Generalised Estimating Equations

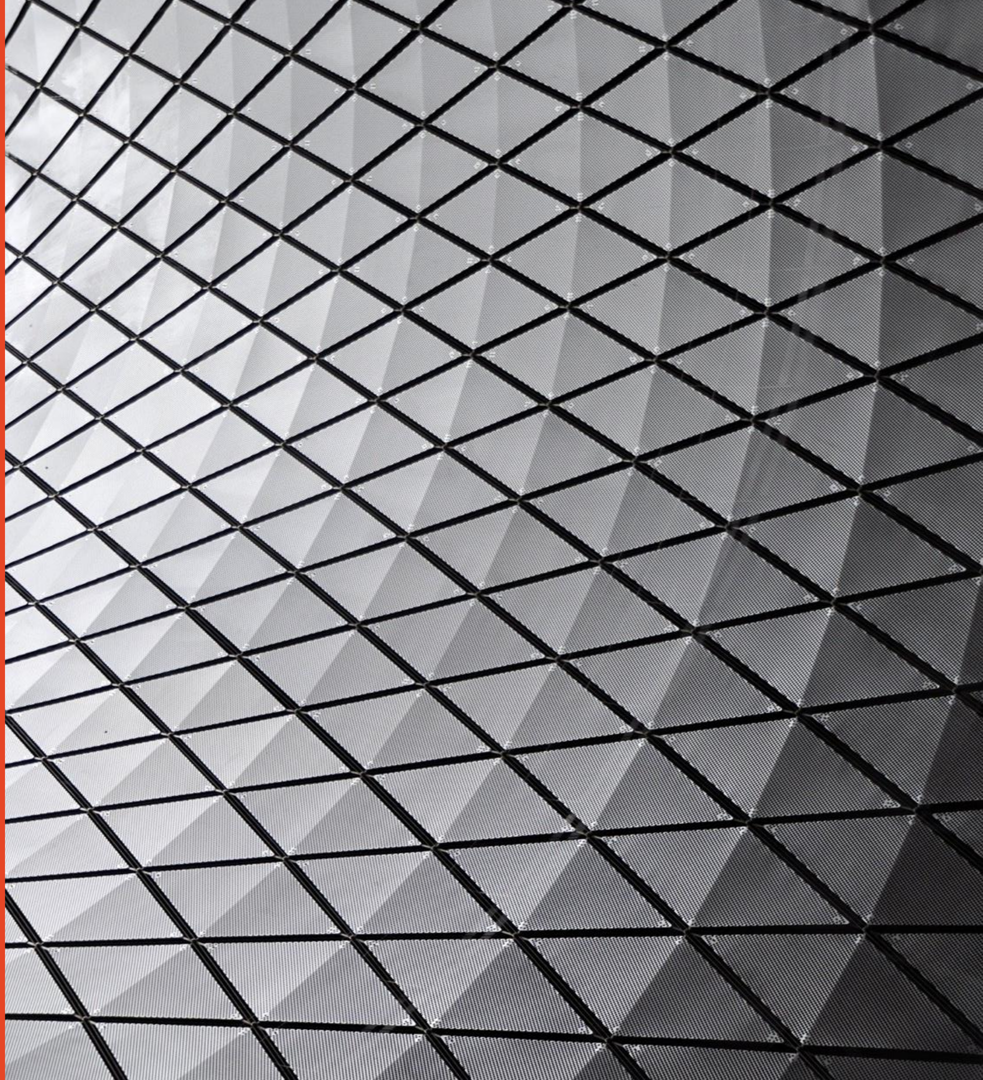
Workshop: Analysis of Longitudinal Data

12th Nov 2024

Jaroslaw Harezlak

Armando Teixeira-Pinto

THE UNIVERSITY OF SYDNEY

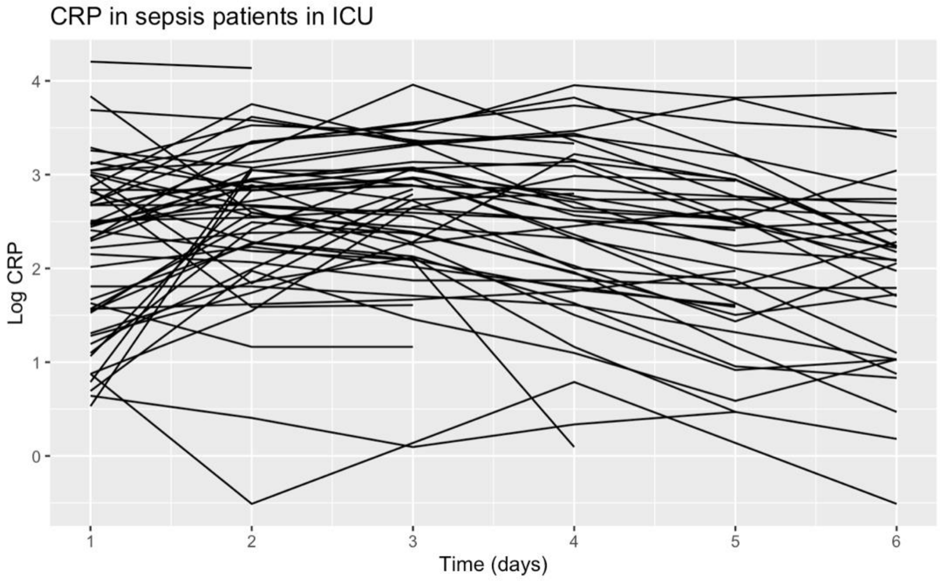


Generalised Estimating Equations

- An alternative to the mixed models approach for clustered data introduced by Liang and Zeger (1986)
- Estimation occurs in two steps
  - Fit a GLM model assuming some correlation (covariance) structure – this is called the "working" correlation
  - Correct the standard errors using the sandwich (or robust) estimator

CRP in ICU example

- Longitudinal dataset
- C-reactive protein (CRP) measured for the first 6 days on admission to ICU for sepsis for 120 patients
- We will use the log(CRP) due to skewness of CRP
- The measurements are clustered by patient



Generalised Estimating Equations

- In the CRP data example, we can fit the usual linear regression

$$logCRP_{ij} = \beta_0 + \beta_1 day_{ij} + \varepsilon_{ij}$$

- Allowing some $\boxed{\text{correlation for the } \varepsilon_{ij}}$

We need to choose a correlation structure

- I.e., the CRPs measured in different days are allowed to be correlated
- We need to specify a structure (a model) for the correlation between the CRP measurements

GEE working correlation structure

– Independent

$$\begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

GEE working correlation structure

