

model.rmd

Yanran Li, Yijin Wang, Shubo Zhang

2024-03-20

```
library(mice)
```

```
##
```

```
## Attaching package: 'mice'
```

```
## The following object is masked from 'package:stats':
```

```
##
```

```
##      filter
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      cbind, rbind
```

```
library(tidyr)
```

```
library(dplyr)
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
##      filter, lag
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      intersect, setdiff, setequal, union
```

```
library(visdat)
```

```
library(naniar)
```

```
library(ggplot2)
```

```
library(lme4)
```

```
## Loading required package: Matrix
```

```
## Warning: package 'Matrix' was built under R version 4.2.3
```

```
##
```

```
## Attaching package: 'Matrix'
```

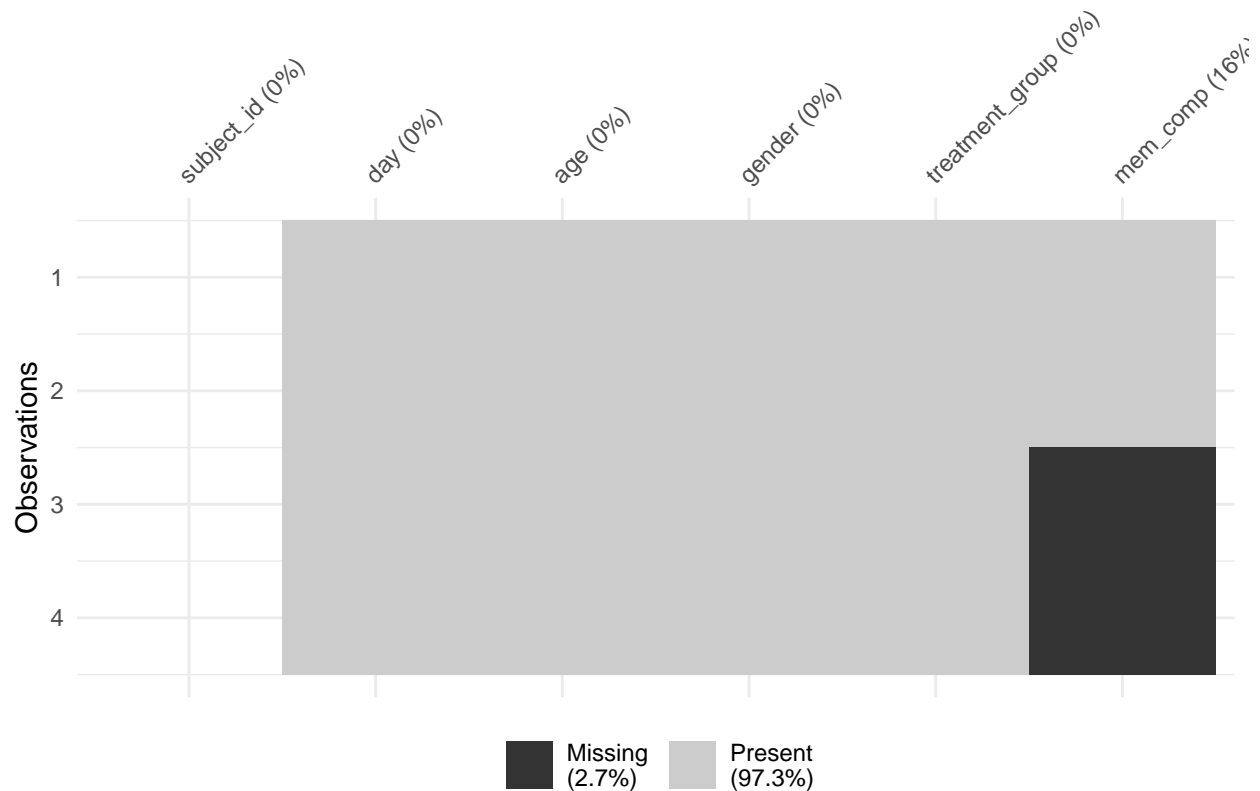
```
## The following objects are masked from 'package:tidyr':
##
##   expand, pack, unpack
```

```
library(broom.mixed)
```

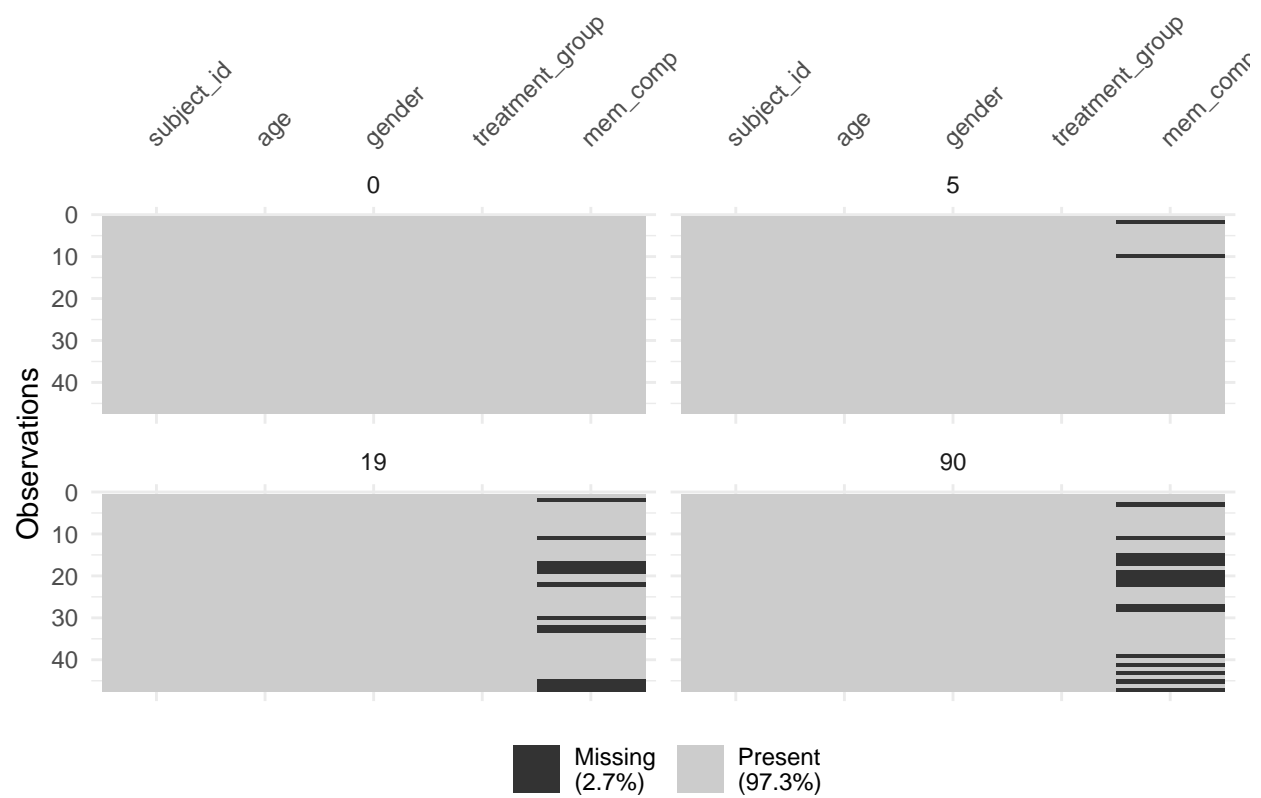
```
data <- read.csv("data.csv") %>%
  janitor::clean_names() %>%
  mutate(day = as.factor(day))
```

```
# pre-processing: expand design matrix and fill in NAs
full_data <- data %>%
  complete(day,
            nesting(subject_id, treatment_group)) %>%
  relocate(day, .after = subject_id) %>%
  relocate(treatment_group, .after = gender) %>%
  group_by(subject_id) %>%
  fill(age, gender, .direction = "downup") %>%
  arrange(subject_id, day)
```

