Thesis: 09/12/2021

Implemented lag

Georgy Gomon

December 9, 2021

Simulated data

Model 3

$$
\begin{aligned}
m_{i,j} &= \beta_0^x + u_{0,i}^x + (\beta_t^x + u_{t,i}^x)t_{i,j} \\
x_{i,j} &= m_{i,j} + \epsilon_{i,j} \\
y_{i,j} &= \gamma_0 m_{i,j} + \gamma_1 m_{i,(j-1)} + \gamma_2 m_{i,(j-2)} + \beta_0^y + u_{0,i}^y + (\beta_t^y + u_{t,i}^y)t_{i,j} \\
&\left(u_{0,i}^x, u_{t,i}^x\right) \sim \mathcal{N}(0, unstructured) \\
&\left(u_{0,i}^y, u_{t,i}^y\right) \sim \mathcal{N}(0, unstructured)
\end{aligned}
$$

- Simulation parameters: $N = 100$, $n = 30$, $\gamma_0 = 0$, $\gamma_1 = 1.2$, $\gamma_2 = 0.5$.
- Compared fits of Model 0, Model 2 and Model 3

Simulation results

```
$coefficients
         true       INLA       CI_l       CI_u
beta^x_0 2.00  1.6928667  1.6178152  1.7655318
beta^x_t 1.20  0.5431013  0.3204972  0.7382384
beta^y_0 1.50  1.8187135  1.4872715  2.1472323
beta^y_t 2.25  2.6028271  2.2143983  2.9906073
gamma_0  0.00 -0.3320000 -0.5800000 -0.0930000
gamma_1  1.30  1.7070000  1.5160000  1.9000000
gamma_2  0.50  0.1640000  0.1000000  0.2280000
```

```
$U1
$U1$true
         [,1]     [,2]
[1,] 1.921836 1.177215
[2,] 1.177215 2.207637

$U1$INLA
         [,1]     [,2]
[1,] 2.227668 1.869187
[2,] 1.869187 3.853131


$U2
$U2$true
         [,1]     [,2]
[1,] 2.725013 1.706097
[2,] 1.706097 4.371469

$U2$INLA
         [,1]     [,2]
[1,] 2.631158 1.535422
[2,] 1.535422 4.691864
```

Comparison with other Models

- Model with lag is considerably better than other models when lag is present!

	model_0	model_2	model_3
mlik	-7419.92	-7395.39	-6725.56
dic_approx	14839.84	14790.78	13451.12
dic	11967.01	11958.82	10592.85
waic	12083.93	12066.18	10665.37
pit.D	0.05	0.05	0.02
cpo	6047.19	6037.93	5337.06

Lag in the LUMC data

- In the LUMC data not all lag is present!
- The percentages of lag are:

	x
Lag 1	0.53
Lag 2	44.25
Lag 3	23.54
Lag 4	17.17
Lag 5	42.83

- I thus decided to incorporate lag of up to 5

Lag in the LUMC data: Model 3

- Model 3 with lag of order 5 has the following form:

$$
\begin{cases}
m_{i,j} = & (\beta_0^{(Cy)} + u_{0,i}^{(Cy)}) + (\beta_t^{(Cy)} + u_{t,i}^{(Cy)})t_{i,j} + (\beta_t^{(Cy),28} + u_{t,i}^{(Cy),28})(t_{i,j} - 28)_+ \\
Cy_{i,j} = & m_{i,j} + \epsilon_{i,j}^2 \\
Sc_{i,j} = & \gamma_0 m_{i,j} + \gamma_1 m_{i,(j-1)} + \gamma_2 m_{i,(j-2)} + \gamma_3 m_{i,(j-3)} + \gamma_4 m_{i,(j-4)} + \gamma_5 m_{i,(j-5)} + \\
& (\beta_0^{(Sc)} + u_{0,i}^{(Sc)}) + (\beta_t^{(Sc)} + u_{t,i}^{(Sc)})t_{i,j} + (\beta_t^{(Sc),28} + u_{t,i}^{(Sc),28})(t_{i,j} - 28)_+ + \epsilon_{i,j}^2
\end{cases}
$$

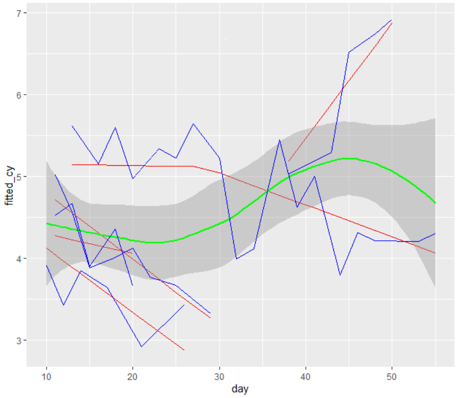
$$
\left(u_{0,i}^{(Cy)}, u_{t,i}^{(Cy)}, u_{t,i}^{(Cy),28}\right) \sim \mathcal{N}(0, unstructured)
$$