– Independent

$$\begin{array}{c|cccc} & D1 & D2 & \cdots & D6 \\ \hline D1 & & & & \\ D2 & & & & \\ \vdots & & & & \\ D6 & & & & \end{array}$$

$$\begin{array}{cc} & \begin{array}{cccc} D1 & D2 & \cdots & D6 \end{array} \\ \begin{array}{c} D1 \\ D2 \\ \vdots \\ D6 \end{array} & \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \end{array}$$

GEE working correlation structure

- Exchangeable

The correlation between measures is the same for all the days

$$\begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}$$

GEE working correlation structure

- Exchangeable

$$
\begin{array}{c|cccc}
 & D1 & D2 & \cdots & D6 \\
\hline
D1 & & & & \\
D2 & & & & \\
\vdots & & & & \\
D6 & & & &
\end{array}
\quad
\begin{array}{c} D1 \\ D2 \\ \vdots \\ D6 \end{array}
\begin{pmatrix}
1 & .5 & \cdots & .5 \\
.5 & 1 & \cdots & .5 \\
\vdots & \vdots & & \vdots \\
.5 & .5 & \cdots & 1
\end{pmatrix}
$$

For example, the correlation between CRP at day 1 and day 6 is 0.5

GEE working correlation structure

- Exchangeable
- If there is continuous time:

patid	time	GDS
1	0.00	1.44
1	2.19	1.5
1	2.72	0
2	0.00	0.07
2	0.48	0.23
2	2.81	0
3	0.00	0.07
3	0.59	1.92
4	0.00	0.08
4	1.03	0.31
4	1.97	0.69
4	2.49	0.69

$$\begin{array}{cc} & \begin{array}{cccc} \text{1st} & \text{2nd} & \cdots & \end{array} \\ \begin{array}{c} \text{1st} \\ \text{2nd} \\ \vdots \\ \\ \end{array} & \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix} \end{array}$$