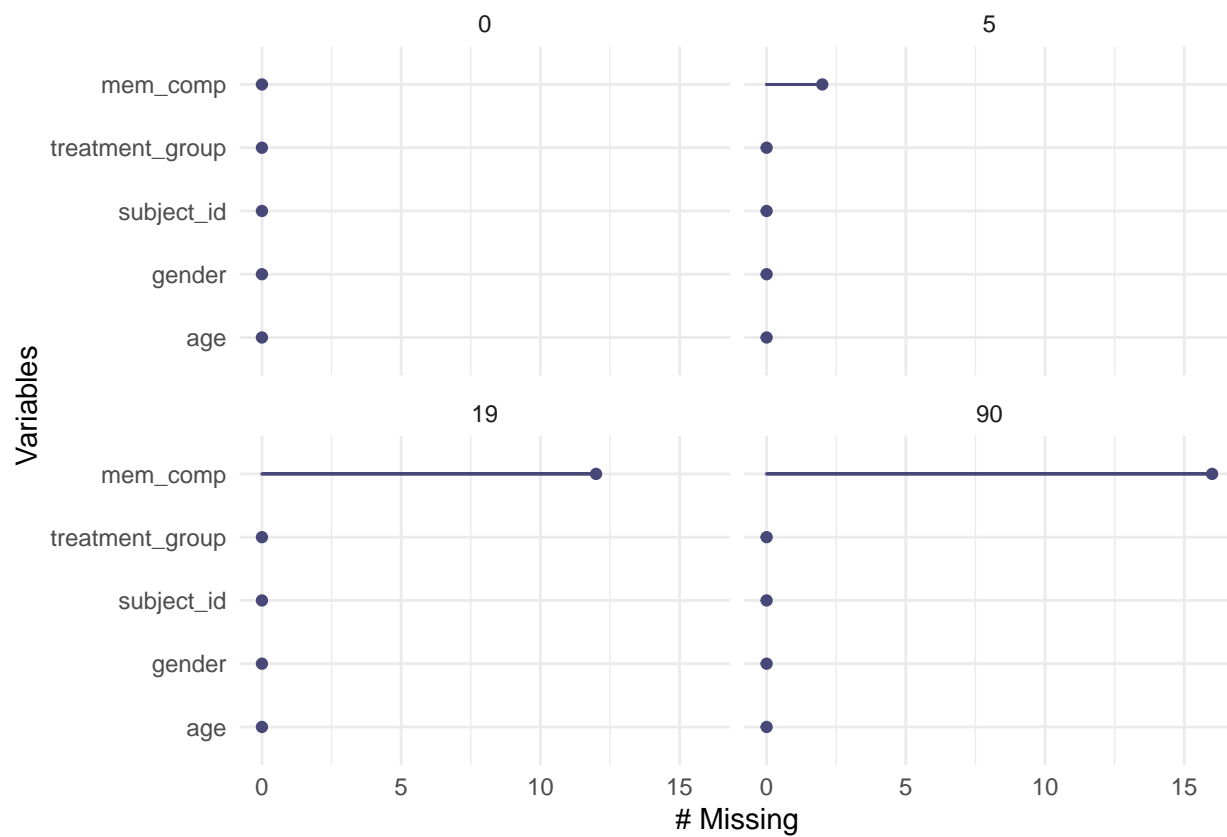
```
vis_miss(full_data)
```



```
vis_miss(full_data, facet = day)
```