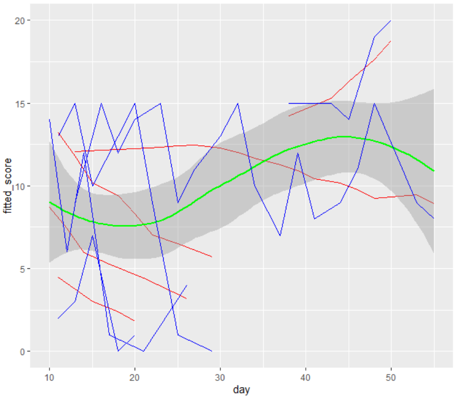
$$
\left(u_{0,i}^{(Sc)}, u_{t,i}^{(Sc)}, u_{t,i}^{(Sc),28}\right) \sim \mathcal{N}(0, unstructured)
$$

Lag in the LUMC data: Model 3 results

	mean	0.025quant	0.975quant
γ_0	1.27	0.73	1.81
γ_1	1.03	0.42	1.63
γ_2	0.93	0.44	1.41
γ_3	0.78	0.26	1.28
γ_4	1.03	0.51	1.55
γ_5	0.76	0.27	1.25

Fitted profiles





Results on the LUMC data: Model 3 comparison

	0_3	2_3	3_3	3_lagg
MLIK	-2411	-2418	**-2407**	-2434
DIC_approx	4821	4836	**4813**	4868
DIC	**4024**	4040	4060	4174
WAIC	**4013**	4031	4046	4159
PIT	**0.0254**	0.0279	0.0291	0.004
CPO	**2046**	2054	2060	2105
MSE	**3.52**	3.70	**3.52**	3.845

Biggest issues so far

- Time!
  - Patients have drastically different time profiles!
  - Breakpoints introduced so far do not solve the problem
  - Possible solutions:
    - Accelerated Faillure Time Models: Failure rate σ independent of time t, while we have no other covariates
    - Random breakpoints:
- Events occurring!
  - Possibly because of events (death) occurring
  - Possible solutions
    - Can I have data on deaths?
    - Joint survival model
    - Estimands
    - Joint Longitudinal model conditional on Being alive

Data

	ID	severity_score_pseudo	day_no	Cytokine_1	Cytokine_2	Cytokine_3
1	1	2.00	11.00	92.03	1053.03	244.03
2	1	3.00	13.00	106.68	1011.68	214.68
3	1	7.00	15.00	49.19	1010.19	183.19
4	1	0.00	18.00	77.63	1473.63	269.63
5	1	1.00	20.00	38.73	1129.73	259.73
6	2	15.00	38.00	152.03	727.03	347.03
7	2	15.00	43.00	199.81	987.81	350.81

- Missing data: only in variable day_no
  - For 2 patients: ID 46 (11 measurements) & ID 98 (8 measurements), both all time points

Transformations

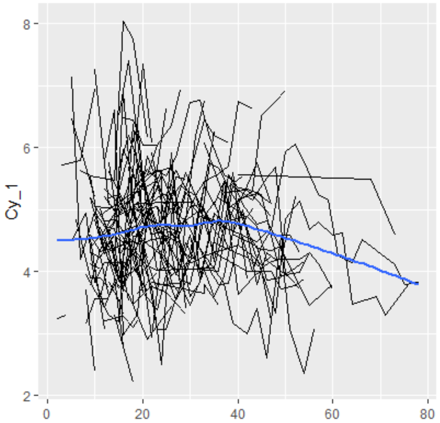
- In the table we see the correlations with severity score for different transformations of the cytokines

	NO	ln	sqrt
Cytokine 1	0.22	0.38	0.31
Cytokine 2	0.04	0.07	0.05
Cytokine 3	0.32	0.34	0.34

- We thus take a log transformation for all Cytokines
- As Cytokine 1 has highest correlation with severity score we focus on Cytokine 1

Time-Transformations

- We have only data about days when patient was at the ICU?
  - Average time-span per patient: 13 days.
    - RIVM: Average time at ICU: 13 days
  - Minimum: 0 days, Maximum: 47 days
- We see that time-patterns are very different

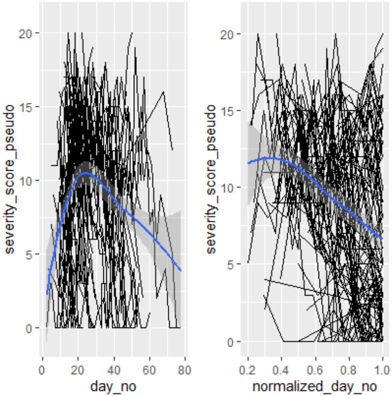


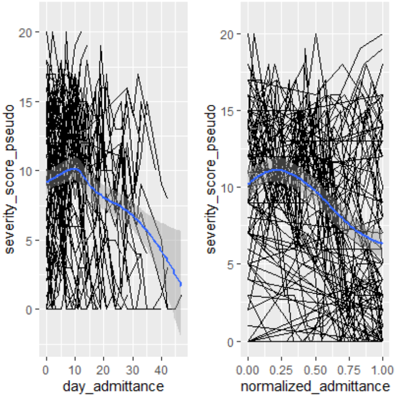
Different Time-scales

- Thus I tried different time-scales
  - Time since infection (day_no)
  - Nomalized time since infection
  - Time since admittance
  - Normalized time since admittance
- Below correlation table: However, correlation measures only linear association

	score	Cy_1	day_no	nor_day_no	day_admit	nor_admit
severity_score_pseudo	1.00	0.38	-0.07	-0.34	-0.20	-0.31
Cy_1	0.38	1.00	-0.04	-0.18	-0.07	-0.14
day_no	-0.07	-0.04	1.00	0.40	0.81	0.48
normalized_day_no	-0.34	-0.18	0.40	1.00	0.39	0.84
day_admittance	-0.20	-0.07	0.81	0.39	1.00	0.64
normalized_admittance	-0.31	-0.14	0.48	0.84	0.64	1.00

