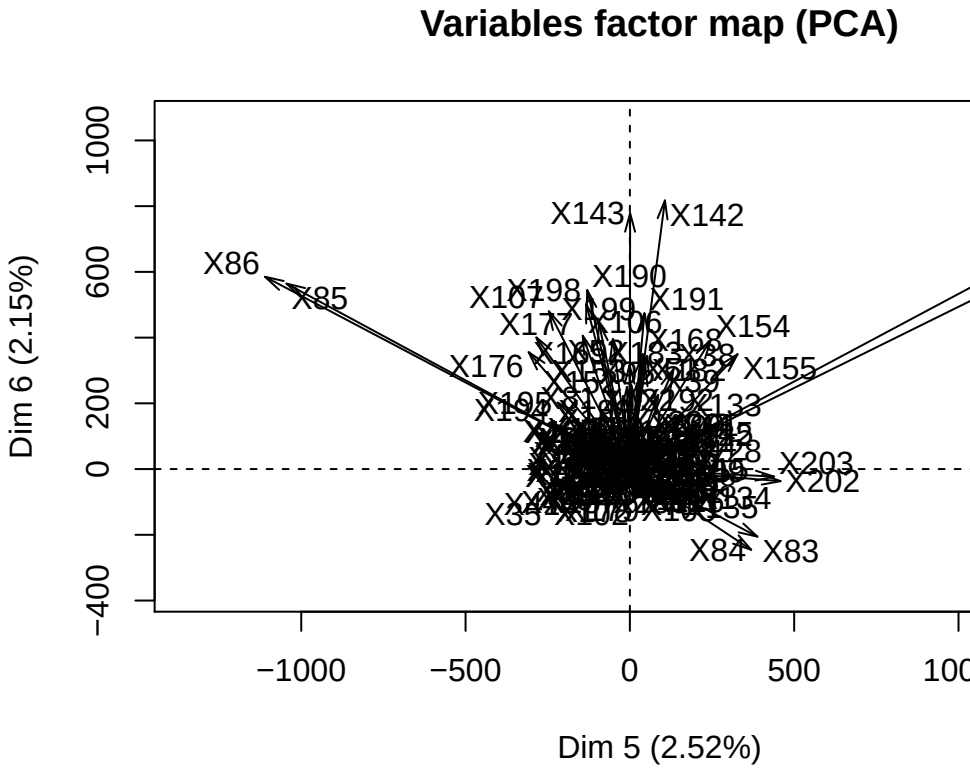



```
## `geom_smooth()` using formula 'y ~ x'
```




```
## [1] "COMBAT"

##
## Call:
## glm(formula = scale(pc4) ~ batch + scale(age) + sex, data = dat_combat)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.9100  -0.6817  -0.0616   0.5715   4.4470
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.01066    0.16782   0.064   0.9494