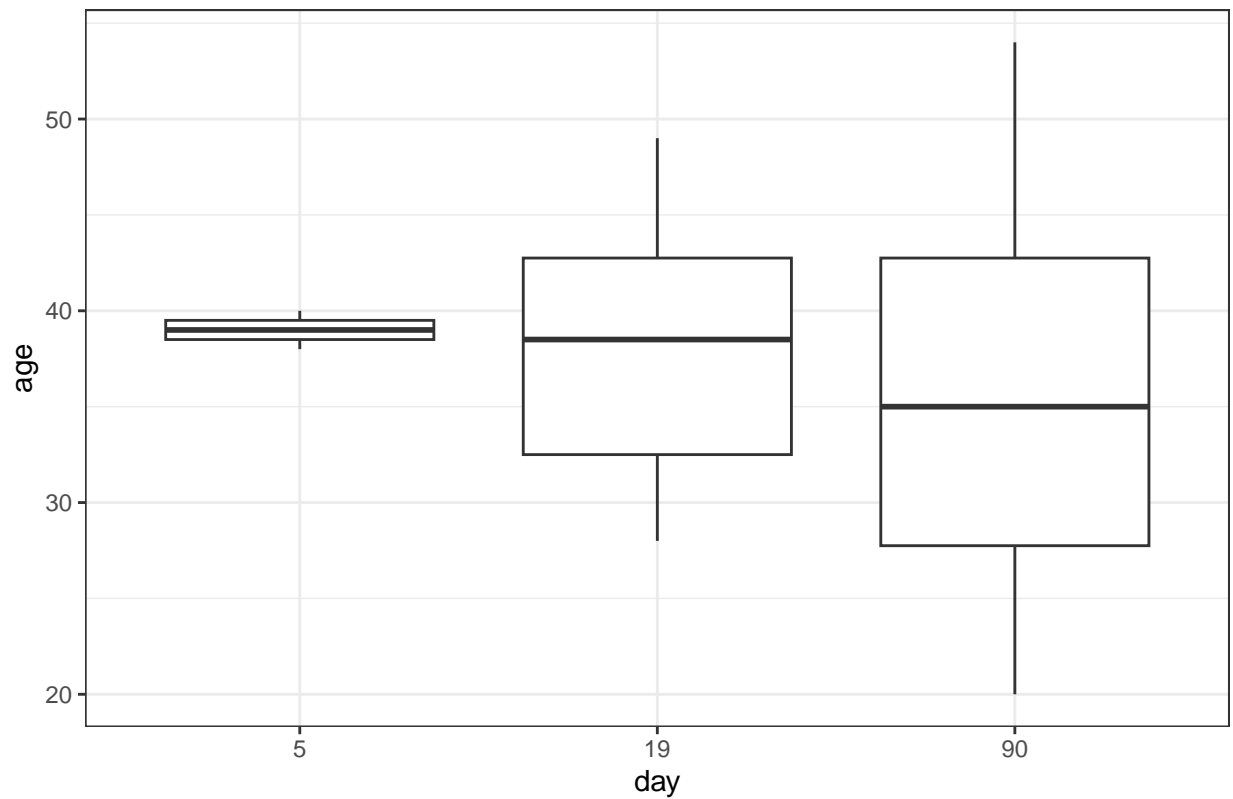


```
gg_miss_var(full_data,
  facet = day)
```



```
full_data %>% filter(is.na(mem_comp)) %>%
  ggplot(aes(x=as.factor(day)))+
  geom_boxplot(aes(y=age))+
  theme_bw()+
  xlab("day")+
  ggtitle("Relationship of age and missing values")
```

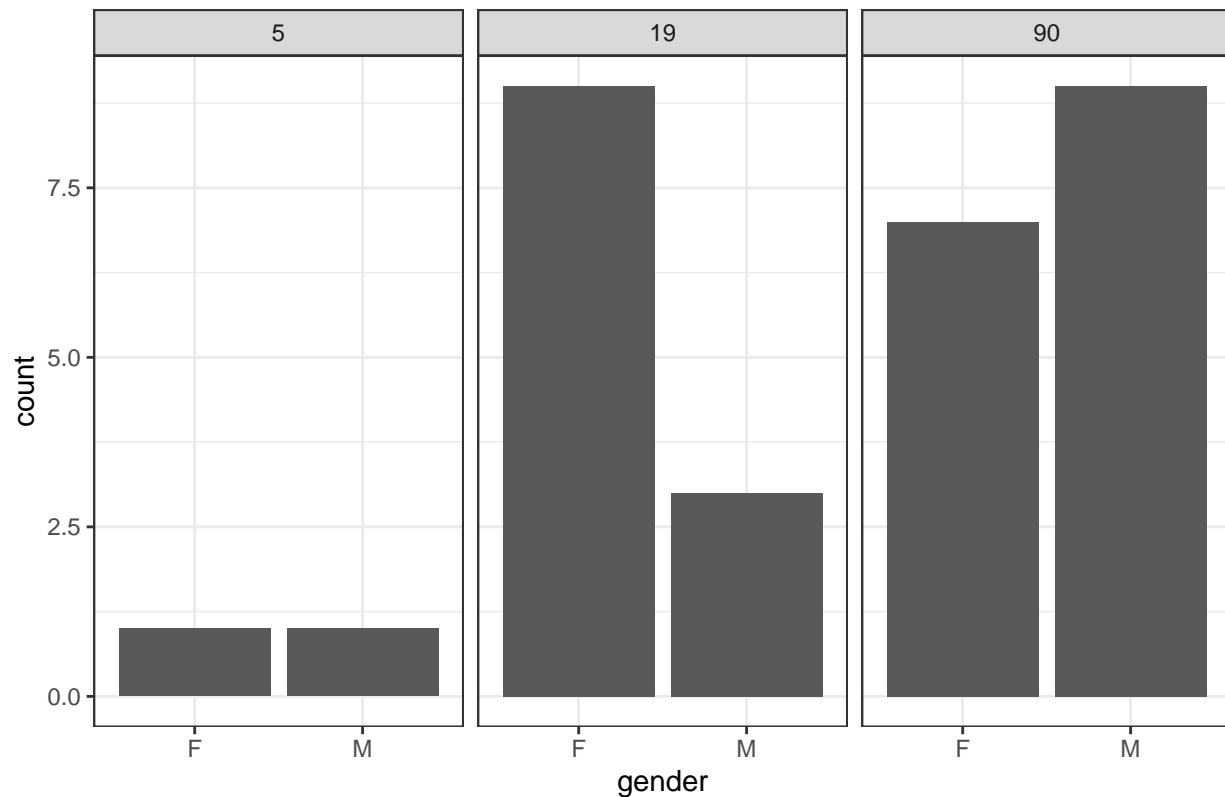
Relationship of age and missing values



```
full_data %>% filter(is.na(mem_comp)) %>%  
  ggplot()+  
  geom_bar(aes(x=gender, stat = "count"))+  
  facet_wrap(day~.)+  
  theme_bw()+  
  xlab("gender")+  
  ggtitle("Relationship of gender and missing values")
```

```
## Warning in geom_bar(aes(x = gender, stat = "count")): Ignoring unknown  
## aesthetics: stat
```

Relationship of gender and missing values



```
# assume MAR
# multiple imputations
imputed_data <- mice::mice(data = full_data, m = 6, seed = 2024, print = FALSE)
```

```
## Warning: Number of logged events: 2
```

```
imputed_data
```