Different Time-scales





Different Time-scales

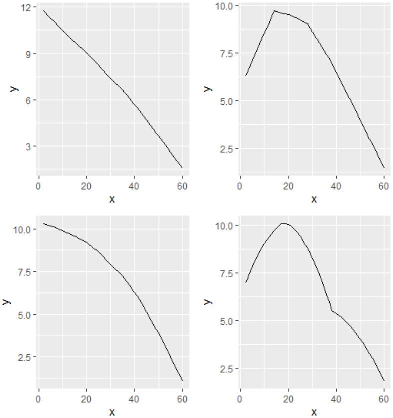
- I decided to focus on days since infection
  - Day since infection is most interesting for a doctor
  - Normalized days can not be used as a prediction model (can be useful for etiology).
- Here you do not account for different disease progression speeds in individuals.
  - Random breakpoints?

Possible Time-models

- For days since infection time is clearly not linear!
  - There are several ways in which to account for this
    - Quadratic change
    - Piecewise linear change (with breakpoints)
    - Quadratic Splines model (with breakpoints)
  - I also added Linear change model as reference
- Piecewise linear change: How to get breakpoints?
  - MARS (Multivariate adaptive regression spline)
    - For Cytokine 1: breakpoints 28 & 40
    - For Severity Score: breakpoints 14 & 28
- Quadratic Splines: location of breakpoints less important
  - I set bp to 1st and 3rd quartiles

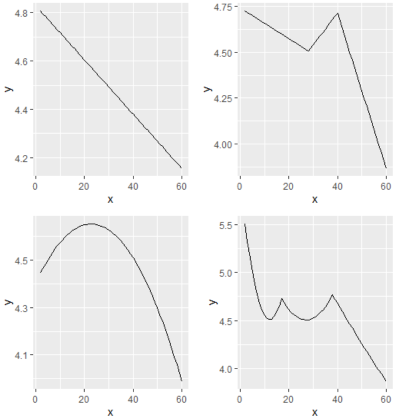
Time-models Severity Score

	Linear	Piecewise	Quadratic	Cubic
AIC	3088	3033	3083	3074
BIC	3109	3072	3109	3117
logLik	-1539	-1507	-1536	-1527



Time-models Cytokine 1

	Linear	Piecewise	Quadratic	Cubic
AIC	1333.00	1318.00	1324.00	1319.00
BIC	1350.00	1344.00	1345.00	1358.00
logLik	-662.00	-653.00	-657.00	-651.00



Modelling time: Conclusion

- I decided to stick with piecewise linear time-model
  - Breakpoints as given by MARS
- Applied Models 0 & 2 & 3 in INLA using piecewise time
  - First without breakpoints (simply linear)
  - No breakpoints within random effects
  - Only breakpoint 28 as random effects
  - All breakpoints as random effects
- Breakpoint 28 is more important than the other because:
  - Cy: Few patients have data further than bp 40 (only 25/99)
  - Score: The breakpoint at 28 seems much stronger than that at 14 (judging by loess fit)

Model 0_3: General Model structure in INLA

Model 0_3

$$
\begin{cases}
Cy_{i,j} = & (\beta_0^{(Cy)} + u_{0,i}^{(Cy)}) + \beta_{Sc}^{(Cy)} Score_{i,j} + (\beta_t^{(Cy)} + u_{t,i}^{(Cy)}) t_{i,j} + \\
& (\beta_t^{(Cy),28} + u_{t,i}^{(Cy),28})(t_{i,j} - 28)_+ + \\
& (\beta_t^{(Cy),40} + u_{t,i}^{(Cy),40})(t_{i,j} - 40)_+ + \epsilon_{i,j}^2 \\
Sc_{i,j} = & (\beta_0^{(Sc)} + u_{0,i}^{(Sc)}) + \beta_{Cy}^{(Sc)} Cy_{i,j} + (\beta_t^{(Sc)} + u_{t,i}^{(Sc)}) t_{i,j} + \\
& (\beta_t^{(Sc),28} + u_{t,i}^{(Sc),28})(t_{i,j} - 28)_+ + \\
& (\beta_t^{(Sc),14} + u_{t,i}^{(Sc),14})(t_{i,j} - 14)_+ + \epsilon_{i,j}^2
\end{cases}
$$

$$
\left(u_{0,i}^{(Cy)}, u_{t,i}^{(Cy)}, u_{t,i}^{(Cy),28}, u_{t,i}^{(Cy),40}\right) \sim \mathcal{N}(0, unstructured)
$$

$$
\left(u_{0,i}^{(Sc)}, u_{t,i}^{(Sc)}, u_{t,i}^{(Sc),28}, u_{t,i}^{(Sc),14}\right) \sim \mathcal{N}(0, unstructured)
$$

