Concepts of multilevel, longitudinal, and mixed models: Group 3

Robrecht Van Der Bauwhede Renée Blanckaert Charlotte Vercammen
Raïsa Carmen

Contents

1 Introduction

The data set results from a longitudinal observational study, the aim of which is to study the post-operative evolution of the cognitive status of elderly hipfracture patients and their pre-operative cognitive status, and to study the effects of housing situation and age on these evolutions. The physical ability is measured using the MMSE (Mini Mental State Examination) score, with values between 0 and 30, where low values correspond to a bad cognitive condition, while high scores correspond to high cognitive condition of the patient. The pre-operative cognitive status is measured through the so-called 'neuro-status' which is a binary indicator for being neuro-psychiatric.

- id: patient identification number
- age: age of the patient at entry
- neuro: neuro-psychiatric status of the patient (1: neuro-psychiatric, 0: not neuro-psychiatric)
- mmse: MMSE score
- time: day after operation at which the MMSE score has been measured (1, 3, 5, 8, or 12)
- housing: the housing situation prior to the hip fracture (1: alone, 2: with family or partner, 3: nursing home)

2 Data exploration

Considering the completeness of the data, it can be observed that 5 persons' housing situation is unknown. Additionally, there was some dropout over time as show in Table 1. The column "Nb" shows the number of respondent at each time instance. The column "Return" shows the number of respondents that participated at time t while they did not participate at time $t-1$; there is, for instance, one respondent that did not participate at $time = 1$ while he did participate at $time = 3$. The column "Dropout" shows the number of respondents that did not participate at time t while they did participate at time $t-1$. It can be seen than

Table 1: Dropout and patient characteristics over time.

Time	Nb	Return	Dropout	Mean age	% neuro-psychiatric	Housing			
						%alone	%family/ partner	%nursing home	%NA
1	58	0	0	78.71	31.03	29.31	39.66	22.41	8.62
3	57	1	2	78.18	33.33	29.82	38.60	22.81	8.77
5	59	2	0	78.59	32.20	28.81	38.98	23.73	8.47
8	52	0	7	77.88	30.77	28.85	36.54	25.00	9.62
12	38	0	14	77.82	28.95	31.58	34.21	23.68	10.53

many drop out at $time = 12$. The other columns show how the patient characteristics change over time as patients are added or lost from the study. Overall, there is not much variation which might indicate that the dropout of patients is not related to MMSE or patient characteristics. We will thus assume dropout is completely random.

Figure 1 shows the average evolution of MMSE over time (Loess curves) for groups of patients with different housing and/or neuro-psychiatric status. Patients that are not neuro-psychiatric seem to have higherMMSE that stays reasonably constant over time (except for nursing home patients). MMSE seems to go up over time for nursing home patients (but they are the smallest group), and down for neuro-spychiatric patients that live with their family or partner (or unknown housing). MMSE for neuro-psychiatric patients that live alone might have a quadratic evolution over time.

Figure 2 shows all patient profiles. There is quite a lot of variation between the patients' evolution. The non-psychiatric patients are also under-represented (except in the nursing home group).

The last variable that is explored, is the age. The loess curves in Figure 3 show that age might be negatively correlated with MMSE.

2.1 Variance analysis and correlation structure

Figure 4 shows that the variance is larger for the neuro-psychiatric patients (the smallest group).

Figure 5 shows that there seems to be an inverse relationship between the mean mmse and the variance for patients with either neuro-psychiatric status.

2.2 Conclusion exploratory analysis

The exploratory analysis has shows that the pattern of mmse over time is likely not constant or even linear. One might attempt quadratic, cubic, or logarithmic transformations of time to accommodate this. Both the level (intercept) of mmse differs quite a lot, depending on neuro-psychiatric status, housing and age.

There seems to be high intraclass correlation???

3 Methodology and results

This section, gradually develops a statistical model that seems to fit the data best. It starts with a simple model where only limited covariates are included in 3.1

3.1 A simple model

This first, simple model will assume a linear relationship between mmse and $log(time)$ for each patient. The model allows for subject-specific intercepts and slopes and the neuro-psychiatric status is the only additional explanatory variable that is taken into account.

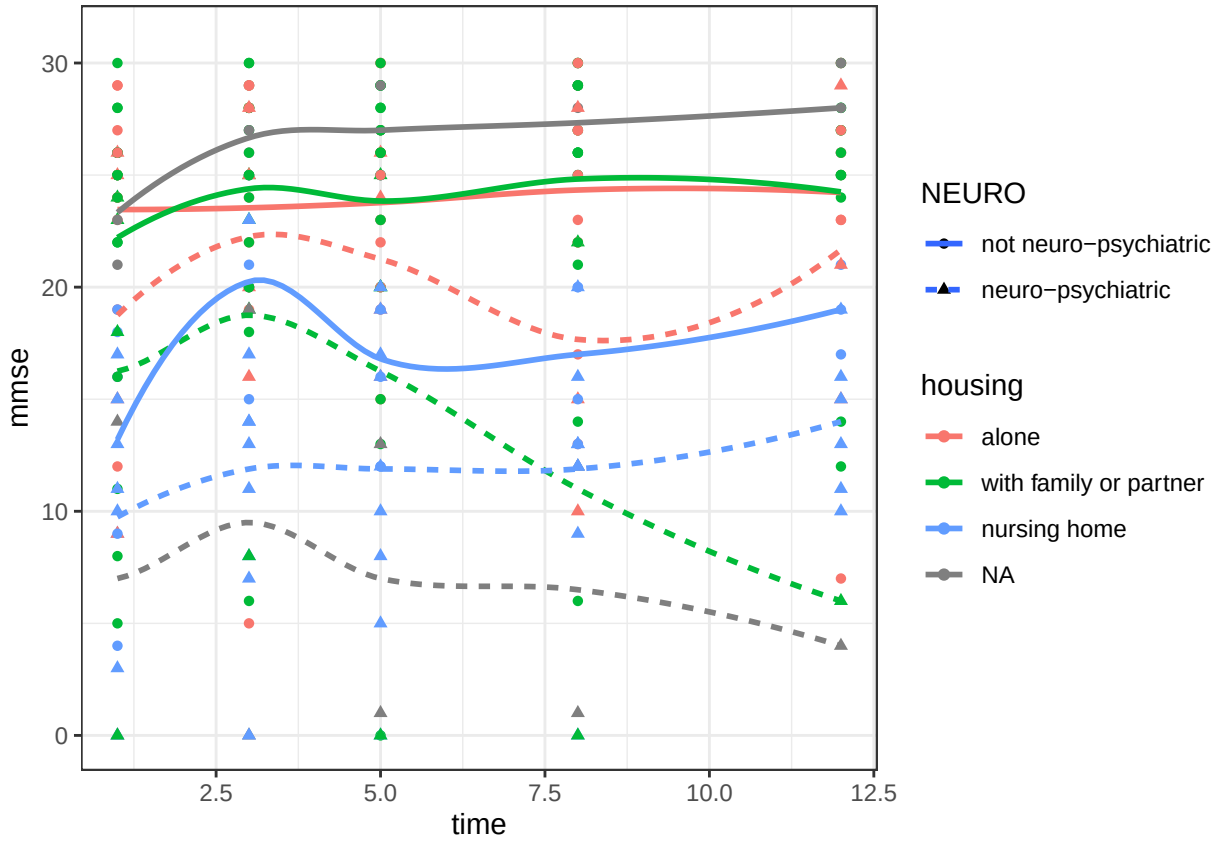

Figure 1: MMSE over time for all patient profiles, including smoothed curves.