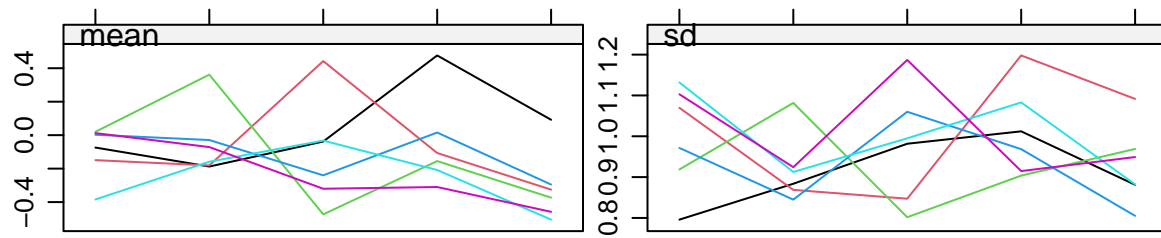
```
## Class: mids
## Number of multiple imputations: 6
## Imputation methods:
##   subject_id      day      age      gender treatment_group
##           ""           ""           ""           ""           ""
##   mem_comp
##   "pmm"
## PredictorMatrix:
##           subject_id day age gender treatment_group mem_comp
## subject_id      0  1  1  0      0      1
## day             1  0  1  0      0      1
## age             1  1  0  0      0      1
## gender          1  1  1  0      0      1
## treatment_group 1  1  1  0      0      1
## mem_comp        1  1  1  0      0      0
## Number of logged events: 2
##   it im dep      meth      out
```

```
## 1 0 0    constant    gender
## 2 0 0    constant treatment_group
```

```
# used pmm method
```

```
# check convergence
```

```
plot(imputed_data)
```