GEE working correlation structure

- Exchangeable
- If there is continuous time

patid	time	GDS
1	0.00 **(1st)**	1.44
1	2.19 **(2nd)**	1.5
1	2.72 **(3rd)**	0
2	0.00 **(1st)**	0.07
2	0.48 **(2nd)**	0.23
2	2.81 **(3rd)**	0
3	0.00 **(1st)**	0.07
3	0.59 **(2nd)**	1.92
4	0.00 **(1st)**	0.08
4	1.03 **(2nd)**	0.31
4	1.97 **(3rd)**	0.69
4	2.49 **(4th)**	0.69

$$\begin{array}{c c} & \begin{array}{cccc} \text{1st} & \text{2nd} & \cdots & \end{array} \\ \begin{array}{c} \text{1st} \\ \text{2nd} \\ \vdots \\ \\ \end{array} & \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix} \end{array}$$

GEE working correlation structure

- Unstructured

$$\begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1J} \\ \rho_{12} & 1 & \cdots & \rho_{2J} \\ \vdots & \vdots & & \vdots \\ \rho_{1J} & \rho_{2J} & \cdots & 1 \end{pmatrix}$$

Use with caution – often leads to convergence problems

GEE working correlation structure

– Autoregressive Order 1 (AR1)

Measurements that are closer are more strongly correlated than measurements far away

$$\begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 & \dots & \rho^p \\ \rho & 1 & \rho & \rho^2 & \dots & \rho^{p-1} \\ \rho^2 & \rho & 1 & \rho & \dots & \rho^{p-2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \rho^p & \rho^{p-1} & \rho^{p-2} & \rho^{p-3} & \dots & 1 \end{pmatrix}$$

GEE working correlation structure

- Autoregressive Order 1 (AR1)

$$\begin{pmatrix} 1 & .5 & .25 & .125 & \dots & \\ .5 & 1 & .5 & .25 & \dots & \\ .25 & .5 & 1 & .5 & \dots & \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \\ \dots & & & & \dots & 1 \end{pmatrix}$$

GEE working correlation structure

- Autoregressive Order 1 (AR1)
- If there is continuous time:

patid	time	GDS
1	0.00 **(1st)**	1.44
1	2.19 **(2nd)**	1.5
1	2.72 **(3rd)**	0
2	0.00 **(1st)**	0.07
2	0.48 **(2nd)**	0.23
2	2.81 **(3rd)**	0
3	0.00 **(1st)**	0.07
3	0.59 **(2nd)**	1.92
4	0.00 **(1st)**	0.08
4	1.03 **(2nd)**	0.31
4	1.97 **(3rd)**	0.69
4	2.49 **(4th)**	0.69

$$\begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 & \dots & \rho^p \\ \rho & 1 & \rho & \rho^2 & \dots & \rho^{p-1} \\ \rho^2 & \rho & 1 & \rho & \dots & \rho^{p-2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \rho^p & \rho^{p-1} & \rho^{p-2} & \rho^{p-3} & \dots & 1 \end{pmatrix}$$

GEE working correlation structure

- Autoregressive Order 1 (AR1)
- If there is continuous time:

patid	time	GDS
1	0.00 **(1st)**	1.44
1	2.19 **(2nd)**	1.5
1	2.72 **(3rd)**	0
2	0.00 **(1st)**	0.07
2	0.48 **(2nd)**	0.23
2	2.81 **(3rd)**	0
3	0.00 **(1st)**	0.07
3	0.59 **(2nd)**	1.92
4	0.00 **(1st)**	0.08
4	1.03 **(2nd)**	0.31
4	1.97 **(3rd)**	0.69
4	2.49 **(4th)**	0.69

$$\begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 & \dots & \rho^p \\ \rho & 1 & \rho & \rho^2 & \dots & \rho^{p-1} \\ \rho^2 & \rho & 1 & \rho & \dots & \rho^{p-2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \rho^p & \rho^{p-1} & \rho^{p-2} & \rho^{p-3} & \dots & 1 \end{pmatrix}$$

GEE working correlation structure

- For very different measurement times across individuals, independence and exchangeable are common choices
- The AR1 and unstructured will use the order of the number of observations per individual, rather than the time itself
- In particular, if a few individuals have many observations the unstructured correlation is not appropriate