## batchHSC-SIEMENSPRISMAFIT -0.02203    0.21569  -0.102   0.9187
## batchPNC-SIEMENSTIMTRIO    0.03065    0.16952   0.181   0.8565
## scale(age)                -0.05188    0.02883  -1.800   0.0722 .
## sexMALE                   -0.08278    0.05643  -1.467   0.1426
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.998844)
##
##     Null deviance: 1278.0  on 1278  degrees of freedom
## Residual deviance: 1272.5  on 1274  degrees of freedom
## AIC: 3635.2
##
## Number of Fisher Scoring iterations: 2
```

Components 5 and 6



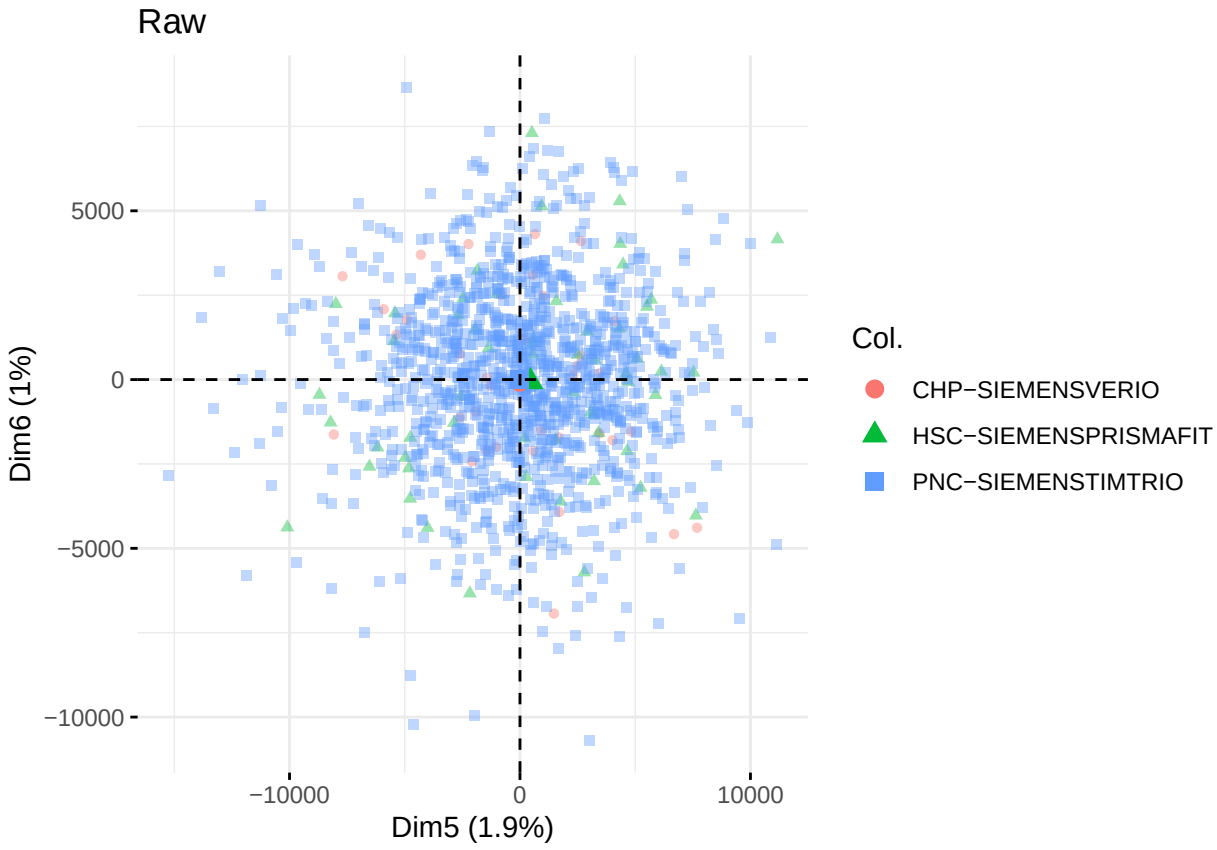
PC5: Temporal vs Parietal Lobe PC6: ?

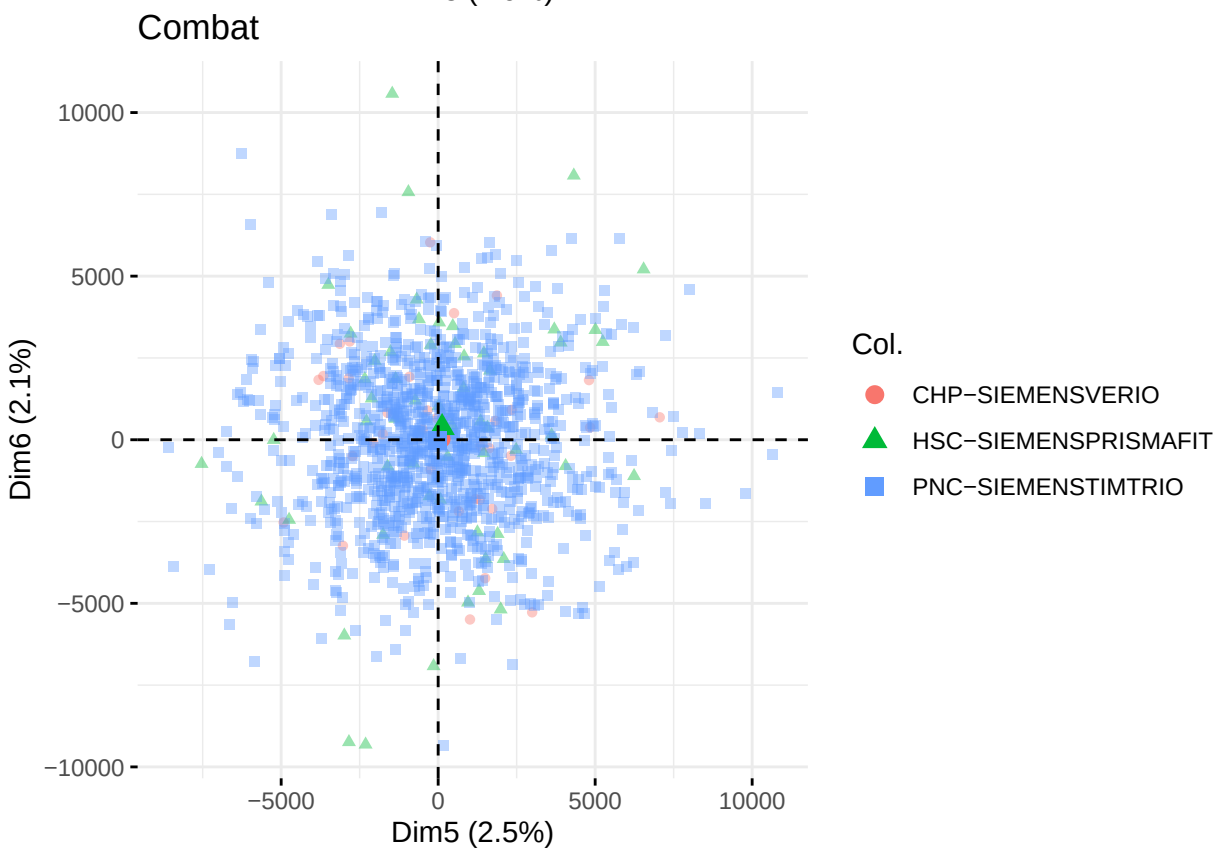
```
## [1] "Dim 5"
```

```
##       contr ROI_INDEX                 ROI_NAME HEMISPHERE TISSUE_SEG
## 1 21.856790        87    temporal lobe WM right      Right         WM
## 2 21.613837        88     temporal lobe WM left       Left         WM
## 3 16.943638        86     parietal lobe WM left       Left         WM
## 4 15.012096        85    parietal lobe WM right      Right         WM
## 5  2.890344       202 Right TMP   temporal pole      Right         GM
## 6  2.649432       203  Left TMP   temporal pole       Left         GM
## 7  2.068942        83   occipital lobe WM right      Right         WM
```

```
## [1] "Dim 6"
```

```
##       contr ROI_INDEX                          ROI_NAME HEMISPHERE TISSUE_SEG
## 1 10.783205       142   Right MFG   middle frontal gyrus      Right         GM
## 2  9.783965       143    Left MFG   middle frontal gyrus       Left         GM
## 3  7.290804        87             temporal lobe WM right      Right         WM
## 4  6.189053        88              temporal lobe WM left       Left         WM
## 5  5.526742        86              parietal lobe WM left       Left         WM
## 6  5.140681        85             parietal lobe WM right      Right         WM
## 7  4.785437       190 Right SFG   superior frontal gyrus      Right         GM
```


Raw
5000
0
−5000
−10000
Dim6 (1%)
−10000
0
10000
Dim5 (1.9%)
Col.