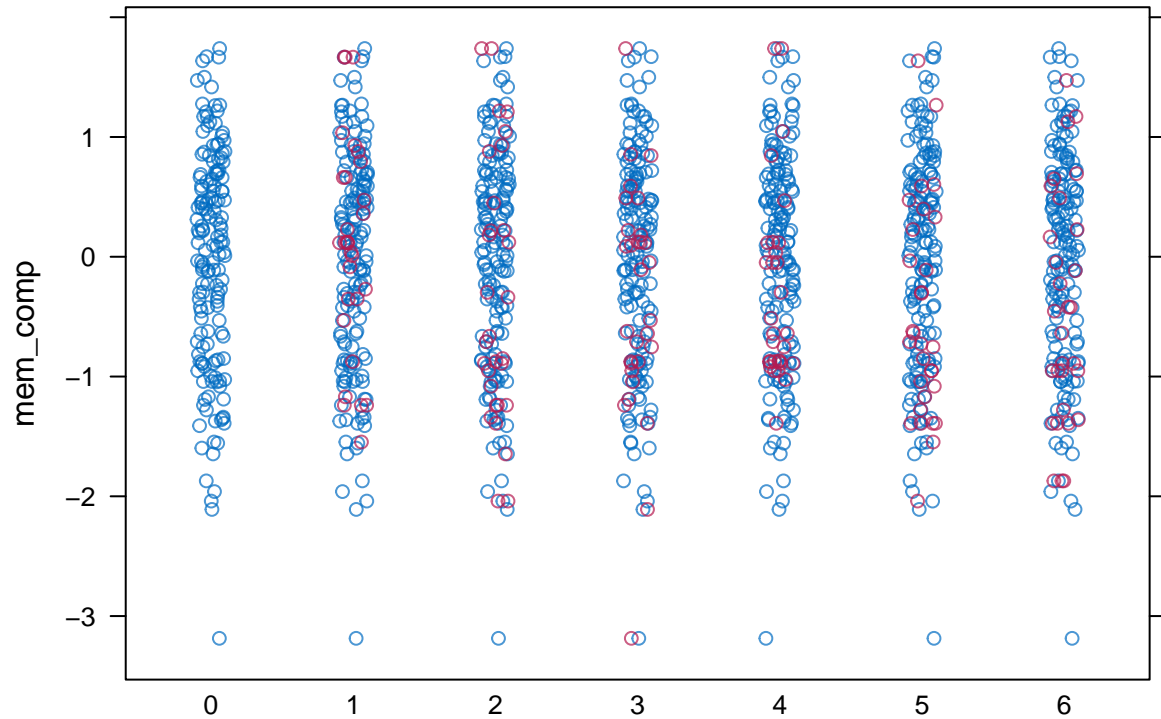


Iteration

```
# check if imputed data makes sense (ie. in a reasonable bound)
```

```
# red is imputed, blue is observed
```

```
mice::stripplot(imputed_data, mem_comp)
```



```
# fit model to each imputed datasets
```

```
model <- with(imputed_data, lmer(mem_comp ~ age + gender + treatment_group+day + +treatment_group:day
```

```
pool.fit <- pool(model)
```

```
summary(pool.fit)
```

```
##           term      estimate std.error statistic      df
## 1      (Intercept)  0.61259148 0.58498731  1.0471876 142.40918
## 2             age -0.01435861 0.01378527 -1.0415912 125.55072
## 3          genderM  0.07118804 0.25591736  0.2781681  88.04616
## 4  treatment_groupB -0.07762312 0.34501125 -0.2249872 169.48896
## 5  treatment_groupC -0.18796858 0.33052182 -0.5687025 169.82702
## 6           day5 -0.09577277 0.20913497 -0.4579472 170.01722
## 7           day19 -0.29378401 0.21785002 -1.3485609 131.38638
## 8           day90 -0.32331261 0.32115763 -1.0067100  12.46684
## 9  treatment_groupB:day5 -0.06276973 0.32782242 -0.1914748 117.93381
## 10 treatment_groupC:day5  0.34998520 0.30725190  1.1390823 152.67981
## 11 treatment_groupB:day19 0.09954652 0.36771610  0.2707157  41.12777
## 12 treatment_groupC:day19 0.57107906 0.31251365  1.8273732 132.74876
## 13 treatment_groupB:day90 0.12771299 0.40146403  0.3181181  24.01220
## 14 treatment_groupC:day90 0.31860668 0.34760696  0.9165716  48.25563
##      p.value
## 1 0.29678712
```



```
## 2 0.29960283
## 3 0.78153553
## 4 0.82226025
## 5 0.57030935
## 6 0.64757513
## 7 0.17979804
## 8 0.33321463
## 9 0.84848291
## 10 0.25645285
## 11 0.78796372
## 12 0.06988959
## 13 0.75314553
## 14 0.36392529
```

```
# check the pooled model covariance
# get the full matrix ubar (instead of only the diagonal)
m <- pool.fit$m
ubar <- Reduce("+", lapply(model$analyses, vcov)) / (m)
b <- pool.fit$pooled$b # this one is still provided by mice

## or by hand as well
# qbar <- getqbar(fil.pooled) # pooled estimates
# b <- 1 / (m-1) * rowSums((sapply(fit.mi$analyses, coef) - qbar)^2)

t <- ubar + (1 + 1 / (m)) * b # this is t as it used to be

# check versus the diagonal of t that is still provided
all.equal(as.numeric(diag(t)), pool.fit$pooled$t) # check
```

```
## [1] TRUE
```

```
# pool coef estimates and standard errors to create 95% CIs
```

```
# sensitivity analysis for each data missing assumptions
# mcar - fit data with completers data
completers_data <- full_data %>%
  group_by(subject_id) %>%
  mutate(missing = any(is.na(mem_comp))) %>%
  ungroup()
completers_fit <- lmer(mem_comp ~ age + gender + treatment_group+day +treatment_group:day + (1|subject_id),
  data=completers_data)
summary(completers_fit)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula:
## mem_comp ~ age + gender + treatment_group + day + treatment_group:day +
## (1 | subject_id)
## Data: completers_data %>% filter(missing == FALSE)
##
## REML criterion at convergence: 173.8
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
```

```
## -1.7688 -0.5756 0.1020 0.6102 1.7294
##
## Random effects:
##   Groups      Name      Variance Std.Dev.
##   subject_id (Intercept) 0.6790 0.8240
##   Residual      0.1597 0.3997
## Number of obs: 96, groups: subject_id, 24
##
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept)    -0.2777878 0.9983716 -0.278
## age              0.0004539 0.0227681 0.020
## genderM          0.2159847 0.3696430 0.584
## treatment_groupB 0.0579773 0.4876764 0.119
## treatment_groupC -0.0553875 0.4461694 -0.124
## day5            -0.2121626 0.1998282 -1.062
## day19           -0.3476050 0.1998282 -1.740
## day90           -0.2289847 0.1998282 -1.146
## treatment_groupB:day5 0.0978170 0.2925186 0.334
## treatment_groupC:day5 0.3899459 0.2746377 1.420
## treatment_groupB:day19 0.3457121 0.2925186 1.182
## treatment_groupC:day19 0.6461374 0.2746377 2.353
## treatment_groupB:day90 0.3018569 0.2925186 1.032
## treatment_groupC:day90 0.2610949 0.2746377 0.951

##
## Correlation matrix not shown by default, as p = 14 > 12.
## Use print(x, correlation=TRUE) or
##      vcov(x)      if you need it
```

```
vcov(completers_fit)
```

```
## 14 x 14 Matrix of class "dpoMatrix"
##               (Intercept)      age      genderM
## (Intercept)    0.996745819 -2.088791e-02 1.689317e-03
## age            -0.020887911 5.183885e-04 -2.037118e-03
## genderM         0.001689317 -2.037118e-03 1.366360e-01
## treatment_groupB -0.212922483 2.575381e-03 -3.229577e-03
## treatment_groupC -0.127542401 6.464675e-04 -7.900043e-03
## day5            -0.019965664 -6.496339e-20 -3.859519e-20
## day19           -0.019965664 -7.395764e-20 3.317334e-20
## day90           -0.019965664 -5.414075e-20 3.555508e-20
## treatment_groupB:day5 0.019965664 1.503083e-19 -5.229356e-20
## treatment_groupC:day5 0.019965664 5.632924e-20 1.435933e-19
## treatment_groupB:day19 0.019965664 1.465875e-19 -1.137372e-19
## treatment_groupC:day19 0.019965664 8.320791e-20 3.355296e-35
## treatment_groupB:day90 0.019965664 1.073095e-19 -1.219031e-19
## treatment_groupC:day90 0.019965664 6.091242e-20 0.000000e+00
##               treatment_groupB treatment_groupC      day5
## (Intercept)    -0.212922483    -0.1275424010 -1.996566e-02
## age              0.002575381      0.0006464675 -6.496339e-20
## genderM         -0.003229577    -0.0079000425 -3.859519e-20
## treatment_groupB 0.237828299      0.1077680039 1.996566e-02
```

```

## treatment_groupC      0.107768004      0.1990671216  1.996566e-02
## day5                   0.019965664      0.0199656643  3.993133e-02
## day19                  0.019965664      0.0199656643  1.996566e-02
## day90                  0.019965664      0.0199656643  1.996566e-02
## treatment_groupB:day5 -0.042783566     -0.0199656643 -3.993133e-02
## treatment_groupC:day5 -0.019965664     -0.0377129214 -3.993133e-02
## treatment_groupB:day19 -0.042783566     -0.0199656643 -1.996566e-02
## treatment_groupC:day19 -0.019965664     -0.0377129214 -1.996566e-02
## treatment_groupB:day90 -0.042783566     -0.0199656643 -1.996566e-02
## treatment_groupC:day90 -0.019965664     -0.0377129214 -1.996566e-02
##                               day19      day90 treatment_groupB:day5
## (Intercept)          -1.996566e-02 -1.996566e-02      1.996566e-02
## age                  -7.395764e-20 -5.414075e-20      1.503083e-19
## genderM              3.317334e-20  3.555508e-20     -5.229356e-20
## treatment_groupB      1.996566e-02  1.996566e-02     -4.278357e-02
## treatment_groupC      1.996566e-02  1.996566e-02     -1.996566e-02
## day5                  1.996566e-02  1.996566e-02     -3.993133e-02
## day19                 3.993133e-02  1.996566e-02     -1.996566e-02
## day90                 1.996566e-02  3.993133e-02     -1.996566e-02
## treatment_groupB:day5 -1.996566e-02 -1.996566e-02      8.556713e-02
## treatment_groupC:day5 -1.996566e-02 -1.996566e-02      3.993133e-02
## treatment_groupB:day19 -3.993133e-02 -1.996566e-02      4.278357e-02
## treatment_groupC:day19 -3.993133e-02 -1.996566e-02      1.996566e-02
## treatment_groupB:day90 -1.996566e-02 -3.993133e-02      4.278357e-02
## treatment_groupC:day90 -1.996566e-02 -3.993133e-02      1.996566e-02
##                               treatment_groupC:day5 treatment_groupB:day19
## (Intercept)          1.996566e-02      1.996566e-02
## age                   5.632924e-20      1.465875e-19
## genderM              1.435933e-19     -1.137372e-19
## treatment_groupB     -1.996566e-02     -4.278357e-02
## treatment_groupC     -3.771292e-02     -1.996566e-02
## day5                 -3.993133e-02     -1.996566e-02
## day19                -1.996566e-02     -3.993133e-02
## day90                -1.996566e-02     -1.996566e-02
## treatment_groupB:day5 3.993133e-02      4.278357e-02
## treatment_groupC:day5 7.542584e-02      1.996566e-02
## treatment_groupB:day19 1.996566e-02      8.556713e-02
## treatment_groupC:day19 3.771292e-02      3.993133e-02
## treatment_groupB:day90 1.996566e-02      4.278357e-02
## treatment_groupC:day90 3.771292e-02      1.996566e-02
##                               treatment_groupC:day19 treatment_groupB:day90
## (Intercept)          1.996566e-02      1.996566e-02
## age                   8.320791e-20      1.073095e-19
## genderM              3.355296e-35     -1.219031e-19
## treatment_groupB     -1.996566e-02     -4.278357e-02
## treatment_groupC     -3.771292e-02     -1.996566e-02
## day5                 -1.996566e-02     -1.996566e-02
## day19                -3.993133e-02     -1.996566e-02
## day90                -1.996566e-02     -3.993133e-02
## treatment_groupB:day5 1.996566e-02      4.278357e-02
## treatment_groupC:day5 3.771292e-02      1.996566e-02
## treatment_groupB:day19 3.993133e-02      4.278357e-02
## treatment_groupC:day19 7.542584e-02      1.996566e-02
## treatment_groupB:day90 1.996566e-02      8.556713e-02