Model 0_0: No breakpoints

$$
\begin{cases}
Cy_{i,j} = & (\beta_0^{(Cy)} + u_{0,i}^{(Cy)}) + \beta_{Sc}^{(Cy)} Score_{i,j} + (\beta_t^{(Cy)} + u_{t,i}^{(Cy)})t_{i,j} + \\
& \cancel{(\beta_t^{(Cy),28} + u_{t,i}^{(Cy),28})(t_{i,j} - 28)_+ +} \\
& \cancel{(\beta_t^{(Cy),40} + u_{t,i}^{(Cy),40})(t_{i,j} - 40)_+} + \epsilon_{i,j}^2 \\
Sc_{i,j} = & (\beta_0^{(Sc)} + u_{0,i}^{(Sc)}) + \beta_{Cy}^{(Sc)} Cy_{i,j} + (\beta_t^{(Sc)} + u_{t,i}^{(Sc)})t_{i,j} + \\
& \cancel{(\beta_t^{(Sc),28} + u_{t,i}^{(Sc),28})(t_{i,j} - 28)_+ +} \\
& \cancel{(\beta_t^{(Sc),14} + u_{t,i}^{(Sc),14})(t_{i,j} - 14)_+} + \epsilon_{i,j}^2
\end{cases}
$$

$$
\left(u_{0,i}^{(Cy)}, u_{t,i}^{(Cy)}, \cancel{u_{t,i}^{(Cy),28}, u_{t,i}^{(Cy),40}}\right) \sim \mathcal{N}(0, unstructured)
$$

$$
\left(u_{0,i}^{(Sc)}, u_{t,i}^{(Sc)}, \cancel{u_{t,i}^{(Sc),28}, u_{t,i}^{(Sc),14}}\right) \sim \mathcal{N}(0, unstructured)
$$

Model 0_1: No breakpoints as random effects

$$
\begin{cases}
Cy_{i,j} = & (\beta_0^{(Cy)} + u_{0,i}^{(Cy)}) + \beta_{Sc}^{(Cy)} Score_{i,j} + (\beta_t^{(Cy)} + u_{t,i}^{(Cy)}) t_{i,j} + \\
& (\beta_t^{(Cy),28} + \cancel{u_{t,i}^{(Cy),28}})(t_{i,j} - 28)_+ + \\
& (\beta_t^{(Cy),40} + \cancel{u_{t,i}^{(Cy),40}})(t_{i,j} - 40)_+ + \epsilon_{i,j}^2 \\
Sc_{i,j} = & (\beta_0^{(Sc)} + u_{0,i}^{(Sc)}) + \beta_{Cy}^{(Sc)} Cy_{i,j} + (\beta_t^{(Sc)} + u_{t,i}^{(Sc)}) t_{i,j} + \\
& (\beta_t^{(Sc),28} + \cancel{u_{t,i}^{(Sc),28}})(t_{i,j} - 28)_+ + \\
& (\beta_t^{(Sc),14} + \cancel{u_{t,i}^{(Sc),14}})(t_{i,j} - 14)_+ + \epsilon_{i,j}^2
\end{cases}
$$

$$
\left(u_{0,i}^{(Cy)}, u_{t,i}^{(Cy)}, \cancel{u_{t,i}^{(Cy),28}, u_{t,i}^{(Cy),40}}\right) \sim \mathcal{N}(0, unstructured)
$$

$$
\left(u_{0,i}^{(Sc)}, u_{t,i}^{(Sc)}, \cancel{u_{t,i}^{(Sc),28}, u_{t,i}^{(Sc),14}}\right) \sim \mathcal{N}(0, unstructured)
$$

Model 0_2: Only breakpoint 28 as random effect

$$
\begin{cases}
Cy_{i,j} = & (\beta_0^{(Cy)} + u_{0,i}^{(Cy)}) + \beta_{Sc}^{(Cy)} Score_{i,j} + (\beta_t^{(Cy)} + u_{t,i}^{(Cy)}) t_{i,j} + \\
& (\beta_t^{(Cy),28} + u_{t,i}^{(Cy),28})(t_{i,j} - 28)_+ + \\
& (\beta_t^{(Cy),40} + \cancel{u_{t,i}^{(Cy),40}})(t_{i,j} - 40)_+ + \epsilon_{i,j}^2 \\
Sc_{i,j} = & (\beta_0^{(Sc)} + u_{0,i}^{(Sc)}) + \beta_{Cy}^{(Sc)} Cy_{i,j} + (\beta_t^{(Sc)} + u_{t,i}^{(Sc)}) t_{i,j} + \\
& (\beta_t^{(Sc),28} + u_{t,i}^{(Sc),28})(t_{i,j} - 28)_+ + \\
& (\beta_t^{(Sc),14} + \cancel{u_{t,i}^{(Sc),14}})(t_{i,j} - 14)_+ + \epsilon_{i,j}^2
\end{cases}
$$

$$
\left(u_{0,i}^{(Cy)}, u_{t,i}^{(Cy)}, u_{t,i}^{(Cy),28}, \cancel{u_{t,i}^{(Cy),40}}\right) \sim \mathcal{N}(0, unstructured)
$$

$$
\left(u_{0,i}^{(Sc)}, u_{t,i}^{(Sc)}, u_{t,i}^{(Sc),28}, \cancel{u_{t,i}^{(Sc),14}}\right) \sim \mathcal{N}(0, unstructured)
$$