CHP−SIEMENSVERIO
HSC−SIEMENSPRISMAFIT
PNC−SIEMENSTIMTRIO

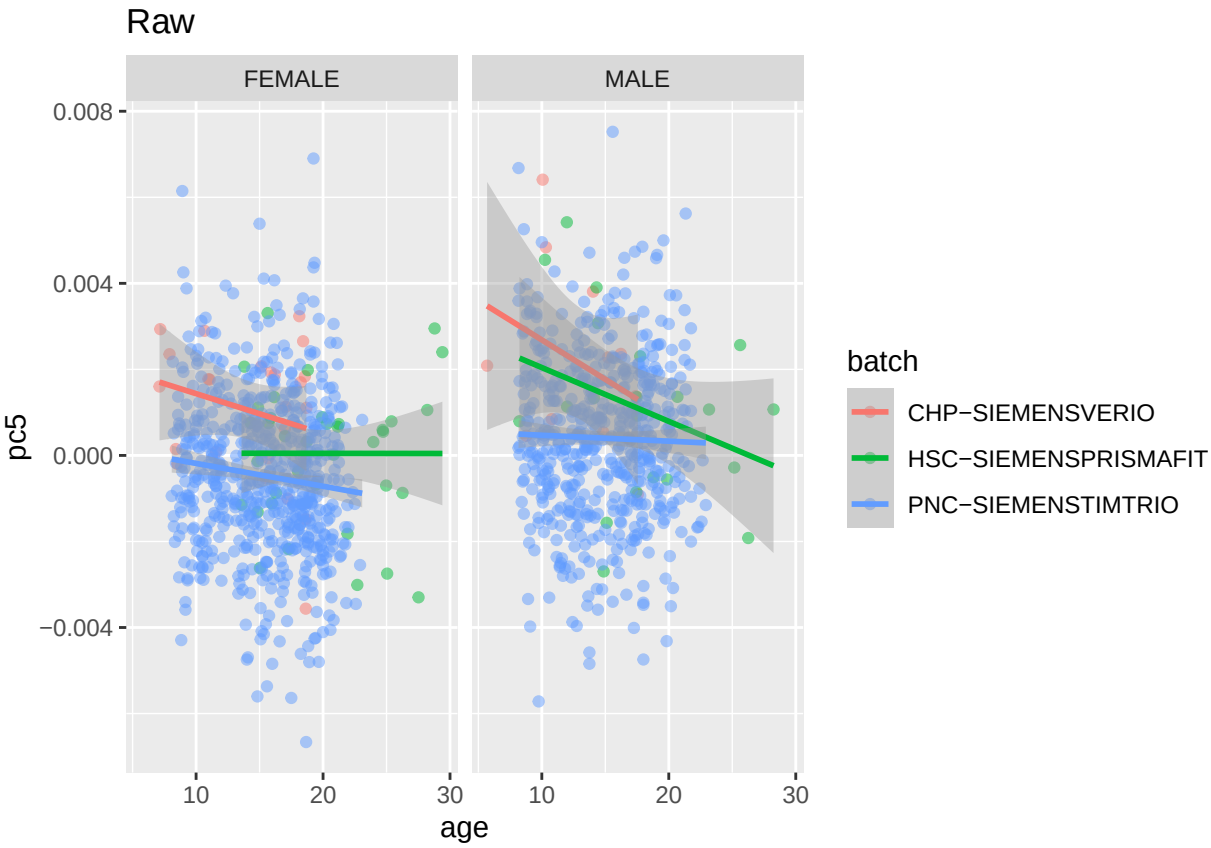

Combat
10000
5000
0
−5000
−10000
Dim6 (2.1%)
−5000
0
5000
10000
Dim5 (2.5%)
Col.
CHP−SIEMENSVERIO
HSC−SIEMENSPRISMAFIT
PNC−SIEMENSTIMTRIO

PC5

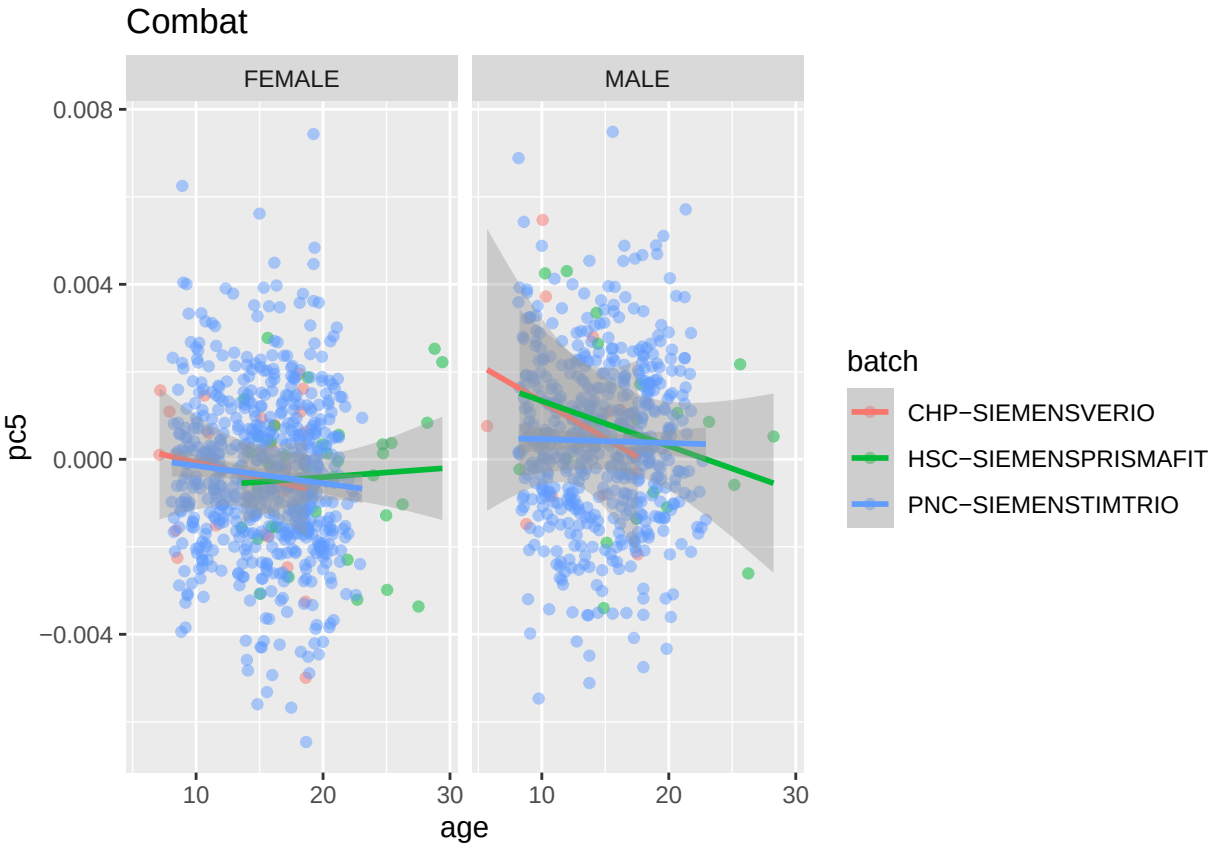
```
## [1] "RAW"
```

```
##
## Call:
## glm(formula = scale(pc5) ~ batch + scale(age) + sex, data = dat)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.2810  -0.6360  -0.0219   0.6247   3.9072
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.54554    0.16164   3.375  0.00076 ***
## batchHSC-SIEMENSPRISMAFIT -0.40954    0.20775  -1.971  0.04890 *
## batchPNC-SIEMENSTIMTRIO   -0.79011    0.16328  -4.839 1.46e-06 ***
## scale(age)                -0.07948    0.02777  -2.862  0.00428 **
## sexMALE                    0.44607    0.05435   8.208 5.47e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.9266395)
##
##     Null deviance: 1278.0  on 1278  degrees