Supplement1

Akihiro Shiroshita

2021/9/2

1. Main analysis

We used the following packages; tidyverse (version 1.3.1), lme4 (version 1.1-27.1), geepack (version 1.3-2), mice(version 3.13.0), norm2 (version 2.0.4), mitools(version 2.4).

1-0. Complete case analysis

First, we performed a complete case analysis using a log-linked Gamma model to evaluate the association between each variable and the length of the hospital stay. We estimated the parameters using the generalized estimating equation method with an exchangeable correlation structure and robust standard error estimation. The following variables were selected as potential confounding factors: age, body mass index, home oxygen therapy use, respiratory rate, heart rate, oxygen use on admission, mental status, systemic steroid use on admission or the next day regardless of the dose, activities of daily living, and the number of recurrences. Hospital-specific effects were implemented as fixed effects.

```
fit_cc <- geeglm(los ~ broad
                 + age
                 + bmi
                 + count
                 + o2
                 + hot
                 + adl
                 + ams
                 + rr
                 + hr
                 + steroid
                 + factor(hospital),
                 id = id,
                 std.err = "san.se",
                 corstr = "exchangeable",
                 family = Gamma(link = "log"),
                 data = data_cc)
summary(fit_cc)
```

```
##
## Call:
## geeglm(formula = los ~ broad + age + bmi + count + o2 + hot +
##     adl + ams + rr + hr + steroid + factor(hospital), family = Gamma(link = "log"),
##     data = data_cc, id = id, corstr = "exchangeable", std.err = "san.se")
##
##  Coefficients:
##                  Estimate    Std.err   Wald Pr(>|W|)
## (Intercept)     2.6617515  0.3875091 47.181 6.47e-12 ***
## broadTRUE      -0.0216821  0.0749919  0.084  0.77249
## age             0.0046328  0.0038785  1.427  0.23229
```

```
## bmi                -0.0219792  0.0076132  8.335  0.00389 **
## count               0.0015627  0.0378927  0.002  0.96710
## o2TRUE             -0.0273225  0.0858043  0.101  0.75016
## hotTRUE            -0.0031426  0.0601726  0.003  0.95835
## adl                -0.0055182  0.0008417 42.982 5.53e-11 ***
## amsTRUE             0.0763017  0.1250161  0.373  0.54164
## rr                 -0.0004810  0.0052290  0.008  0.92671
## hr                 -0.0136294  0.0551359  0.061  0.80476
## steroidTRUE         0.0398069  0.0880865  0.204  0.65134
## factor(hospital)2   0.3785102  0.0846625 19.988 7.79e-06 ***
## factor(hospital)4   0.4206083  0.1417393  8.806  0.00300 **
## factor(hospital)5   0.2136787  0.1084503  3.882  0.04880 *
## factor(hospital)6   0.3204698  0.1389944  5.316  0.02113 *
## factor(hospital)7   1.0671119  0.2706493 15.546 8.05e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation structure = exchangeable
## Estimated Scale Parameters:
##
##             Estimate Std.err
## (Intercept)   0.5483   0.146
##   Link = identity
##
## Estimated Correlation Parameters:
##       Estimate Std.err
## alpha   0.3011  0.3076
## Number of clusters:   254  Maximum cluster size: 13
```

1-1. Multiple imputation by chained equations

Based on missing at random assumption, we created 100 datasets using multiple imputation by chained equations.

We used the following variables to estimate the missing data.

- covariates of the outcome analysis: number of recurrence, age, home oxygen therapy use, oxygen use on admission, heart rate, respiratory rate, mental status, systemic steroid use, admitting hospital

- treatment variable: anti-pseudomonal antibiotics use

- the outcome variable: length of hospital stay

- auxiliary variable: systolic blood pressure

```
data_imp <- mice(data,
                 m = 100,
                 predictorMatrix = predictorMatrix100,
                 maxit = 20,
                 printFlag = FALSE,
                 seed = 1234)
```

1-2. Outcome analysis within each dataset

We used a log-linked Gamma model to evaluate the association between each variable and the length of the hospital stay. We estimated the parameters using the generalized estimating equation method with an exchangeable correlation structure and robust standard error estimation. The following variables were