Model 2_3: General Model structure in INLA

Model 2_3

$$
\begin{cases}
Cy_{i,j} = & (\beta_0^{(Cy)} + u_{0,i}^{(Cy)}) + \beta_{Sc}^{(Cy)} Score_{i,j} + (\beta_t^{(Cy)} + u_{t,i}^{(Cy)}) t_{i,j} + \\
& (\beta_t^{(Cy),28} + u_{t,i}^{(Cy),28})(t_{i,j} - 28)_+ + \\
& (\beta_t^{(Cy),40} + u_{t,i}^{(Cy),40})(t_{i,j} - 40)_+ + \epsilon_{i,j}^2 \\
Sc_{i,j} = & (\beta_0^{(Sc)} + u_{0,i}^{(Sc)}) + \beta_{Cy}^{(Sc)} Cy_{i,j} + (\beta_t^{(Sc)} + u_{t,i}^{(Sc)}) t_{i,j} + \\
& (\beta_t^{(Sc),28} + u_{t,i}^{(Sc),28})(t_{i,j} - 28)_+ + \\
& (\beta_t^{(Sc),14} + u_{t,i}^{(Sc),14})(t_{i,j} - 14)_+ + \epsilon_{i,j}^2
\end{cases}
$$

$$
\left(u_{0,i}^{(Cy)}, u_{t,i}^{(Cy)}, u_{0,i}^{(Sc)}, u_{t,i}^{(Sc)}\right) \sim \mathcal{N}(0, unstructured)
$$

$$
\left(u_{t,i}^{(Cy),28}, u_{t,i}^{(Cy),40}, u_{t,i}^{(Sc),28}, u_{t,i}^{(Sc),14}\right) \sim \mathcal{N}(0, unstructured)
$$

Model 2_0: No breakpoints

$$
\begin{cases}
Cy_{i,j} = & (\beta_0^{(Cy)} + u_{0,i}^{(Cy)}) + \beta_{Sc}^{(Cy)} Score_{i,j} + (\beta_t^{(Cy)} + u_{t,i}^{(Cy)}) t_{i,j} + \\
& \cancel{(\beta_t^{(Cy),28} + u_{t,i}^{(Cy),28})(t_{i,j} - 28)_+} + \\
& \cancel{(\beta_t^{(Cy),40} + u_{t,i}^{(Cy),40})(t_{i,j} - 40)_+} + \epsilon_{i,j}^2 \\
Sc_{i,j} = & (\beta_0^{(Sc)} + u_{0,i}^{(Sc)}) + \beta_{Cy}^{(Sc)} Cy_{i,j} + (\beta_t^{(Sc)} + u_{t,i}^{(Sc)}) t_{i,j} + \\
& \cancel{(\beta_t^{(Sc),28} + u_{t,i}^{(Sc),28})(t_{i,j} - 28)_+} + \\
& \cancel{(\beta_t^{(Sc),14} + u_{t,i}^{(Sc),14})(t_{i,j} - 14)_+} + \epsilon_{i,j}^2
\end{cases}
$$

$$
\left(u_{0,i}^{(Cy)}, u_{t,i}^{(Cy)}, u_{0,i}^{(Sc)}, u_{t,i}^{(Sc)}\right) \sim \mathcal{N}(0, unstructured)
$$

$$
\cancel{\left(u_{t,i}^{(Cy),28}, u_{t,i}^{(Cy),40}, u_{t,i}^{(Sc),28}, u_{t,i}^{(Sc),14}\right) \sim \mathcal{N}(0, unstructured)}
$$

Model 2_1: No breakpoints as random effects

$$
\begin{cases}
Cy_{i,j} = & (\beta_0^{(Cy)} + u_{0,i}^{(Cy)}) + \beta_{Sc}^{(Cy)} Score_{i,j} + (\beta_t^{(Cy)} + u_{t,i}^{(Cy)}) t_{i,j} + \\
& (\beta_t^{(Cy),28} + \cancel{u_{t,i}^{(Cy),28}})(t_{i,j} - 28)_+ + \\
& (\beta_t^{(Cy),40} + \cancel{u_{t,i}^{(Cy),40}})(t_{i,j} - 40)_+ + \epsilon_{i,j}^2 \\
Sc_{i,j} = & (\beta_0^{(Sc)} + u_{0,i}^{(Sc)}) + \beta_{Cy}^{(Sc)} Cy_{i,j} + (\beta_t^{(Sc)} + u_{t,i}^{(Sc)}) t_{i,j} + \\
& (\beta_t^{(Sc),28} + \cancel{u_{t,i}^{(Sc),28}})(t_{i,j} - 28)_+ + \\
& (\beta_t^{(Sc),14} + \cancel{u_{t,i}^{(Sc),14}})(t_{i,j} - 14)_+ + \epsilon_{i,j}^2
\end{cases}
$$

$$
\left(u_{0,i}^{(Cy)}, u_{t,i}^{(Cy)}, u_{0,i}^{(Sc)}, u_{t,i}^{(Sc)}\right) \sim \mathcal{N}(0, unstructured)
$$

$$
\cancel{\left(u_{t,i}^{(Cy),28}, u_{t,i}^{(Cy),40}, u_{t,i}^{(Sc),28}, u_{t,i}^{(Sc),14}\right)} \sim \cancel{\mathcal{N}(0, unstructured)}
$$