of freedom
## Residual deviance: 1180.5  on 1274  degrees of freedom
## AIC: 3539.2
##
## Number of Fisher Scoring iterations: 2
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
## `geom_smooth()` using formula 'y ~ x'
```



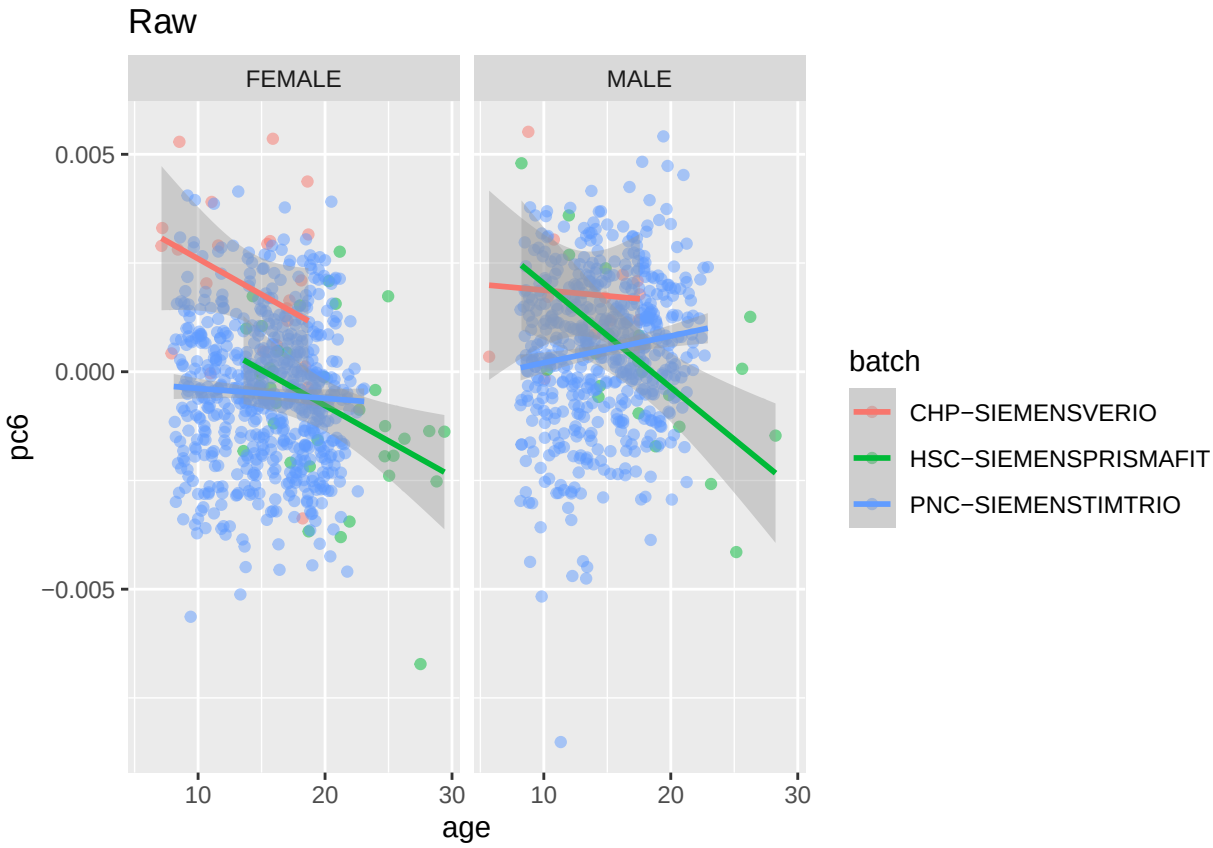

```
## [1] "COMBAT"

##
## Call:
## glm(formula = scale(pc5) ~ batch + scale(age) + sex, data = dat_combat)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.1570  -0.6461  -0.0133   0.6080   4.1388
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                -0.14652    0.16423  -0.892   0.3725
## batchHSC-SIEMENSPRISMAFIT   0.02305    0.21108   0.109   0.9131
## batchPNC-SIEMENSTIMTRIO    -0.04499    0.16590  -0.271   0.7863
## scale(age)                 -0.05802    0.02821  -2.057   0.0399 *
## sexMALE                     0.40709    0.05522   7.372 3.01e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.956572)
##
##     Null deviance: 1278.0  on 1278  degrees of freedom
## Residual deviance: 1218.7  on 1274  degrees of freedom
## AIC: 3579.8
##
## Number of Fisher Scoring iterations: 2
```