selected as potential confounding factors: age, body mass index, home oxygen therapy use, respiratory rate, heart rate, oxygen use on admission), mental status, systemic steroid use on admission or the next day regardless of the dose, activities of daily living, and the number of recurrences. Hospital-specific effects were implemented as fixed effects.

```r
est <- se <- vector(length = data_imp$m,
                    mode = "list")
for (i in 1:data_imp$m) {
  com <- complete(data_imp, i)
  com <- com %>%
    mutate(bmi = weight/(height/100)^2) %>%
    mutate(hr = if_else(hr >= 109, 1, 0))
  fit <- geeglm(los ~ broad
                + age
                + bmi
                + count
                + o2
                + hot
                + adl
                + ams
                + rr
                + hr
                + steroid
                + hospital,
                id = id,
                std.err = "san.se",
                corstr = "exchangeable",
                family = Gamma(link = "log"),
                data = com)
  s <- summary(fit)
  est[[i]] <- s[["coefficients"]][["Estimate"]]
  se[[i]] <- s[["coefficients"]][["Std.err"]]
}
```
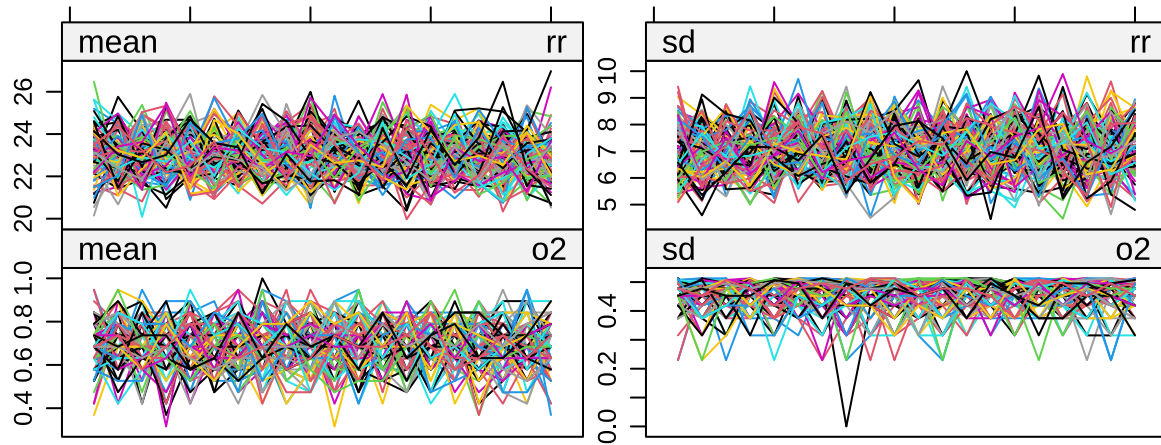
Here, we checked the convergence using spagetti plots.

```r
plot(data_imp)
```