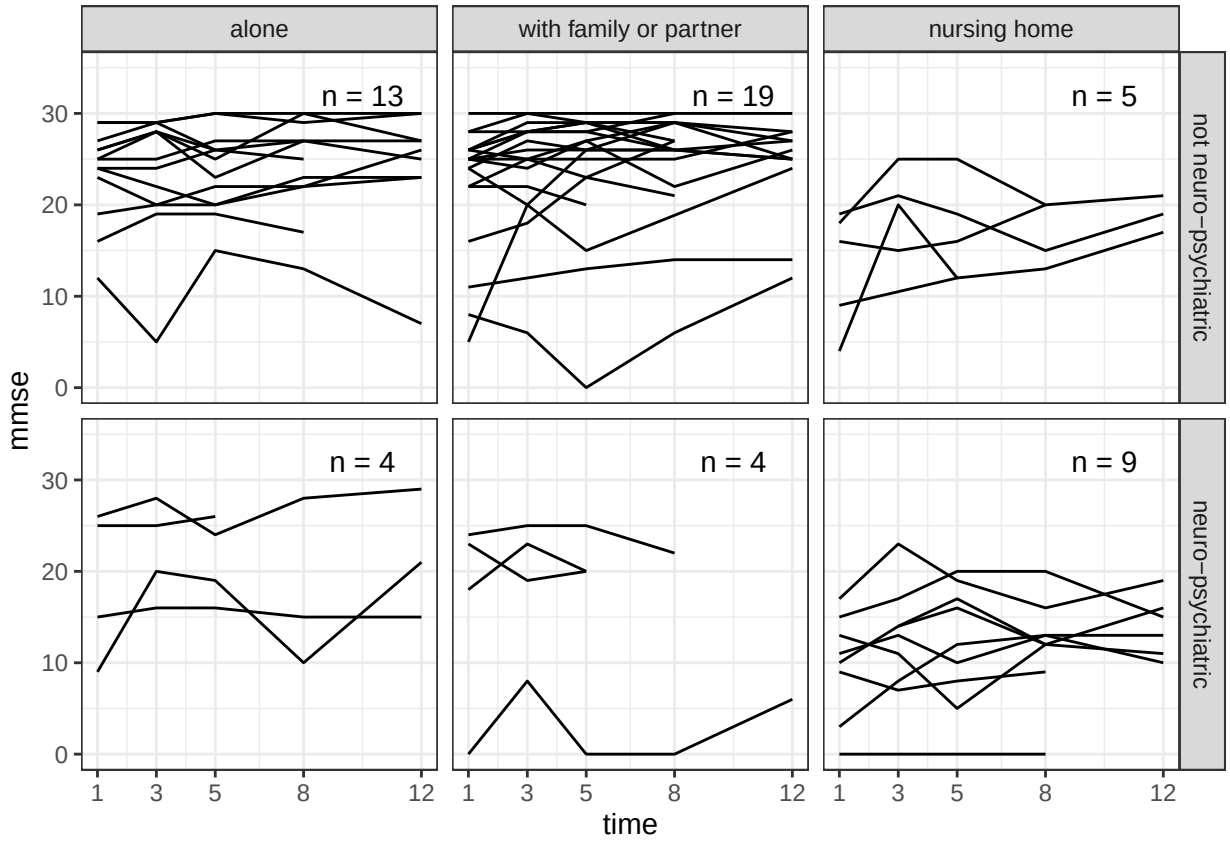

Figure 2: Patient profiles of MMSE over time.

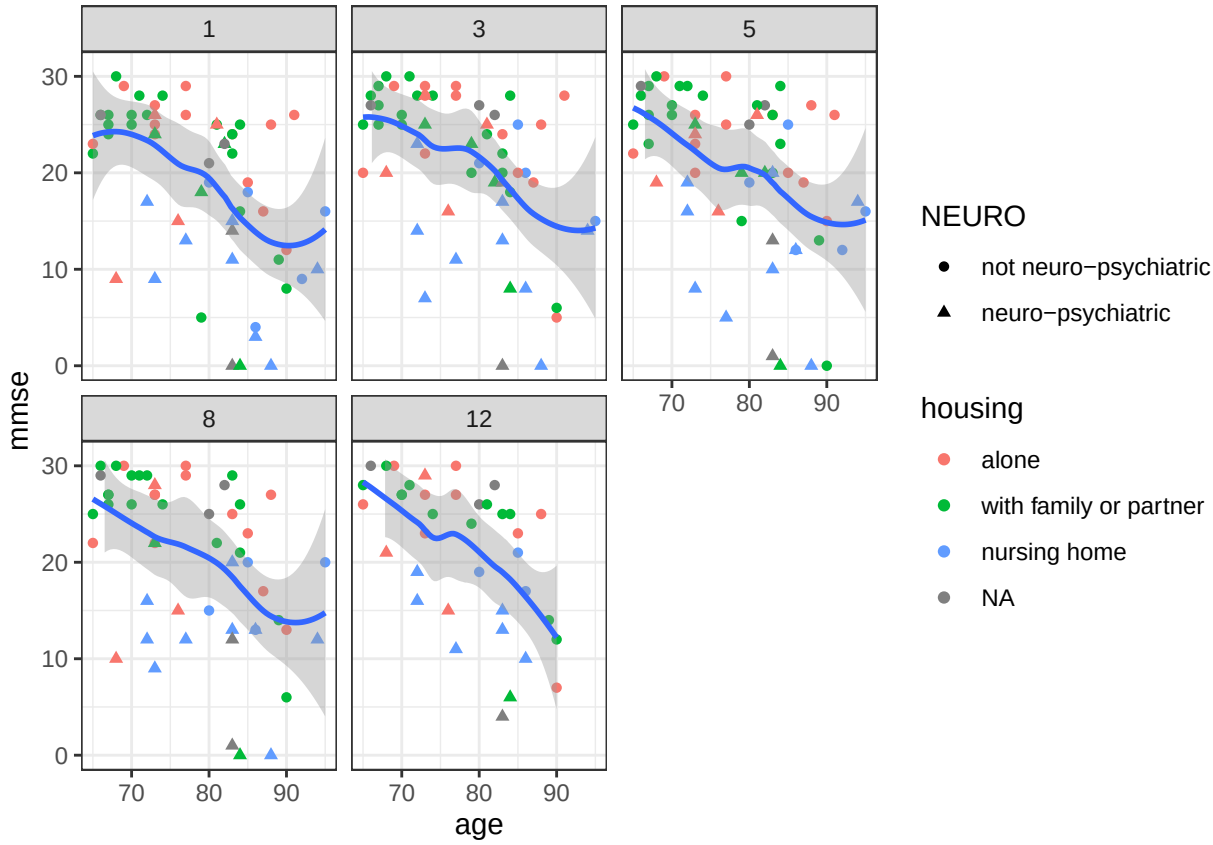

Figure 3: MMSE versus age. Facets show the time instances.