PC6

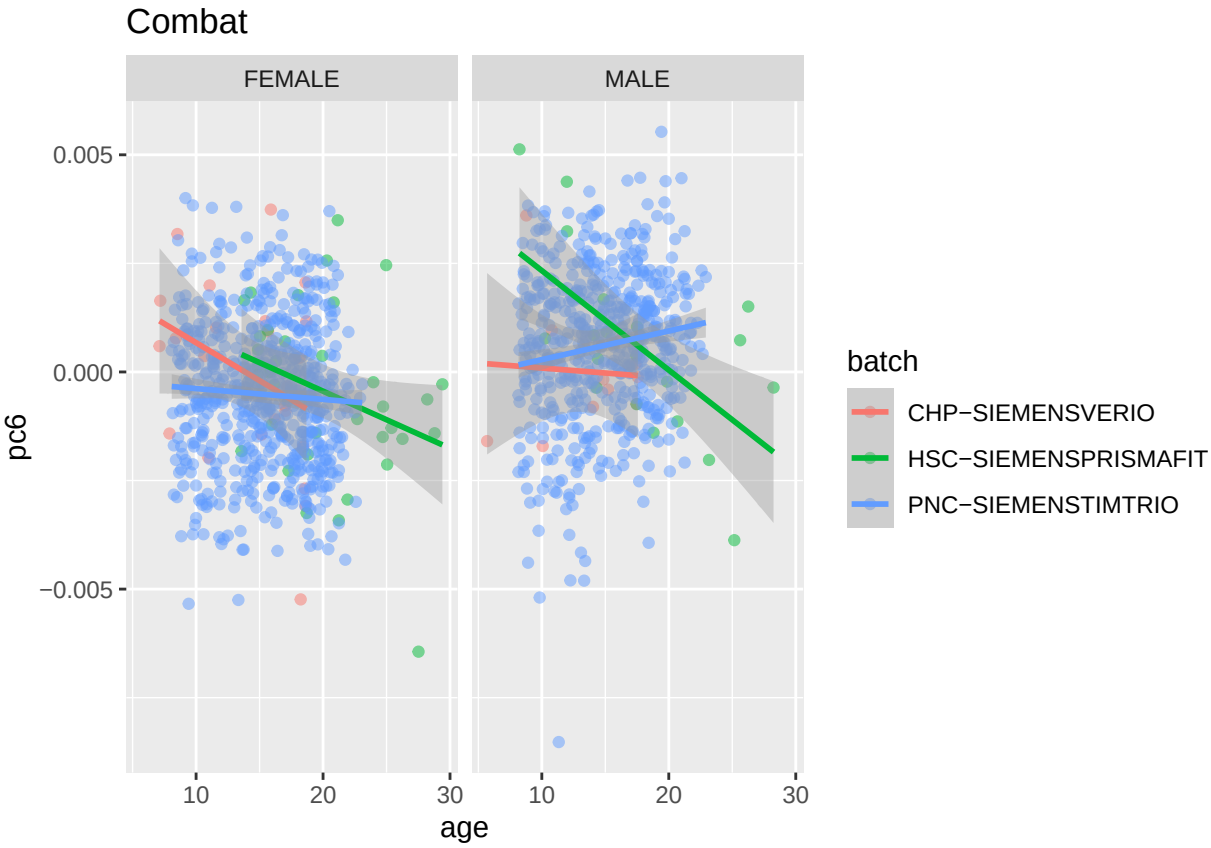
```
## [1] "RAW"
```

```
##
## Call:
## glm(formula = scale(pc6) ~ batch + scale(age) + sex, data = dat)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -4.9646  -0.6189   0.0290   0.6324   2.7241
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.843452   0.159153   5.300 1.37e-07 ***
## batchHSC-SIEMENSPRISMAFIT -1.272553   0.204545  -6.221 6.67e-10 ***
## batchPNC-SIEMENSTIMTRIO   -1.114422   0.160765  -6.932 6.58e-12 ***
## scale(age)                -0.008908   0.027341  -0.326    0.745
## sexMALE                    0.536773   0.053511  10.031  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.8982965)
##
##     Null deviance: 1278.0  on 1278  degrees of freedom
## Residual deviance: 1144.4  on 1274  degrees of freedom
## AIC: 3499.5
##
## Number of Fisher Scoring iterations: 2
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
## `geom_smooth()` using formula 'y ~ x'
```




```
## [1] "COMBAT"

##
## Call:
## glm(formula = scale(pc6) ~ batch + scale(age) + sex, data = dat_combat)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -5.1047  -0.6036   0.0306   0.6208   2.7757
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                -0.227376   0.160269  -1.419    0.156
## batchHSC-SIEMENSPRISMAFIT -0.039197   0.205979  -0.190    0.849
## batchPNC-SIEMENSTIMTRIO   -0.054389   0.161893  -0.336    0.737
## scale(age)                 -0.003702   0.027532  -0.134    0.893
## sexMALE                     0.608059   0.053886  11.284   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.9109415)
##
##     Null deviance: 1278.0  on 1278  degrees of freedom
## Residual deviance: 1160.5  on 1274  degrees of freedom
## AIC: 3517.3
##
## Number of Fisher Scoring iterations: 2
```