Model 2_2: Only breakpoint 28 as random effect

$$
\begin{cases}
Cy_{i,j} = & (\beta_0^{(Cy)} + u_{0,i}^{(Cy)}) + \beta_{Sc}^{(Cy)} Score_{i,j} + (\beta_t^{(Cy)} + u_{t,i}^{(Cy)})t_{i,j} + \\
& (\beta_t^{(Cy),28} + u_{t,i}^{(Cy),28})(t_{i,j} - 28)_+ + \\
& (\beta_t^{(Cy),40} + \cancel{u_{t,i}^{(Cy),40}})(t_{i,j} - 40)_+ + \epsilon_{i,j}^2 \\
Sc_{i,j} = & (\beta_0^{(Sc)} + u_{0,i}^{(Sc)}) + \beta_{Cy}^{(Sc)} Cy_{i,j} + (\beta_t^{(Sc)} + u_{t,i}^{(Sc)})t_{i,j} + \\
& (\beta_t^{(Sc),28} + u_{t,i}^{(Sc),28})(t_{i,j} - 28)_+ + \\
& (\beta_t^{(Sc),14} + \cancel{u_{t,i}^{(Sc),14}})(t_{i,j} - 14)_+ + \epsilon_{i,j}^2
\end{cases}
$$

$$
\left(u_{0,i}^{(Cy)}, u_{t,i}^{(Cy)}, u_{0,i}^{(Sc)}, u_{t,i}^{(Sc)}\right) \sim \mathcal{N}(0, unstructured)
$$

$$
\left(u_{t,i}^{(Cy),28}, u_{t,i}^{(Sc),28}, \cancel{u_{t,i}^{(Cy),40}, u_{t,i}^{(Sc),14}}\right) \sim \cancel{\mathcal{N}(0, unstructured)}
$$

Model 3_3: General Model structure in INLA

$$
\begin{cases}
m_{i,j} = & (\beta_0^{(Cy)} + u_{0,i}^{(Cy)}) + (\beta_t^{(Cy)} + u_{t,i}^{(Cy)})t_{i,j} + \\
& (\beta_t^{(Cy),28} + u_{t,i}^{(Cy),28})(t_{i,j} - 28)_+ + \\
& (\beta_t^{(Cy),40} + u_{t,i}^{(Cy),40})(t_{i,j} - 40)_+ \\
Cy_{i,j} = & m_{i,j} + \epsilon_{i,j}^2 \\
Sc_{i,j} = & \gamma m_{i,j} + (\beta_0^{(Sc)} + u_{0,i}^{(Sc)}) + \beta_{Cy}^{(Sc)} Cy_{i,j} + (\beta_t^{(Sc)} + u_{t,i}^{(Sc)})t_{i,j} + \\
& (\beta_t^{(Sc),28} + u_{t,i}^{(Sc),28})(t_{i,j} - 28)_+ + \\
& (\beta_t^{(Sc),14} + u_{t,i}^{(Sc),14})(t_{i,j} - 14)_+ + \epsilon_{i,j}^2
\end{cases}
$$

$$
\left(u_{0,i}^{(Cy)}, u_{t,i}^{(Cy)}, u_{t,i}^{(Cy),28}, u_{t,i}^{(Cy),40}\right) \sim \mathcal{N}(0, unstructured)
$$

$$
\left(u_{0,i}^{(Sc)}, u_{t,i}^{(Sc)}, u_{t,i}^{(Sc),28}, u_{t,i}^{(Sc),14}\right) \sim \mathcal{N}(0, unstructured)
$$

Models 3

- Similarly to Models 0 and 2 we also introduced Models 3_0, 3_1 and 3_2.
- A problem with Model 3 is that we can not put Score into the linear predictor of Cy
  - $Cy_{i,j} = \beta_{Sc}^{(Cy)} Score_{i,j}$ is not possible
  - INLA then simply ensures that $\gamma \cdot \beta_{Sc}^{(Cy)} = 1$, and all other terms are set to 0.
  - This problem can be averted when using Lagged values!
- Fitting these models takes a lot of computing power!
  - I therefore do a lot of the programming on the LUMC-test computer, is that allowed?
  - Can I install git on the LUMC computer?