Generalised Estimating Equations

- We **don't need** to have the structure for the correlation correctly specified
- If this happens, the SE computed using that correlation will be incorrect
- However, we can "fix" it using the **sandwich estimator**
- Also know as **robust** $V(\hat{\beta})$

$$\underbrace{[X^T V_w^{-1} X]^{-1}}_{\text{Naïve variance}} \underbrace{\left[\sum_i X_i^T V_w^{-1} (y_i - \mu_i)(y_i - \mu_i)^T V_w^{-1} X_i\right]}_{\text{Empirical variance}} \underbrace{[X^T V_w^{-1} X]^{-1}}_{\text{Naïve variance}}$$

GEE: Exchangeable Correlation Matrix

```
library(geepack)

gee.exch <- geeglm(logcrp ~ day,                   #model
                   id=ID,                          #cluster variable
                   corstr = "exchangeable",        #working correlation
                   data=crp.Data)

summary(gee.exch)
```

GEE: Exchangeable Correlation Matrix

```
library(geepack)

gee.exch <- geeglm(logcrp ~ day,
                   id=ID,
                   corstr = "exchangeable",
                   data=crp.Data)
  summary(gee.exch)
```

```
 Coefficients:
            Estimate Std.err   Wald Pr(>|W|)
(Intercept)   2.5832  0.0792 1064.4  < 2e-16 ***
day          -0.0893  0.0221   16.2  5.6e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'
0.1 ' ' 1

Correlation structure = exchangeable
Estimated Scale Parameters:

            Estimate Std.err
(Intercept)    0.807   0.102
  Link = identity

Estimated Correlation Parameters:
      Estimate Std.err
alpha    0.612   0.058
Number of clusters:   120  Maximum cluster size: 6
```

GEE: Exchangeable Correlation Matrix

```
library(geepack)

gee.exch <- geeglm(logcrp ~ day,
                   id=ID,
                   corstr = "exchangeable",
                   data=crp.Data)
  summary(gee.exch)
```

```
 Coefficients:
            Estimate  Std.err   Wald Pr(>|W|)
(Intercept)   2.5832   0.0792 1064.4  < 2e-16 ***
day          -0.0893   0.0221   16.2  5.6e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'
0.1 ' ' 1

Correlation structure = exchangeable
Estimated Scale Parameters:

            Estimate Std.err
(Intercept)    0.807   0.102
  Link = identity

Estimated Correlation Parameters:
      Estimate Std.err
alpha    0.612   0.058
Number of clusters:   120  Maximum cluster size: 6
```

Robust standard errors reported by default

GEE: Exchangeable Correlation Matrix

– Other software/packages may present both the sandwich and naïve SE

```
library(gee)

gee.exch2 <- gee(logcrp ~ day,
                 data = crp.Data,
                 id = ID,
                 corstr = "exchangeable")
summary(gee.exch2)
```

```
Coefficients:
            Estimate Naive S.E. Naive z Robust S.E. Robust z
(Intercept)   2.5832     0.0815   31.71      0.0792    32.62
day          -0.0893     0.0138   -6.49      0.0221    -4.03
```

GEE: Exchangeable Correlation Matrix

$$
\begin{array}{c c}
 & \begin{array}{cccc} D1 & D2 & \cdots & D6 \end{array} \\
\begin{array}{c} D1 \\ D2 \\ \vdots \\ D6 \end{array} & \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}
\end{array}
$$

```
Call:
geeglm(formula = logcrp ~ day, data = crp.Data,
id = ID, corstr = "exchangeable")

 Coefficients:
            Estimate Std.err   Wald Pr(>|W|)
(Intercept)   2.5832  0.0792 1064.4  < 2e-16 ***
day          -0.0893  0.0221   16.2  5.6e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'
0.1 ' ' 1

Correlation structure = exchangeable
Estimated Scale Parameters:

            Estimate Std.err
(Intercept)    0.807   0.102
  Link = identity

Estimated Correlation Parameters:
      Estimate Std.err
alpha    0.612   0.058
Number of clusters:   120  Maximum cluster size: 6
```