```

```
## treatment_groupC:day90      3.771292e-02      3.993133e-02
##      treatment_groupC:day90
## (Intercept)      1.996566e-02
## age      6.091242e-20
## genderM      0.000000e+00
## treatment_groupB      -1.996566e-02
## treatment_groupC      -3.771292e-02
## day5      -1.996566e-02
## day19      -1.996566e-02
## day90      -3.993133e-02
## treatment_groupB:day5      1.996566e-02
## treatment_groupC:day5      3.771292e-02
## treatment_groupB:day19      1.996566e-02
## treatment_groupC:day19      3.771292e-02
## treatment_groupB:day90      3.993133e-02
## treatment_groupC:day90      7.542584e-02
```

```
# calculate CIs
```

```
# mnar - fill na's with different delta --> fit model --> CI's
```

```
# Create a delta vector that represent the following adjustment values for mmHg: 0 for MAR, and -0.7, -
delta <- c(0, -0.7, -0.2, 0.2, 0.7 )
```

```
imp.all <- vector("list", length(delta))
post <- imputed_data$post
for (i in 1:length(delta)){
  d <- delta[i]
  cmd <- paste("imp[[j]][,i] <- imp[[j]][,i] +", d)
  post["mem_comp"] <- cmd
  imp <- mice::mice(full_data, post = post, m = 6, seed = i, print = FALSE)
  imp.all[[i]] <- imp
}
```

```
## Warning: Number of logged events: 2
```

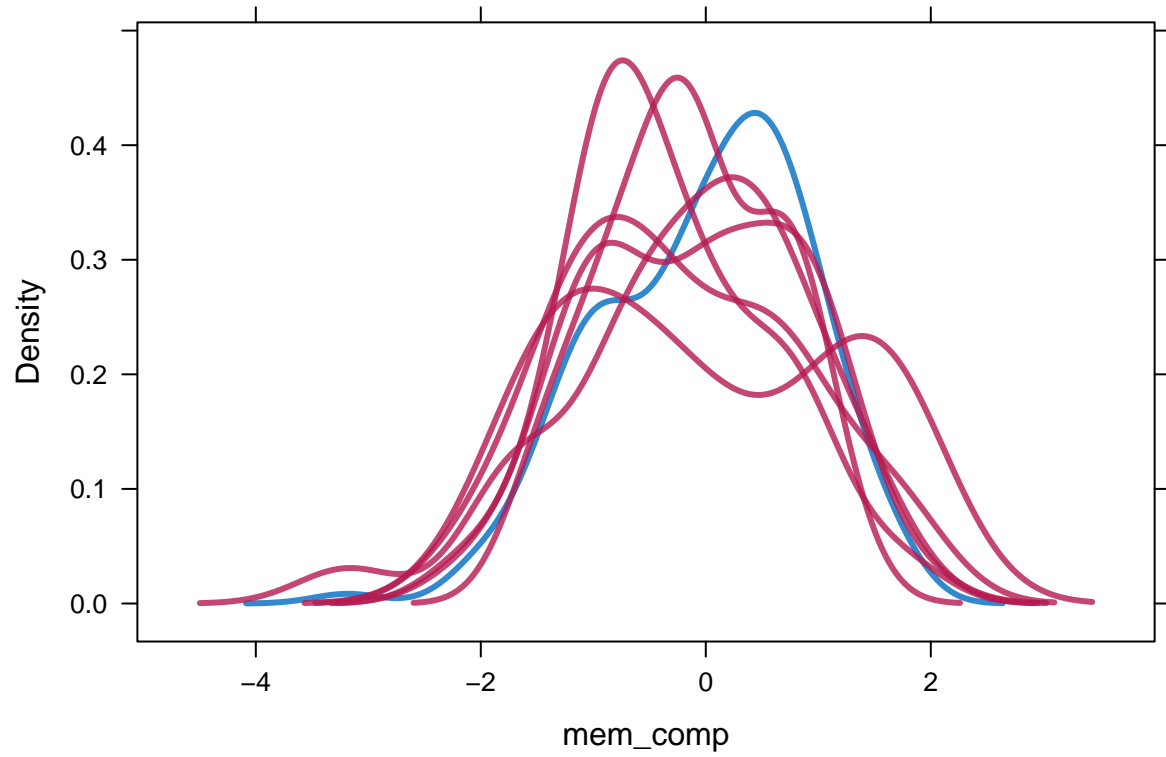
```
## Warning: Number of logged events: 2
```

```
## Warning: Number of logged events: 2
```