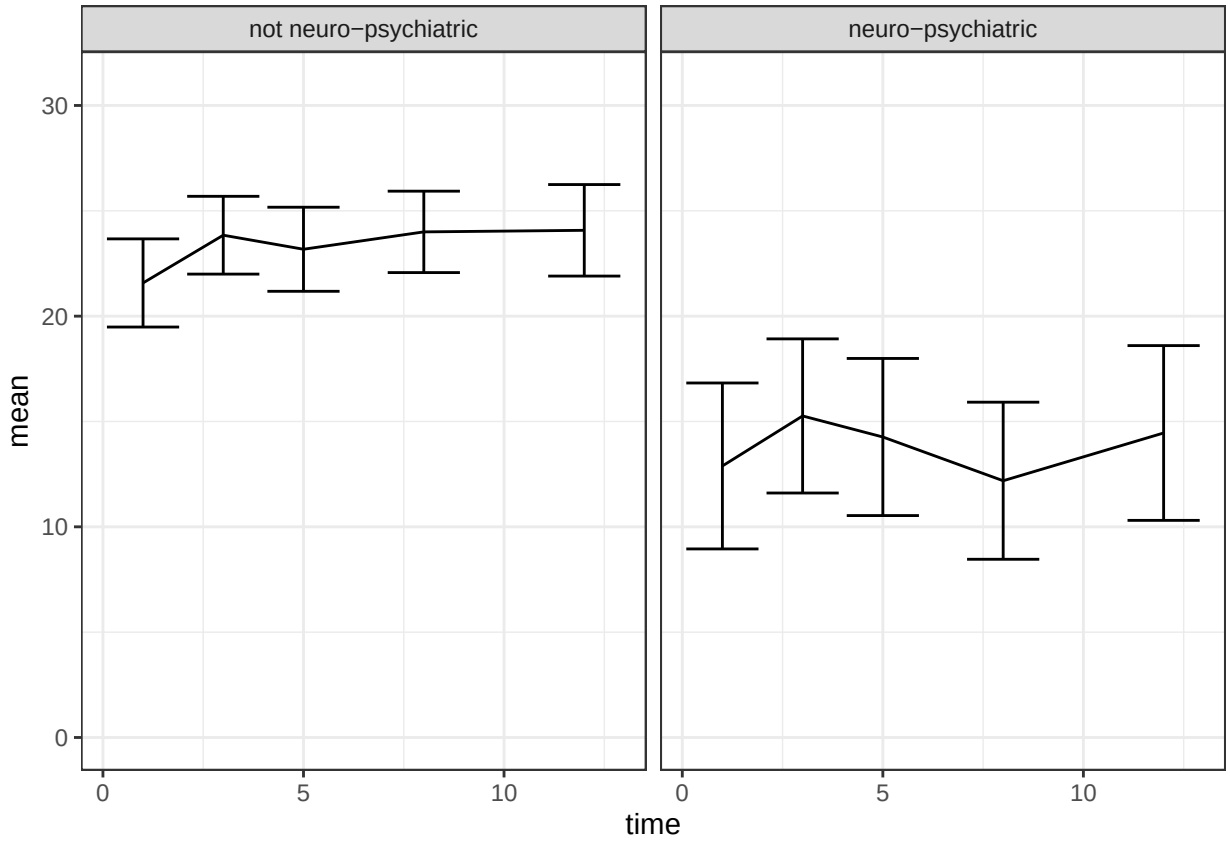

Figure 4: Error bars with 95 percent confidence intervals.

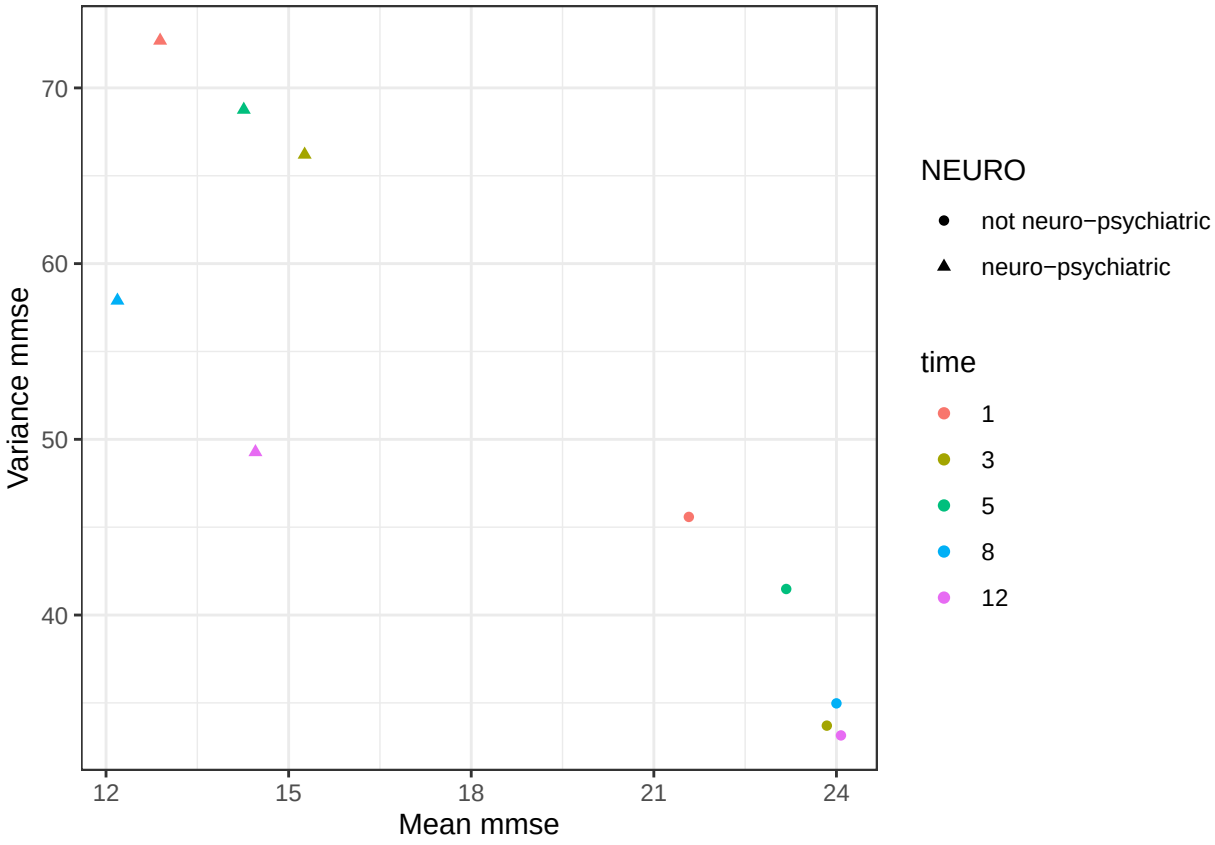

Figure 5: Variance and mean of the mmse, plotted for each of the groups.

Table 2: Fixed effects for the simple model with random intercept and slope

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	21.762	1.142	56.891	19.065	0.000
logtime	0.953	0.227	53.841	4.194	0.000
NEUROneuro-psychiatric	-8.381	2.016	57.314	-4.158	0.000
logtime:NEUROneuro-psychiatric	-0.327	0.415	57.976	-0.788	0.434

Table 3: variance-coviarance matrix.

	(Intercept)	logtime
(Intercept)	47.383	-2.370
logtime	-2.370	0.203

Table 2 shows the fixed effects for the simple model where time is on a log scale, with a random intercept and slope. For the average not-neuro-psychiatric patient, the slope is positive and the intercept is at 21.7624198. For the average neuro-psychiatric patients, the intercept is at 13.3817701 and the slope is less steep (as shown in Figure 7) but the difference in slope is not significant.

Figure 6 shows that the model performs reasonably well; there are no real remaining trends (all curves going up or down) and most residuals seem to randomly fluctuate around zero (no profiles are consistently above or below zero). There are, however, some individuals whose residuals are far larger (both positive and negative) than others.

Figure 8 shows the inverse relationship between the slope and the intercept: the higher the intercept, the lower the slope for each of the neuro-psychiatric statuses. This makes sense because MMSE is not a real continuous variable. It has a maximum value of 30 which means that there is less room for an increase in MMSE (meaning a lower slope) if the starting value (intercept) for MMSE is already high.

```
## $id
```

Table 4: Standard deviations.

	x
(Intercept)	6.884
logtime	0.451

Table 5: Correlations.