GEE: Exchangeable Correlation Matrix

$$
\begin{array}{cc}
 & \begin{array}{cccc} D1 & D2 & \cdots & D6 \end{array} \\
\begin{array}{c} D1 \\ D2 \\ \vdots \\ D6 \end{array} &
\begin{pmatrix}
1 & \rho & \cdots & \rho \\
\rho & 1 & \cdots & \rho \\
\vdots & \vdots & & \vdots \\
\rho & \rho & \cdots & 1
\end{pmatrix}
\end{array}
$$

```
Call:
geeglm(formula = logcrp ~ day, data = crp.Data,
id = ID, corstr = "exchangeable")

 Coefficients:
            Estimate Std.err   Wald Pr(>|W|)
(Intercept)   2.5832  0.0792 1064.4  < 2e-16 ***
day          -0.0893  0.0221   16.2  5.6e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'
0.1 ' ' 1

Correlation structure = exchangeable
Estimated Scale Parameters:

            Estimate Std.err
(Intercept)    0.807   0.102
  Link = identity

Estimated Correlation Parameters:
      Estimate Std.err
alpha    0.612   0.058
Number of clusters:   120  Maximum cluster size: 6
```

GEE: Exchangeable Correlation Matrix

- The random intercept model implies an exchangeable correlation structure

$$logCRP_{ij} = \beta_0 + \beta_1 day_{ij} + \boxed{b_{i0}} + \varepsilon_{ij}$$

The observations of each patient share this random effect. This induces correlation between the observations

- So the coefficients should be the same (similar) as the GEE with exchangeable correlation structure

- And the correlation from the random effect model is obtained as the $\rho = \frac{var(b_{i0})}{var(b_{i0})+var(\varepsilon_{ij})}$

GEE: Exchangeable Correlation Matrix

GEE

```
Call:
geeglm(formula = logcrp ~ day, data = crp.Data,
id = ID, corstr = "exchangeable")

 Coefficients:
            Estimate Std.err   Wald Pr(>|W|)
(Intercept)   2.5832  0.0792 1064.4  < 2e-16 ***
day          -0.0893  0.0221   16.2  5.6e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation structure = exchangeable
Estimated Scale Parameters:

            Estimate Std.err
(Intercept)    0.807   0.102
  Link = identity

Estimated Correlation Parameters:
      Estimate Std.err
alpha    0.612   0.058
Number of clusters:   120  Maximum cluster size: 6
```

Random intercept model

```
Linear mixed model fit by REML ['lmerMod']
Formula: logcrp ~ day + (1 | ID)
   Data: crp.Data

REML criterion at convergence: 1292

Scaled residuals:
   Min     1Q Median     3Q    Max
-3.598 -0.475  0.073  0.597  2.764

Random effects:
 Groups   Name        Variance Std.Dev.
 ID       (Intercept) 0.480    0.693
 Residual             0.321    0.566
Number of obs: 606, groups:  ID, 120


Fixed effects:
            Estimate Std. Error t value
(Intercept)   2.5832     0.0810   31.89
day          -0.0892     0.0139   -6.43
```

GEE: Exchangeable Correlation Matrix

GEE

```
Call:
geeglm(formula = logcrp ~ day, data = crp.Data,
id = ID, corstr = "exchangeable")

 Coefficients:
            Estimate Std.err   Wald Pr(>|W|)
(Intercept)   2.5832  0.0792 1064.4  < 2e-16 ***
day          -0.0893  0.0221   16.2  5.6e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation structure = exchangeable
Estimated Scale Parameters:

            Estimate Std.err
(Intercept)    0.807   0.102
  Link = identity

Estimated Correlation Parameters:
      Estimate Std.err
alpha    0.612   0.058
Number of clusters:   120  Maximum cluster size: 6
```

Random intercept model

```
Linear mixed model fit by REML ['lmerMod']
Formula: logcrp ~ day + (1 | ID)
   Data: crp.Data

REML criterion at convergence: 1292

Scaled residuals:
   Min      1Q  Median      3Q     Max
-3.598                          2.764

Random effects:
 Groups   Name        Variance Std.Dev.
 ID       (Intercept) 0.480    0.693
 Residual             0.321    0.566
Number of obs: 606, groups:  ID, 120

Fixed effects:
            Estimate Std. Error t value
(Intercept)   2.5832     0.0810   31.89
day          -0.0892     0.0139   -6.43
```