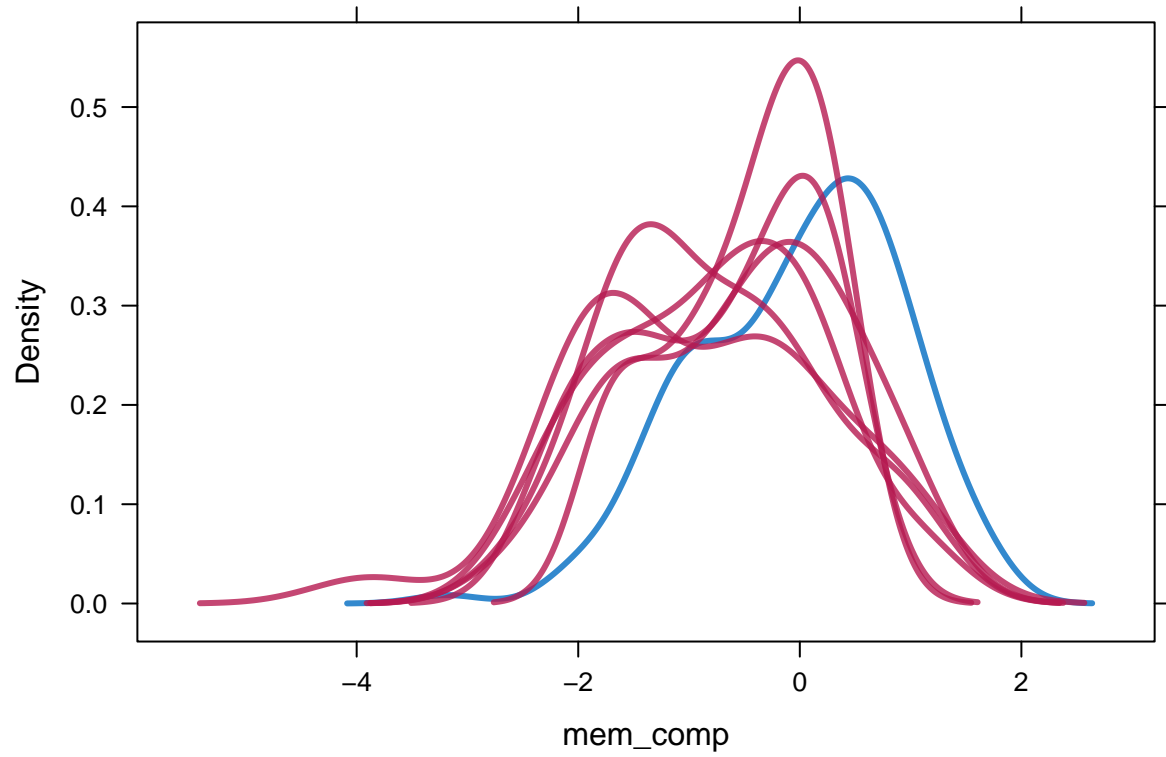
```
## Warning: Number of logged events: 2
```

```
## Warning: Number of logged events: 2
```

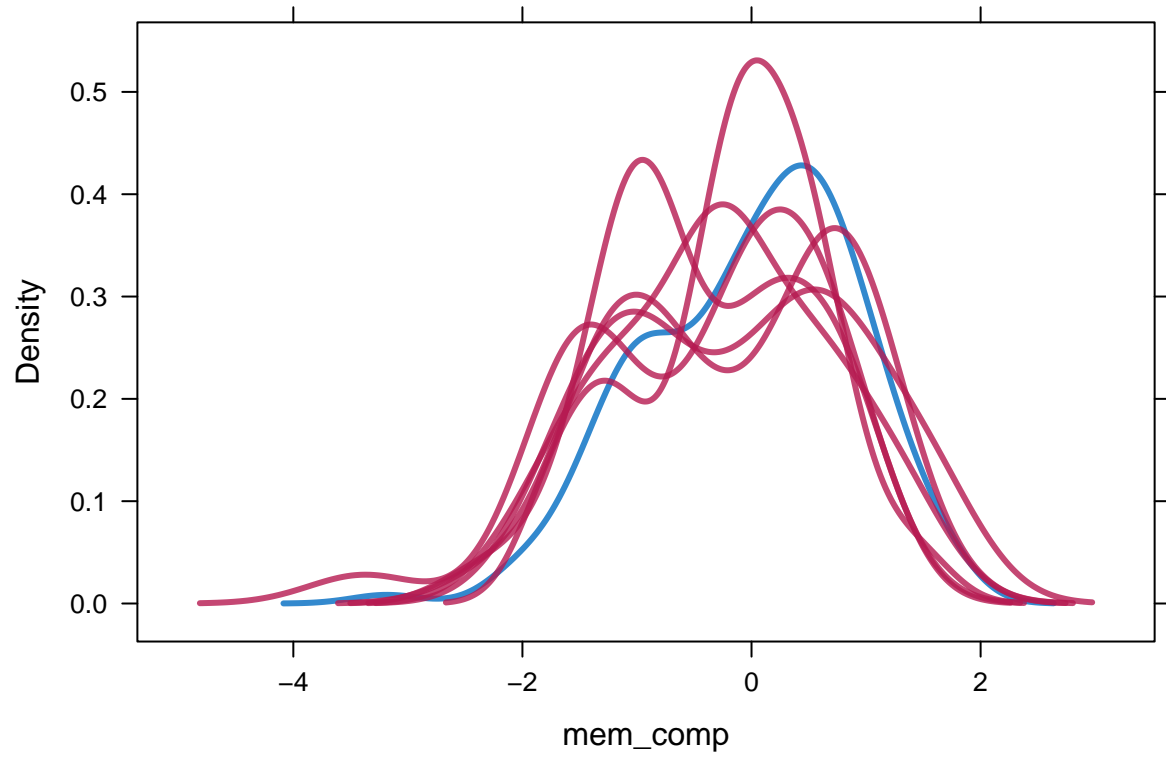
```
# imputation with no adjustment (delta = 0)
densityplot(imp.all[[1]], lwd = 3)
```