	(Intercept)	logtime
(Intercept)	1.000	-0.764
logtime	-0.764	1.000

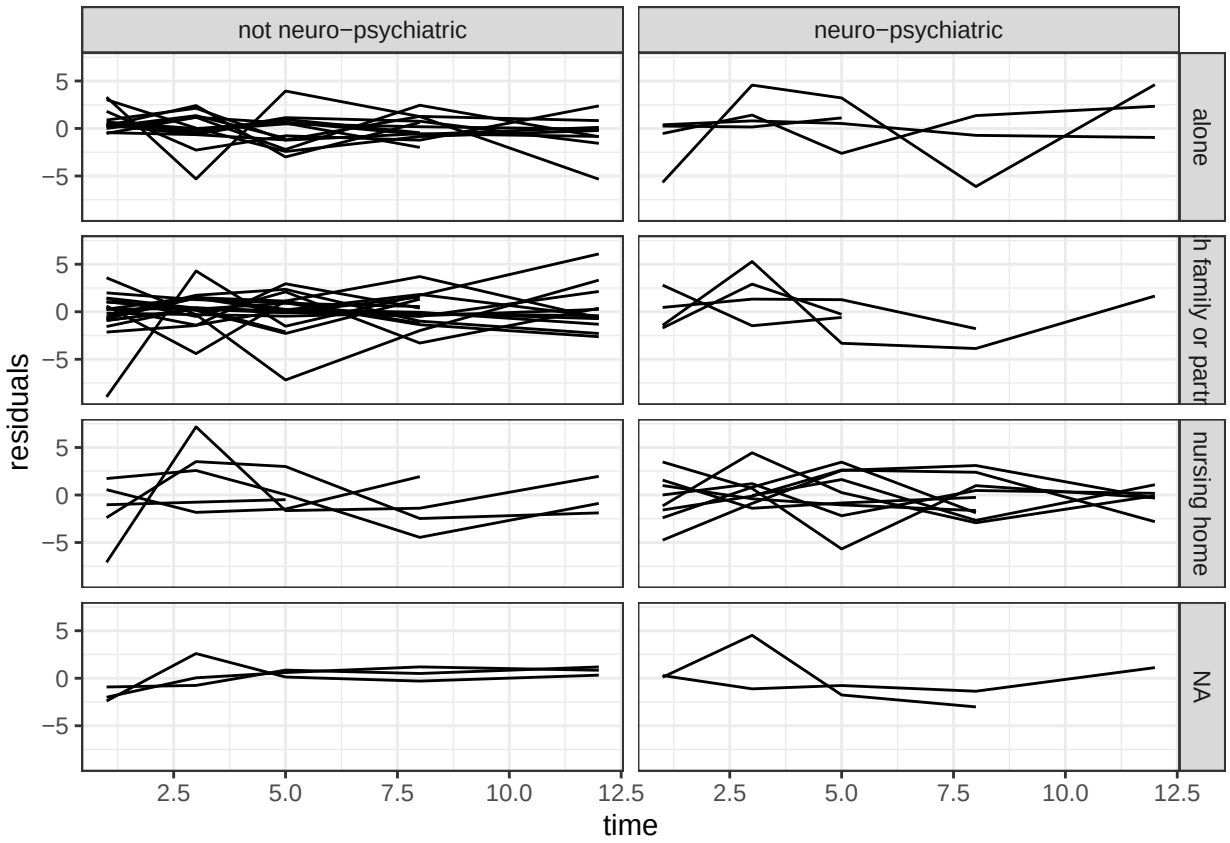

Figure 6: Residuals and time.

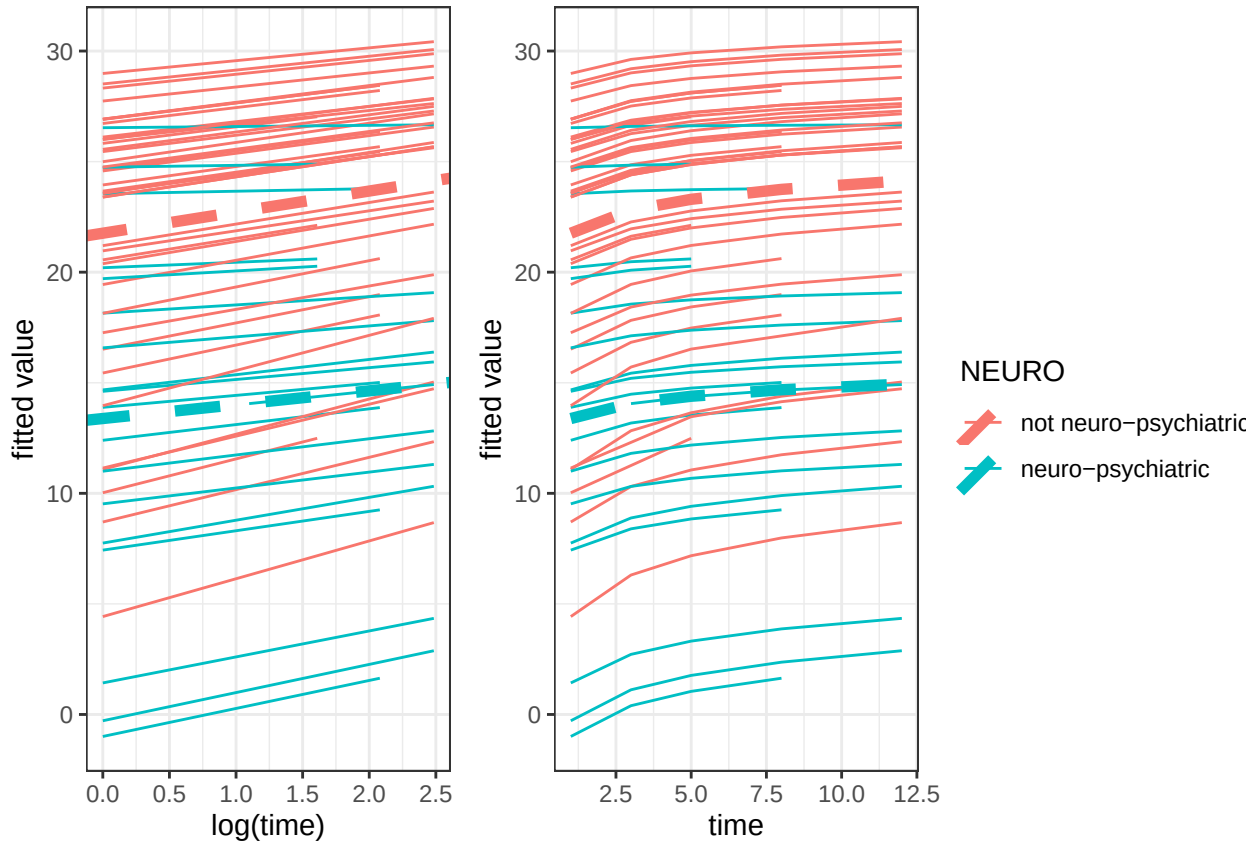

Figure 7: Fitted values plotted against log(time) and time. Dashed, thicker lines are the predicted trends based on the fixed effects only, for each group.

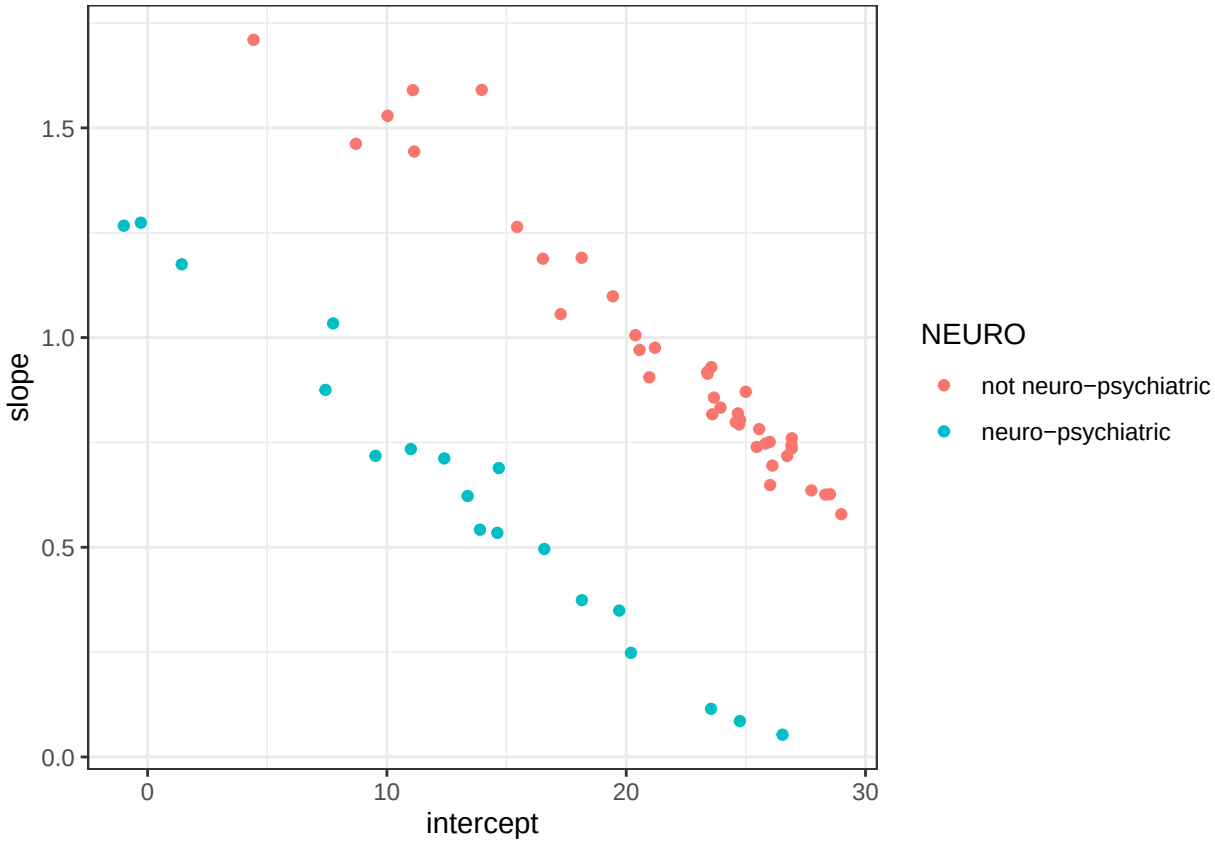

Figure 8: Scatterplot of the fitted intercept and slopes (log(time)).