Intraclass correlation

$$\frac{0.480}{0.480 + 0.321} = 0.60$$

GEE: Independence Correlation Matrix

```
library(geepack)

gee.ind <- geeglm(logcrp ~ day,
                  id=ID,
                  corstr = "independence",
                  data=crp.Data)
summary(gee.ind)
```

```
Call:
geeglm(formula = logcrp ~ day, data = crp.Data, id = ID,
corstr = "independence")

 Coefficients:
            Estimate Std.err  Wald Pr(>|W|)
(Intercept)   2.5707  0.0823 975.8  < 2e-16 ***
day          -0.0833  0.0234  12.7  0.00037 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
' ' 1

Correlation structure = independence
Estimated Scale Parameters:

            Estimate Std.err
(Intercept)    0.807   0.102
Number of clusters:   120  Maximum cluster size: 6
```

GEE: Independence Correlation Matrix

– A **GEE with independent correlation** structure is equivalent to fitting a standard linear regression

GEE

```
Call:
geeglm(formula = logcrp ~ day, data = crp.Data,
id = ID, corstr = "independence")

 Coefficients:
            Estimate Std.err  Wald Pr(>|W|)
(Intercept)   2.5707  0.0823 975.8  < 2e-16 ***
day          -0.0833  0.0234  12.7  0.00037 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 1

Correlation structure = independence
Estimated Scale Parameters:

            Estimate Std.err
(Intercept)    0.807   0.102
Number of clusters:   120  Maximum cluster size: 6
```

Standard linear regression

```
Call:
lm(formula = logcrp ~ day, data = crp.Data)


Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.5707     0.0798   32.21  < 2e-16 ***
day          -0.0833     0.0214   -3.89  0.00011 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'

Residual standard error: 0.9 on 604 degrees of freedom
Multiple R-squared:  0.0245, Adjusted R-squared:
0.0228
F-statistic: 15.1 on 1 and 604 DF,  p-value: 0.000111
```

Comparing models fitted with GEE

- AIC, BIC and the log-likelihood ratio are based on the likelihood
- They cannot be used with GEE (it is not a likelihood-based method)
- QIC is an alternative measure to compare models fitted with GEE
- We can compare models with different covariates but also with different covariance structure

Comparing models fitted with GEE

```
> QIC(gee.bk.ind, gee.bk.exch, gee.bk.unst)
              QIC QICu Quasi Lik  CIC params QICC
gee.bk.ind   1642 1640      -818 2.81      2 1642
gee.bk.exch  1645 1642      -819 3.33      2 1645
gee.bk.unst  1644 1641      -819 3.23      2 1650
```

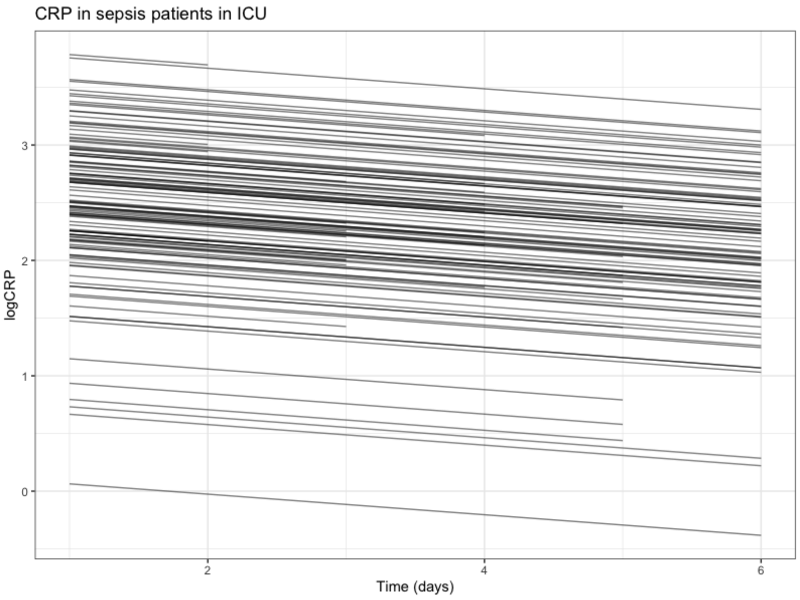
Mixed model vs GEE

- For the linear model, the random intercept model and the linear regression fitted with GEE have similar interpretation
- The regression parameters for the GEE have a marginal interpretation
- In other words, the effects estimated with the GEE are averaged (over everyone) effects
- For the random effects model, the effects are subject-specific but they are also the average effect (this is only true for the linear model)

Mixed model vs GEE

– For the random effect model, β_1 is the change of logCRP per day, for each individual