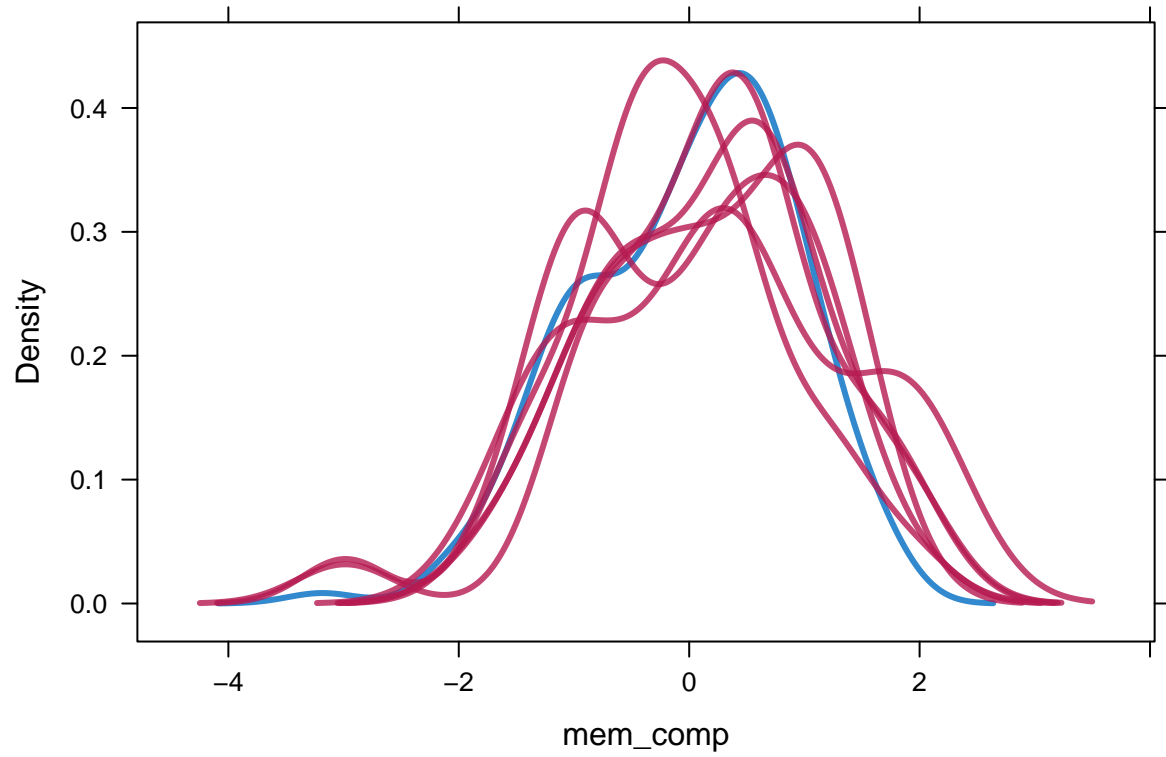
```
# imputation with adjustment (delta = -0.7)
densityplot(imp.all[[2]], lwd = 3)
```



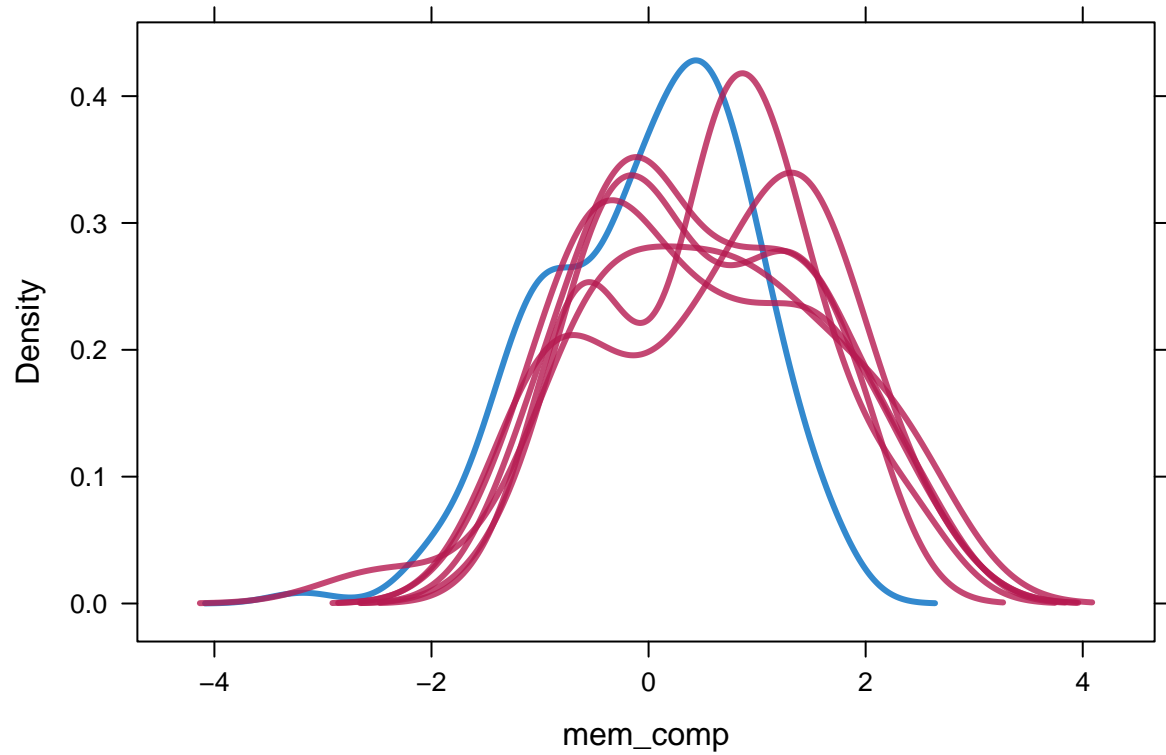
```
# imputation with adjustment (delta = -0.2)
densityplot(imp.all[[3]], lwd = 3)
```



```
# imputation with adjustment (delta = 0.2)  
densityplot(imp.all[[4]], lwd = 3)
```



```
# imputation with adjustment (delta = 0.7)  
densityplot(imp.all[[5]], lwd = 3)
```

```
# fit model to each imputed datasets  
# pool coef estimates and standard errors to create 95% CIs  
# sensitivity analysis for each data missing assumptions  
# mcar - fit data with completers data
```