mean | rr

sd | rr

mean | o2

sd | o2
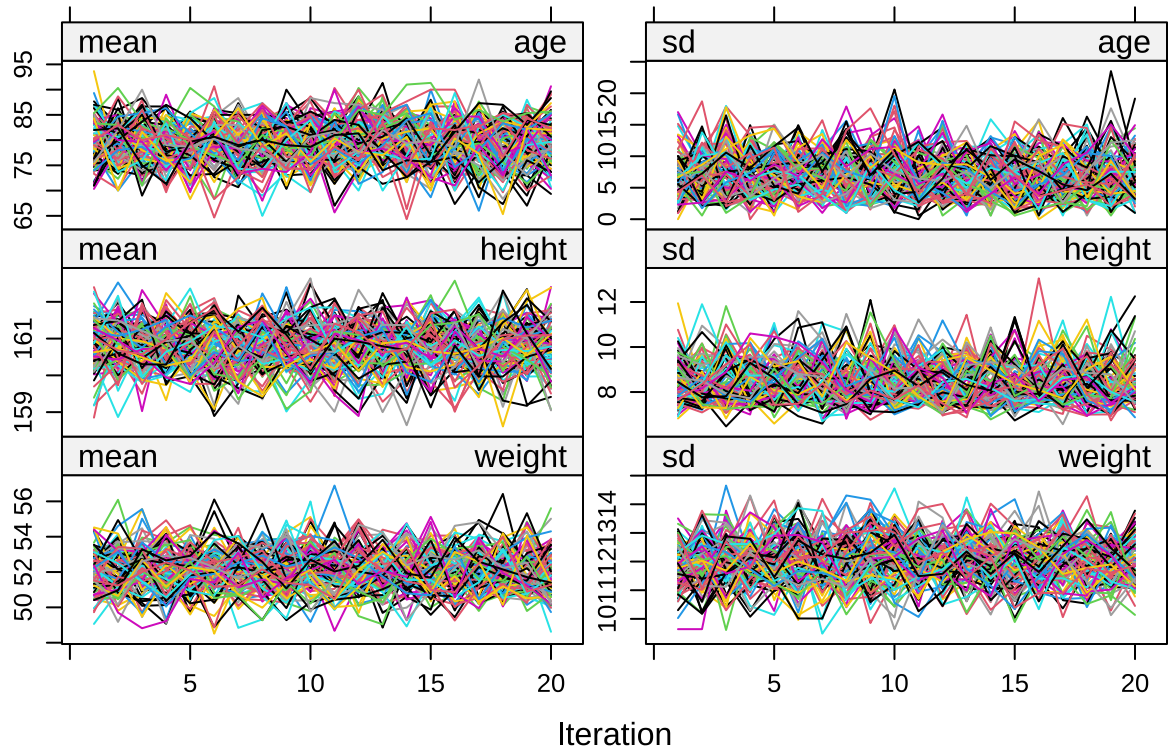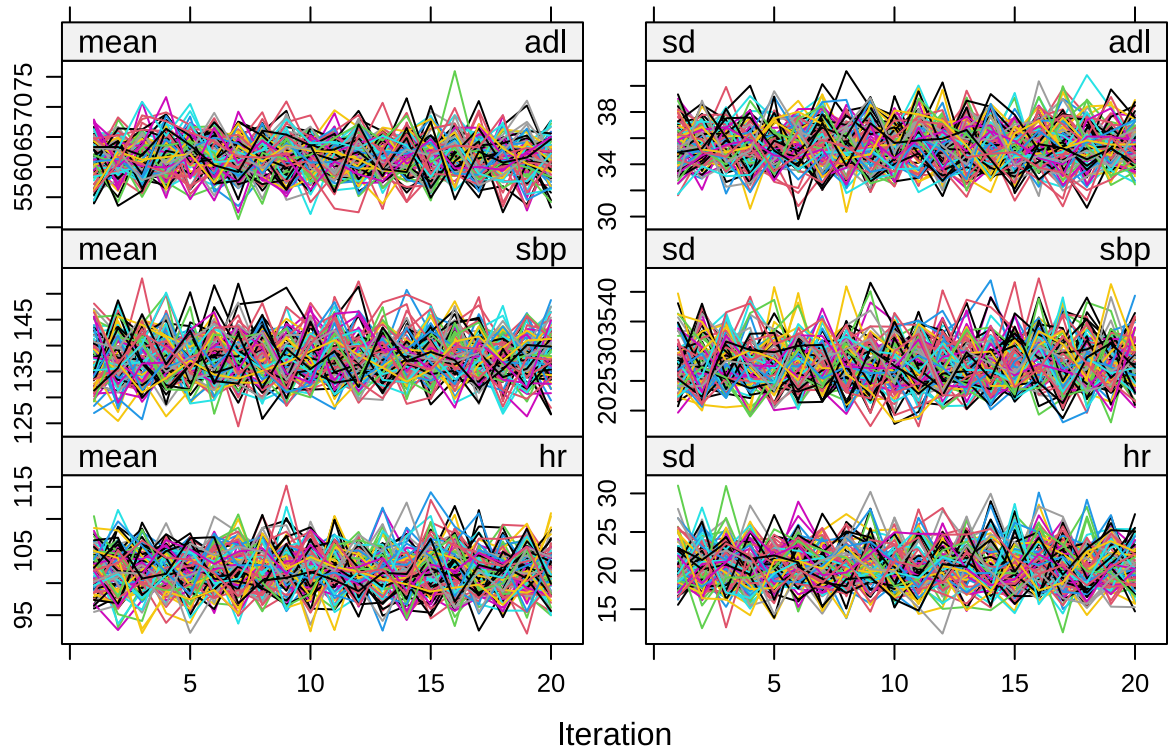
Iteration

1-3. Combining the results using Rubin's rule

We calculated the point estimate and 95% confidence interval using Rubin's combining rule.

```r
miinf <- miInference(est, se)
print(miinf)
```

```
##              Est         SE Est/SE         df      p Pct.mis
##  [1,]   2.291800 0.3445300  6.652      10282 0.000     9.8
##  [2,]   0.040971 0.0627040  0.653     204411 0.514     2.2
##  [3,]   0.008567 0.0031062  2.758      58783 0.006     4.1
##  [4,]  -0.014071 0.0084542 -1.664       1498 0.096    25.7
##  [5,]   0.003183 0.0291250  0.109   16153086 0.913     0.2
##  [6,]   0.010999 0.0651070  0.169      90827 0.866     3.3
##  [7,]   0.006017 0.0509250  0.118     128402 0.906     2.8
##  [8,]  -0.005660 0.0007351 -7.700       1748 0.000    23.8
##  [9,]   0.073310 0.0807130  0.908      69576 0.364     3.8
## [10,]  -0.002696 0.0040909 -0.659      10488 0.510     9.7
## [11,]   0.049666 0.0537230  0.924       5937 0.355    12.9
## [12,]  -0.108030 0.0732730 -1.474      99199 0.140     3.2
## [13,]   0.438030 0.0792680  5.526    1166795 0.000     0.9
## [14,]   0.467630 0.1084800  4.311      57858 0.000     4.1
## [15,]   0.464890 0.0985390  4.718      67901 0.000     3.8
## [16,]   0.379000 0.1005400  3.770     183402 0.000     2.3
## [17,]   0.432330 0.1018800  4.244      29256 0.000     5.8
## [18,]   0.979260 0.1320700  7.415     257010 0.000     2.0
```

```
# "Intercept",
# "Anti-pseudomonal antibiotics",
# "Age",
# "Body mass index",
# "The number of recurrences",
# "Oxygen demand",
# "Home oxygen therapy",
# "Activity of daily living",
# "Altered mental status",
# "Respiratory rate",
# "Heart rate",
# "Steroid therapy",
# "Hospital1",
# "Hospital2",
# "Hospital3",
# "Hospital4",
# "Hospital5",
# "Hospital6"
```