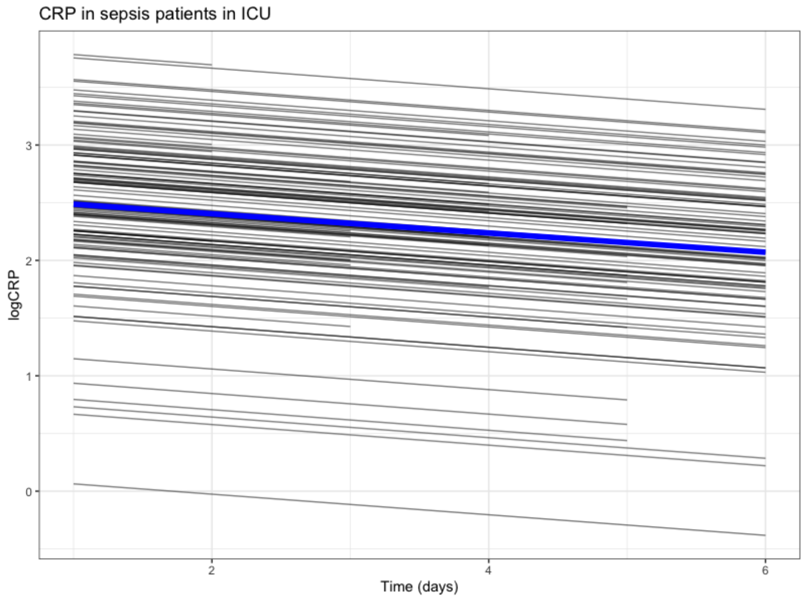
$$logCRP_{ij} = \beta_0 + \beta_1 \text{day}_{ij} + b_{i0} + \varepsilon_{ij}$$



Mixed model vs GEE

- For the random effect model, β_1 is the change of logCRP per day, for each individual
- But it is also the average change per day, across all the individuals
- β_1, in the linear case, has both a conditional (subject-specific) and a marginal (average) interpretation

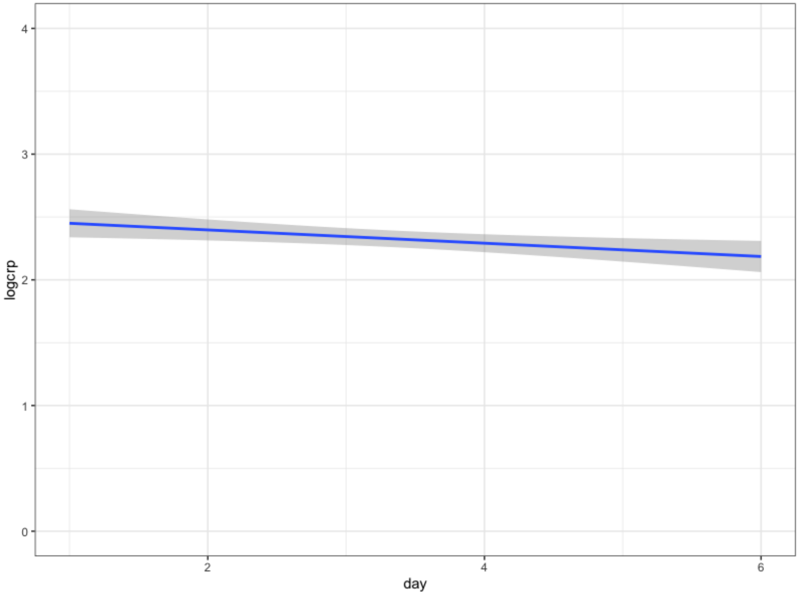
$$logCRP_{ij} = \beta_0 + \beta_1 \mathrm{day}_{ij} + b_{i0} + \varepsilon_{ij}$$



Mixed model vs GEE

- When using a GEE, β_1 is the **average change** of logCRP per day
- β_1 has a marginal (average) interpretation

$$logCRP_{ij} = \beta_0 + \beta_1 \mathrm{day}_{ij} + \varepsilon_{ij}$$



Disadvantages of GEE

- Less efficient than maximum likelihood methods
- Because it is not a likelihood-based method, it requires that the any missing date is **missing completely at random (MCAR)** while likelihood methods assume missing at random (MAR)
- It treats the correlation of observations as a nuisance rather than a feature of the data (maybe what we want!)
- Designed for large number of clusters