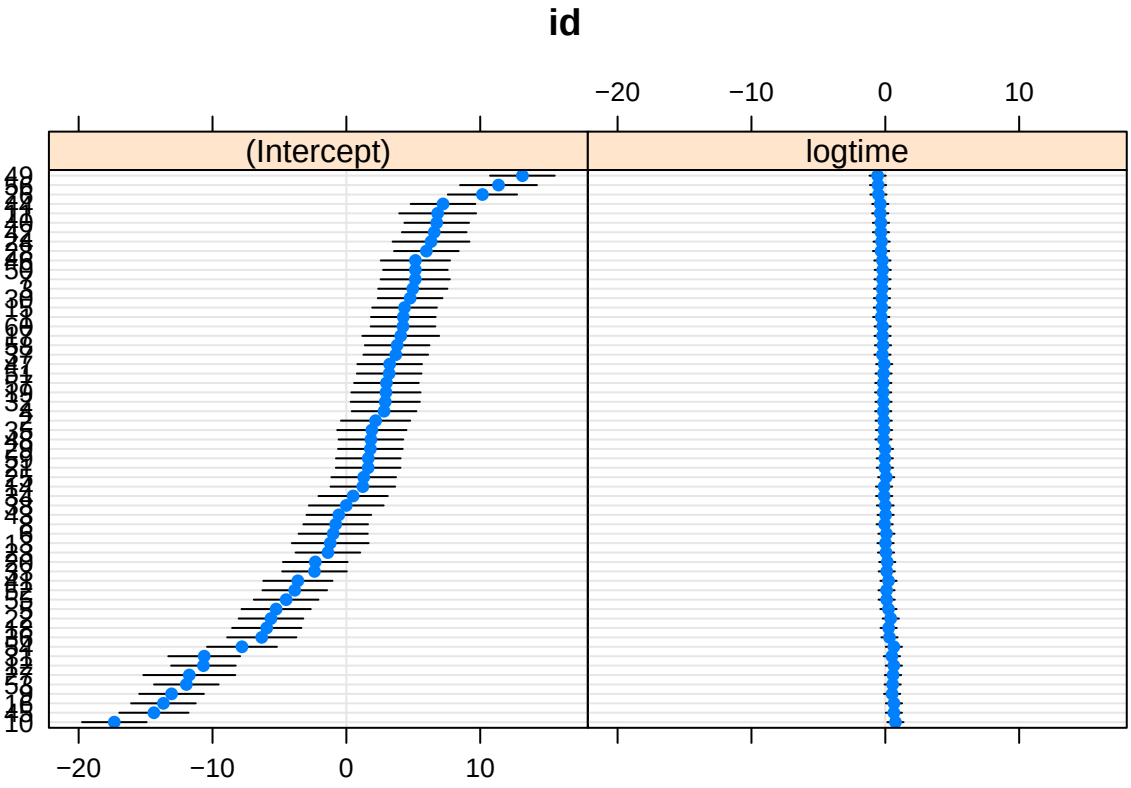

Figure 9: This shows both random effects in the model for each subject (y-axis). It is clear that there is very little variation in the slopes.

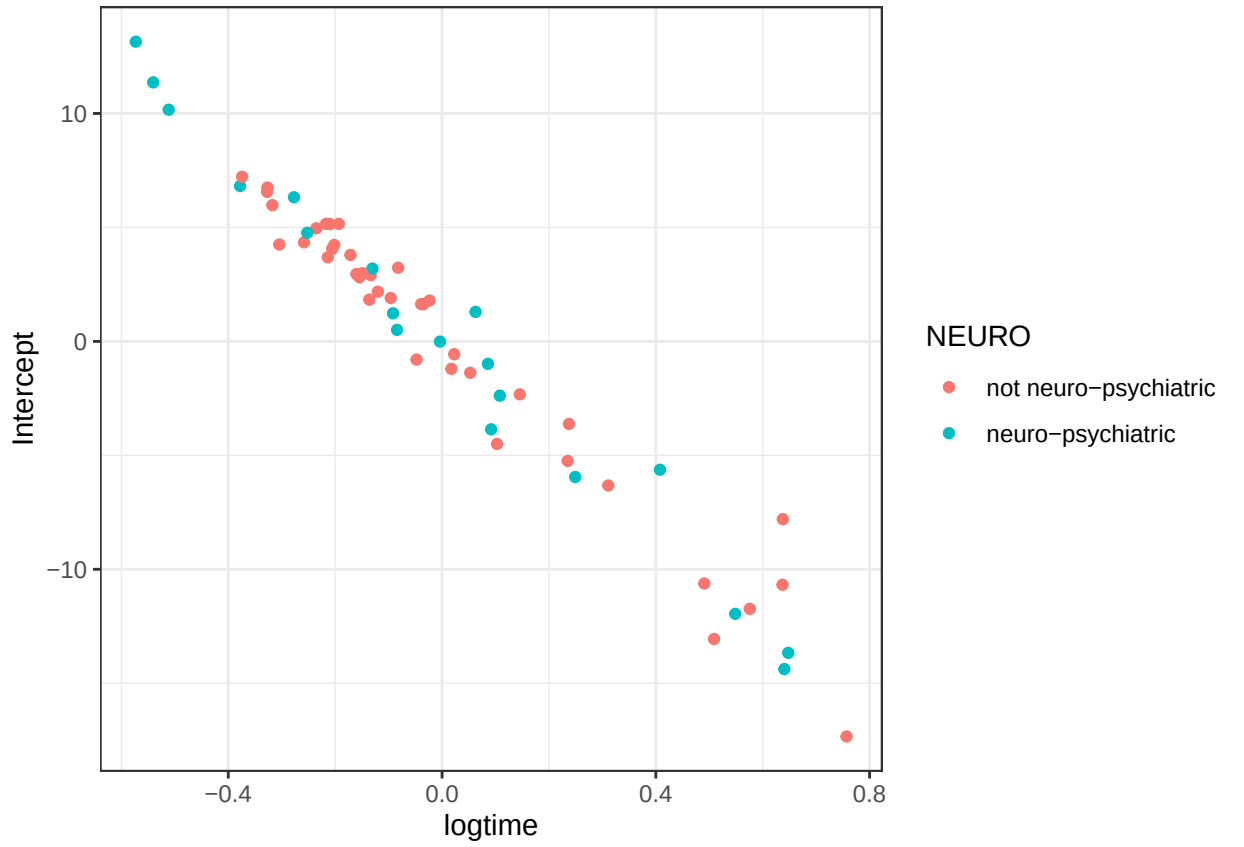

Figure 10: This shows both random effects in the model. notice that there is very little variation in the slope.

Table 6: Fixed effects for the simple model with random intercept and random neuro effect

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	21.767	0.976	46.909	22.303	0.000
logtime	0.949	0.219	203.477	4.334	0.000
NEUROneuro-psychiatric	-8.414	2.084	32.110	-4.037	0.000
logtime:NEUROneuro-psychiatric	-0.276	0.401	203.538	-0.688	0.492

Table 7: variance-coviarance matrix for the simple model with random intercept and random neuro effect.