General form of Model 3

- Association supplied via γ.

$$
\begin{cases}
m_{ij} = (\beta_0^{(x)} + u_{0,i}^{(x)}) + \beta_v^{(x)} \cdot v_i + \left(\beta_t^{(x)} + u_{t,i}^{(x)}\right) \cdot t_{i,j} \\
x_{i,j} = m_{ij} + \epsilon_{i,j}^{(x)} \\
y_{i,j} = \gamma \cdot m_{ij} + (\beta_0^{(y)} + u_{0,i}^{(y)}) + \beta_v^{(y)} \cdot v_i + \left(\beta_t^{(y)} + u_{t,i}^{(y)}\right) \cdot t_{i,j} + \epsilon_{i,j}^{(y)}
\end{cases}
$$

with

$$
\begin{bmatrix} u_{0,i}^{(x)} \\ u_{t,i}^{(x)} \end{bmatrix} \sim \mathcal{N}_2 \left[\mathbf{0}, \begin{pmatrix} \sigma_{x,0}^2 & \sigma_{x,(0,t)} \\ \sigma_{x,(t,0)} & \sigma_{x,t}^2 \end{pmatrix} \right]; \quad \begin{bmatrix} u_{0,i}^{(y)} \\ u_{t,i}^{(y)} \end{bmatrix} \sim \mathcal{N}_2 \left[\mathbf{0}, \begin{pmatrix} \sigma_{y,0}^2 & \sigma_{y,(0,t)} \\ \sigma_{y,(t,0)} & \sigma_{y,t}^2 \end{pmatrix} \right];
$$

General form of Model 3

- This can be rewritten as:

$$
\begin{aligned}
y_{i,j} &= \gamma \cdot m_{ij} + (\beta_0^{(y)} + u_{0,i}^{(y)}) + \beta_v^{(y)} \cdot v_i + \left(\beta_t^{(y)} + u_{t,i}^{(y)}\right) \cdot t_{i,j} + \epsilon_{i,j}^{(y)} = \\
&= (\gamma \beta_0^{(x)} + \beta_0^{(y)}) + (\gamma \beta_v^{(x)} + \beta_v^{(y)}) v_i + \left(\gamma \beta_t^{(x)} + \beta_t^{(y)}\right) \cdot t_{i,j} + (\gamma u_{0,i}^{(x)} + u_{0,i}^{(y)}) + \\
&\quad + (\gamma u_{t,i}^{(x)} + u_{t,i}^{(y)}) t_{i,j} + \epsilon_{i,j}^{(y)} = \\
&= \beta_0^{(y)'} + \beta_v^{(y)'} v_i + \beta_t^{(y)'} t_{i,j} + u_{0,i}^{(y)'} + u_{t,i}^{(y)'} \cdot t_{i,j} + \epsilon_{i,j}^{(y)}
\end{aligned}
$$

with $\beta_0^{(y)'} = \gamma \beta_0^{(x)} + \beta_0^{(y)}$, $\beta_v^{(y)'} = \gamma \beta_v^{(x)} + \beta_v^{(y)}$ and $\beta_t^{(y)'} = \gamma \beta_t^{(x)} + \beta_t^{(y)}$.
For the random effects, they now have distribution:

$$
\begin{bmatrix} u_{0,i}^{(y)'} \\ u_{t,i}^{(y)'} \end{bmatrix} \sim \mathcal{N}_2 \left[\mathbf{0}, \begin{pmatrix} \gamma^2 \sigma_{x,0}^2 + \sigma_{y,0}^2 & \gamma^2 \sigma_{x,(0,t)} + \sigma_{y,(0,t)} \\ \gamma^2 \sigma_{x,(t,0)} + \sigma_{y,(t,0)} & \gamma^2 \sigma_{x,t}^2 + \sigma_{y,t}^2 \end{pmatrix} \right]
$$

General form of Model 3

- We obtain:

$$x_{i,j} = (\beta_0^{(x)} + u_{0,i}^{(x)}) + \beta_v^{(x)} \cdot v_i + \left(\beta_t^{(x)} + u_{t,i}^{(x)}\right) \cdot t_{i,j} + \epsilon_{i,j}^{(x)}$$

$$y_{i,j} = (\beta_0^{(y)'} + u_{0,i}^{(y)'}) + \beta_v^{(y)'} v_i + (\beta_t^{(y)'} + u_{t,i}^{(y)'}) t_{i,j} + \epsilon_{i,j}^{(y)}$$

with

$$\begin{bmatrix} u_{0,i}^{(x)} \\ u_{t,i}^{(x)} \end{bmatrix} \sim \mathcal{N}_2 \left[\mathbf{0}, \begin{pmatrix} \sigma_{x,0}^2 & \sigma_{x,(0,t)} \\ \sigma_{x,(t,0)} & \sigma_{x,t}^2 \end{pmatrix} \right],$$

$$\beta_0^{(y)'} = \gamma\beta_0^{(x)} + \beta_0^{(y)}, \quad \beta_v^{(y)'} = \gamma\beta_v^{(x)} + \beta_v^{(y)} \text{ and } \beta_t^{(y)'} = \gamma\beta_t^{(x)} + \beta_t^{(y)}$$

$$\begin{bmatrix} u_{0,i}^{(y)'} \\ u_{t,i}^{(y)'} \end{bmatrix} \sim \mathcal{N}_2 \left[\mathbf{0}, \begin{pmatrix} \gamma^2\sigma_{x,0}^2 + \sigma_{y,0}^2 & \gamma^2\sigma_{x,(0,t)} + \sigma_{y,(0,t)} \\ \gamma^2\sigma_{x,(t,0)} + \sigma_{y,(t,0)} & \gamma^2\sigma_{x,t}^2 + \sigma_{y,t}^2 \end{pmatrix} \right]$$