2. Sensitivity analysis

As a sensitivity analysis, we added the additional covariates as confounding factors; COPD stage, and the use of "inhaled corticosteroid", "long-acting beta2-agonist", and "long-acting muscarinic antagonist".

```
data_imp_sen <- mice(data_sen,
                     m = 100,
                     predictorMatrix = predictorMatrix_sen,
                     maxit = 20,
                     printFlag = FALSE,
                     seed = 1234)
est_sen <- se_sen <- vector(length = data_imp_sen$m,
                     mode = "list")
for (i in 1:data_imp_sen$m) {
  com <- complete(data_imp_sen, i)
  com <- com %>%
    mutate(bmi = weight/(height/100)^2) %>%
    mutate(hr = if_else(hr >= 109, 1, 0))
  fit <- geeglm(los ~ broad
                + age
                + bmi
                + stage
                + count
                + o2
                + hot
                + adl
                + ams
                + rr
                + hr
                + steroid
                + ics
                + laba
                + lama
                + hospital,
                id = id,
                std.err = "san.se",
```

```r
                corstr = "exchangeable",
                family = Gamma(link = "log"),
                data = com)
  s <- summary(fit)
  est_sen[[i]] <- s[["coefficients"]][["Estimate"]]
  se_sen[[i]] <- s[["coefficients"]][["Std.err"]]
}
miinf_sen <- miInference(est_sen, se_sen)
print(miinf_sen)
```

```
##             Est        SE  Est/SE         df     p Pct.mis
##  [1,]  2.108200 0.3911600   5.390     2893.7 0.000    18.5
##  [2,]  0.036195 0.0616590   0.587    66999.6 0.557     3.8
##  [3,]  0.008935 0.0032358   2.761    15611.7 0.006     8.0
##  [4,] -0.011928 0.0088070  -1.354     1089.2 0.176    30.1
##  [5,]  0.122670 0.1007700   1.217      286.6 0.224    58.8
##  [6,]  0.005062 0.0290870   0.174  1989574.9 0.862     0.7
##  [7,]  0.002325 0.0640030   0.036    57497.4 0.971     4.1
##  [8,]  0.015227 0.0501580   0.304    53865.9 0.761     4.3
##  [9,] -0.005597 0.0007342  -7.623     2255.6 0.000    21.0
## [10,]  0.088926 0.0806290   1.103    47076.7 0.270     4.6
## [11,] -0.002449 0.0041889  -0.585     5354.0 0.559    13.6
## [12,]  0.057907 0.0532430   1.088     4954.9 0.277    14.1
## [13,] -0.090762 0.0717960  -1.264    42939.1 0.206     4.8
## [14,] -0.009425 0.0568370  -0.166     4920.6 0.868    14.2
## [15,]  0.036436 0.0534940   0.681     7637.8 0.496    11.4
## [16,] -0.122830 0.0615030  -1.997     9165.0 0.046    10.4
## [17,]  0.527070 0.0980640   5.375     2344.9 0.000    20.5
## [18,]  0.502530 0.1091200   4.605     8857.8 0.000    10.6
## [19,]  0.538180 0.1046500   5.143    15944.1 0.000     7.9
## [20,]  0.415160 0.1044700   3.974    12295.5 0.000     9.0
## [21,]  0.483700 0.1046900   4.620     9789.8 0.000    10.1
## [22,]  0.990660 0.1320100   7.504    53600.9 0.000     4.3
```

```r
# "Intercept",
# "Anti-pseudomonal antibiotics",
# "Age",
# "Body mass index",
# "Stage",
# "The number of recurrences",
# "Oxygen demand",
# "Home oxygen therapy",
# "Activity of daily living",
# "Altered mental status",
# "Respiratory rate",
# "Heart rate",
# "Steroid therapy",
# "Inhaled corticoteroid"
# "long-acting beta2-agonist"
# "long-acting muscarinic antagonist"
# "Hospital1",
# "Hospital2",
# "Hospital3",
# "Hospital4",
```

```
# "Hospital5",
# "Hospital6"
```