	(Intercept)	NEUROneuro-psychiatric
(Intercept)	33.225	8.367
NEUROneuro-psychiatric	8.367	9.316

3.1.1 Model with random neuro effect and random intercept

```
## $id
```

Table 8: Standard deviations.

	x
(Intercept)	5.764
NEUROneuro-psychiatric	3.052

Table 9: Correlations.

	(Intercept)	NEUROneuro-psychiatric
(Intercept)	1.000	0.476
NEUROneuro-psychiatric	0.476	1.000

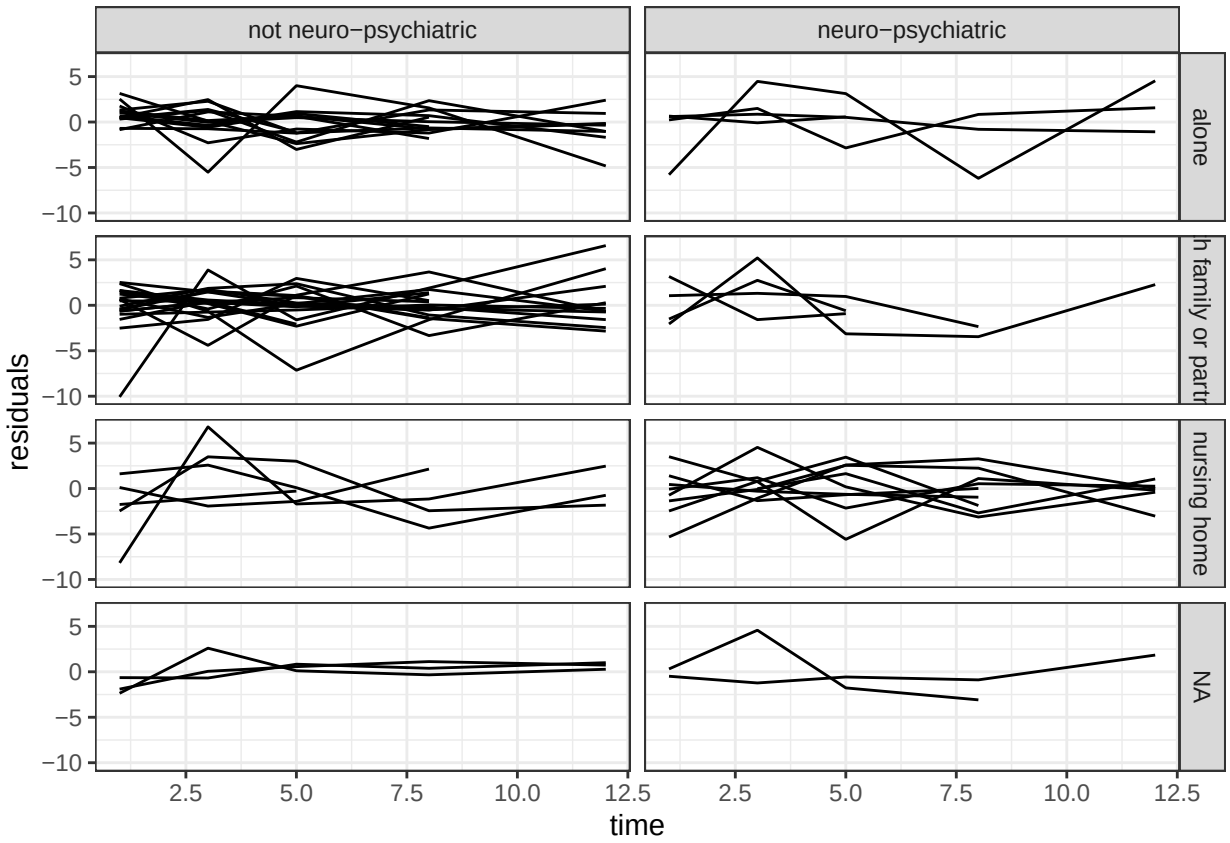

Figure 11: Residuals and time for the simple model with random intercept and random neuro effect.

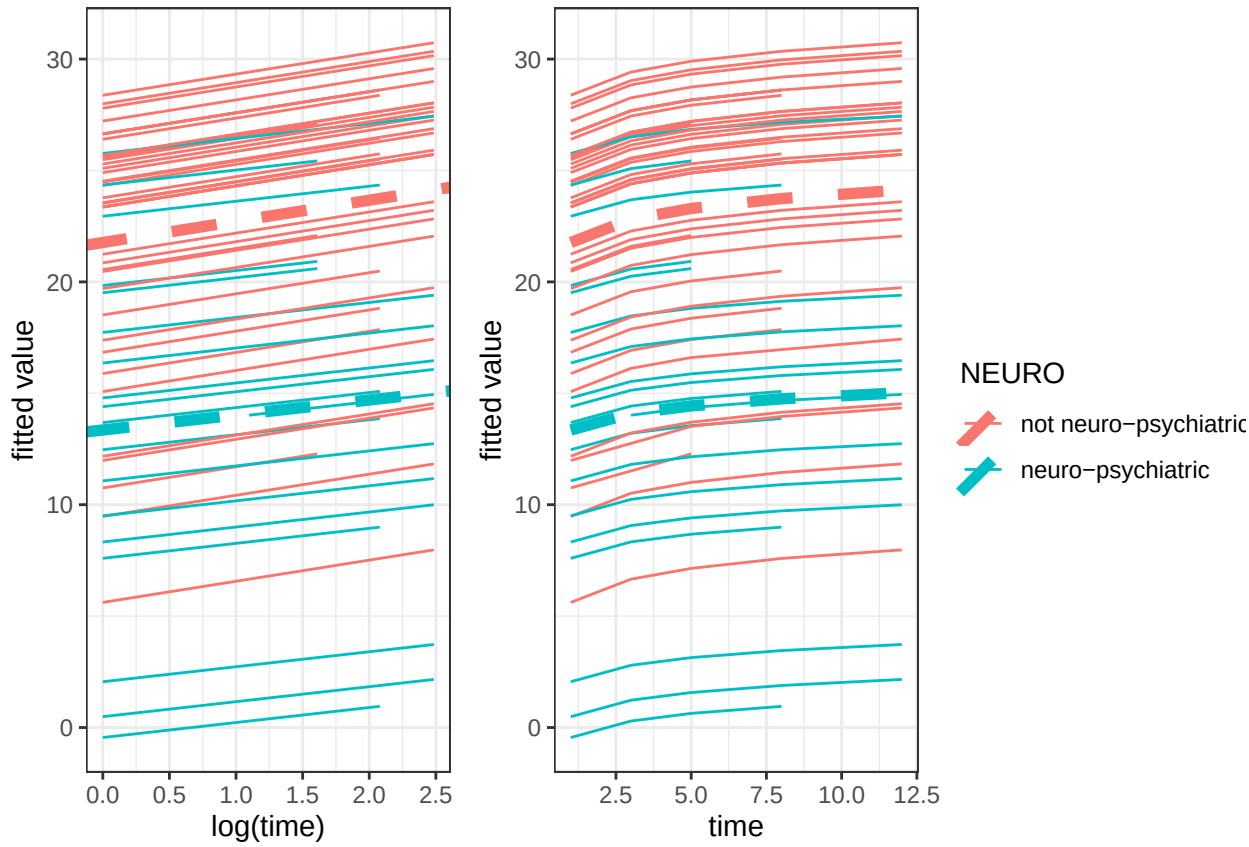

Figure 12: Fitted values for the simple model with random intercept and random neuro effect plotted against log(time) and time. Dashed, thicker lines are the predicted trends based on the fixed effects only, for each group.