Lagged effects

- Via multiple γ's
- Via history of outcome/endogenous covariate

Multiple γ's

$$
\begin{cases}
m_{ij} = (\beta_0^{(x)} + u_{0,i}^{(x)}) + \beta_v^{(x)} \cdot v_i + \left(\beta_t^{(x)} + u_{t,i}^{(x)}\right) \cdot t_{i,j} \\
x_{i,j} = m_{ij} + \epsilon_{i,j}^{(x)} \\
x_{i,(j-1)} = m_{i,(j-1)} + \epsilon_{i,(j-1)}^{(x)} \\
y_{i,j} = \gamma_1 \cdot m_{ij} + \gamma_2 \cdot m_{i,(j-1)} + (\beta_0^{(y)} + u_{0,i}^{(y)}) + \beta_v^{(y)} \cdot v_i + \left(\beta_t^{(y)} + u_{t,i}^{(y)}\right) \cdot t_{i,j} + \epsilon_{i,j}^{(y)}
\end{cases}
$$

- Here the same problem arises as in model with single γ.

Multiple γ's

- We obtain:

$$x_{i,j} = (\beta_0^{(x)} + u_{0,i}^{(x)}) + \beta_v^{(x)} \cdot v_i + \left(\beta_t^{(x)} + u_{t,i}^{(x)}\right) \cdot t_{i,j} + \epsilon_{i,j}^{(x)}$$

$$y_{i,j} = (\beta_0^{(y)'} + u_{0,i}^{(y)'}) + \beta_v^{(y)'} v_i + (\beta_t^{(y)'} + u_{t,i}^{(y)'})t_{i,j} + \epsilon_{i,j}^{(y)}$$

with

$$\beta_0^{(y)'} = (\gamma_1 + \gamma_2)\beta_0^{(x)} + \beta_0^{(y)}, \quad \beta_v^{(y)'} = (\gamma_1 + \gamma_2)\beta_v^{(x)} + \beta_v^{(y)}$$

and $\beta_t^{(y)'} = (\gamma_1 + \gamma_2)\beta_t^{(x)} + \beta_t^{(y)}$

Variable History

- Can be implemented via e.g. AR process

$$\begin{cases} x_{i,j} = (\beta_0^{(x)} + u_{0,i}^{(x)}) + \beta_v^{(x)} \cdot v_i + \left(\beta_t^{(x)} + u_{t,i}^{(x)}\right) \cdot t_{i,j} + \sum_{t=1}^{p} \phi_t x_{j-t} + \epsilon_{i,j}^{(x)} \\ y_{i,j} = (\beta_0^{(y)} + u_{0,i}^{(y)}) + \beta_v^{(y)} \cdot v_i + \left(\beta_t^{(y)} + u_{t,i}^{(y)}\right) \cdot t_{i,j} + \sum_{t=1}^{p} \phi_t y_{j-t} + \epsilon_{i,j}^{(y)} \end{cases}$$

- One can set the depth of the history by varying p
- One can set requirements for ϕ, e.g: $\phi_t = \phi^t$.
- Can also be combined with model 3

Model comparison

- We see that overall Model 0_3 performs best
  - No association between the endogenous covariate and the outcome
  - Maximum number of breakpoints
  - Maximum number of random effects, maximum covariance between random effects

	0_0	0_1	0_2	0_3	2_0	2_1	2_2	2_3	3_0	3_1	3_2	3_3
MLIK	-2399	-2404	-2420	-2411	-2402	**-2394**	-2417	-2418	-2406	-2404	-2417	-2407
DIC_approx	4798	4808	4840	4821	4804	**4789**	4834	4836	4811	4809	4834	4813
DIC	4108	4088	4087	**4024**	4114	4074	4039	4040	4136	4137	4131	4060
WAIC	4103	4085	4081	**4013**	4109	4072	4034	4031	4129	4132	4123	4046
PIT	0.0394	0.0315	0.0255	0.0254	0.0380	0.0317	0.0324	0.0279	0.0290	**0.0252**	0.0277	0.0291
CPO	2083	2072	2074	**2046**	2081	2061	2049	2054	2093	2092	2092	2060
MSE	3.61	3.70	4.03	**3.52**	3.73	3.91	3.71	3.70	3.60	3.68	4.06	**3.52**

To-do

- Focus should be on lagged effects!
  - Using Lagged functional form
    - Using $\gamma_2 \cdot m_i(t_{i,j-1})$
    - Using autoregressive model?
  - Using other residual error covariance structures
- Inspect different time-scales and time-models

- H. Rue, S. Martino, Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations, 2009, JRSS B
- J. van Niekerk, H. Bakka, H. Rue, Joint models as latent Gaussian models - not reinventing the wheel, 2019, arXiv
- J. van Niekerk, H. Bakka, H. Rue, Competing risks joint models using R-INLA, 2021, Statistical Modelling
- R. Weiss, Modelling Longitudinal Data, 2005, Chapter 13