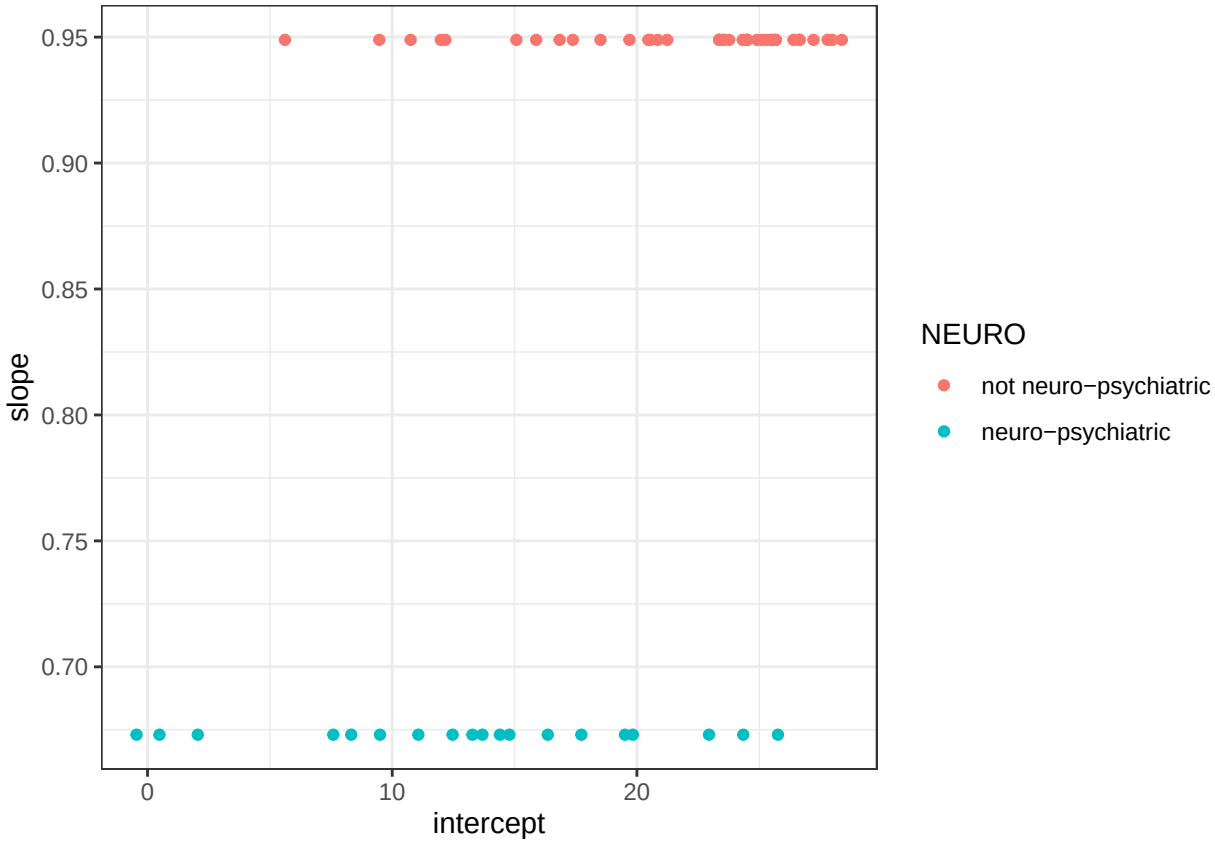

Figure 13: Scatterplot of the fitted intercept and slopes (log(time)).

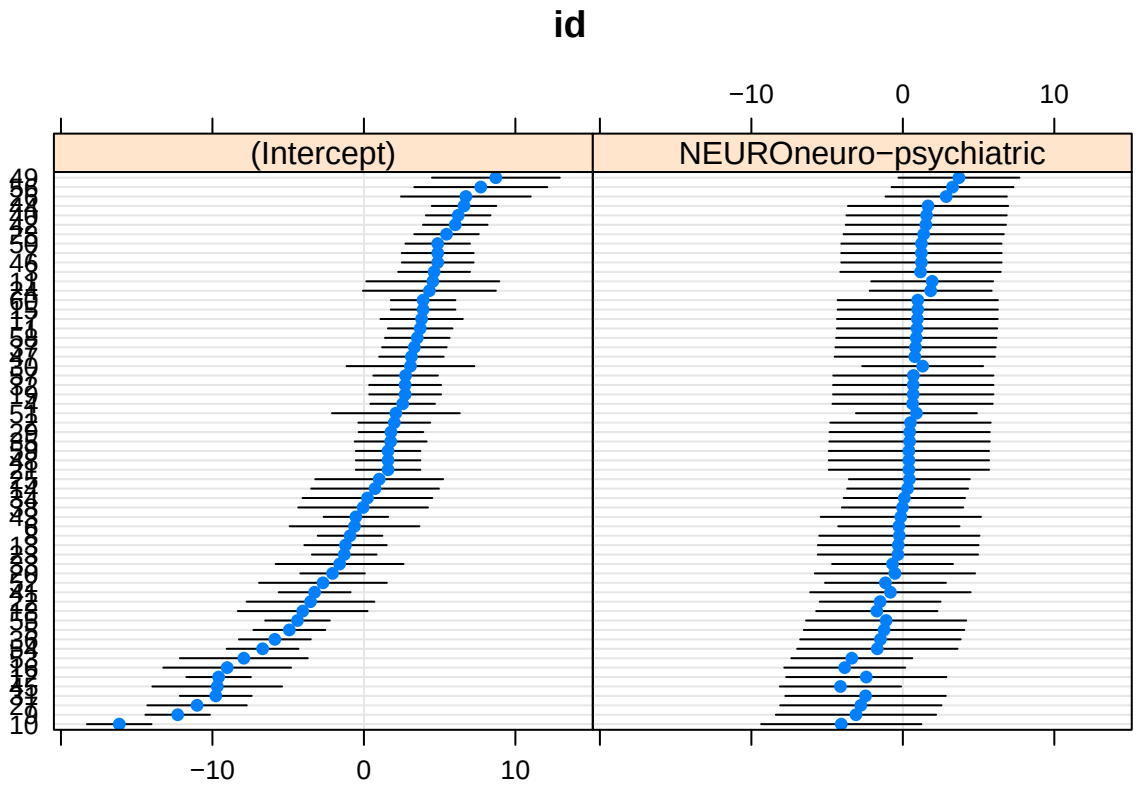

Figure 14: This shows both random effects in the model for each subject (y-axis). It is clear that there is very little variation in the slopes.

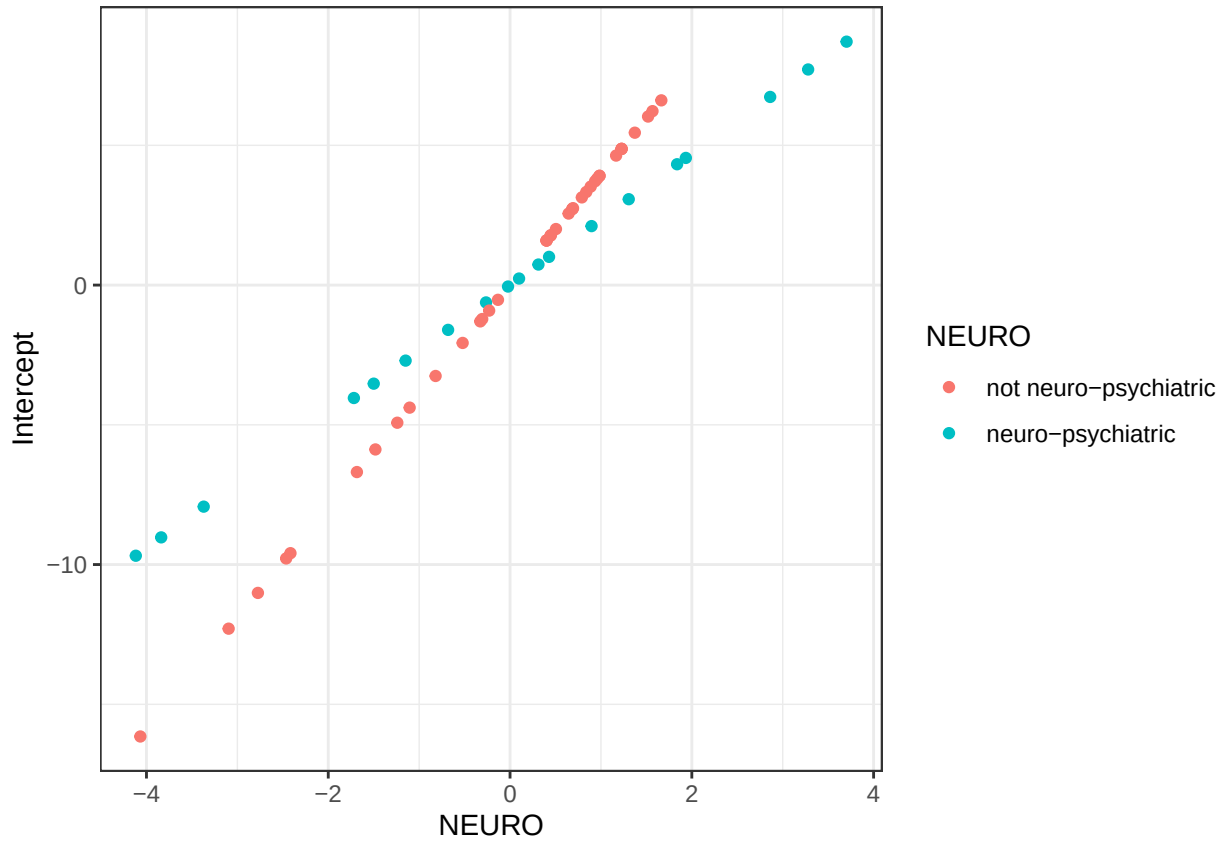

Figure 15: This shows both random effects in the model for each subject (y-axis). It is clear that there is very little variation in the slopes.

3.2 Model with housing and age

3.2.1 Full model with interactions, random slope and intercept

3.2.2 Full model with interactions, random neuro effect and random intercept

This model failed to converge with one negative eigenvalue. Model is nearly unidentifiable: large eigenvalue ratio - Rescale variables? The results, however, are the same as Sas's output

3.2.3 Reduced model

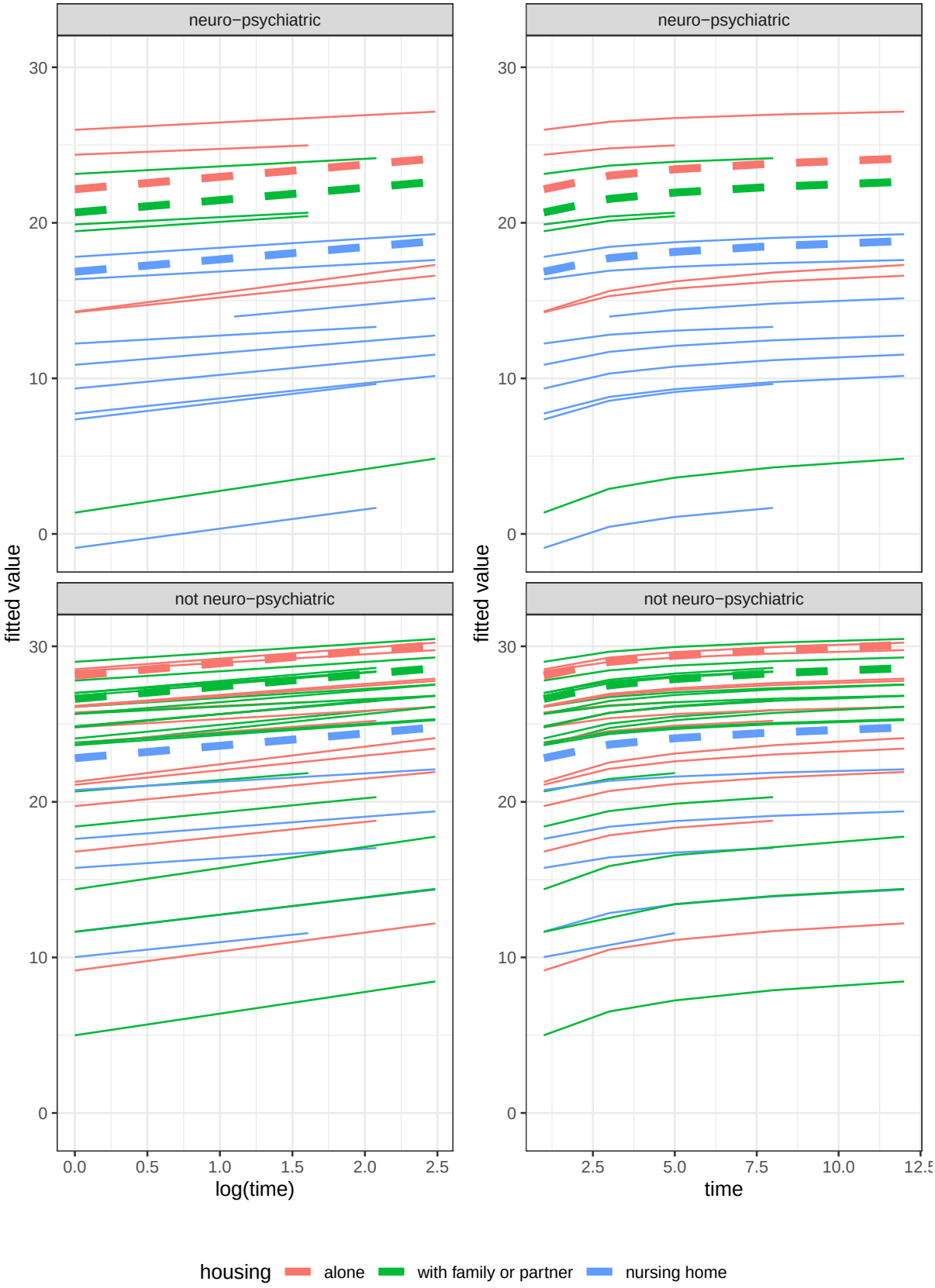

Table 10: Fixed effects for a model with random intercept and slope, including age and housing, dichotomized dependent variable (median split)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.713	4.638	-0.154	0.878
logtime	1.324	15.268	0.087	0.931
NEUROneuro-psychiatric	-2.119	2.931	-0.723	0.470
age65	-0.154	0.156	-0.988	0.323
housinginvalone	4.151	4.137	1.003	0.316
housinginvwith family or partner	3.473	4.121	0.843	0.399
logtime:NEUROneuro-psychiatric	-7.275	9.783	-0.744	0.457
logtime:age65	-0.318	0.504	-0.630	0.529
logtime:housinginvalone	4.281	13.874	0.309	0.758
logtime:housinginvwith family or partner	3.838	13.712	0.280	0.780

Table 11: Fixed effects for a model with random intercept and slope, including age and housing, dichotomized dependent variable (split at 25).

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.988	4.207	0.235	0.814
logtime	0.546	1.482	0.368	0.713
NEUROneuro-psychiatric	-5.611	3.102	-1.809	0.070
age65	-0.401	0.231	-1.736	0.082
housinginvalone	5.417	3.753	1.443	0.149
housinginvwith family or partner	4.267	3.651	1.169	0.242
logtime:NEUROneuro-psychiatric	-0.921	0.987	-0.932	0.351
logtime:age65	-0.039	0.053	-0.744	0.457
logtime:housinginvalone	0.600	1.300	0.462	0.644
logtime:housinginvwith family or partner	1.005	1.276	0.788	0.431

3.2.4 Median split

Lastly, a model is tested where mmse is dichotomized using a median split; all mmse scores equal to or below the median (23) are cetagorized as "Low" and all mmse scores over 23 are categorized as "High". It should be noted that the literature warns against this technique as it is rarely justified from either a conceptual or statistical perspective [@maccallum2002practice].

In this model, nothing is significant. It seems we've indeed lost a lot of information by dichotomizing the dependent variable.

3.2.5 Dichotomized at 25

The maximum score for the MMSE is 30. A score of 25 or higher is classed as normal. If the score is below 24, the result is usually considered to be abnormal, indicating possible cognitive impairment.

In this model, nothing is significant. It seems we've indeed lost a lot of information by dichotomizing the dependent variable.

4 Conclusion

5 Bibliography