Chapter 3
Paired Continuous Data with Predictors (Generalized Linear Models, 50 Patients)

1 General Purpose

Paired t-tests and Wilcoxon signed rank tests (Chap. 2) require, just like multivariate data, two outcome variables, like the effects of two parallel treatments. However, they can not assess the effect of additional predictors like patient characteristics on the outcomes, because they have no separate predictor variables for that purpose. Generalized Linear Models can simultaneously assess the difference between two outcomes, and the overall effect of additional predictors on the outcome data.

2 Schematic Overview of Type of Data File

Outcome 1	outcome 2	predictor…..
.	.	.
.	.	.
.	.	.
.	.	.
.	.	.
.	.	.
.	.	.
.	.	.
.	.	.

Unlike pairedt -tests (Chap. 2) generalized linear models can simultaneously test the difference between two paired continuous outcomes and the paired outcomes for additional predictor effects. For the purpose a normal distribution and a linear link function is adequate.

3   Primary Scientific Question

Can crossover studies of different treatments be adjusted for patients' age and other patient characteristics. Can this methodology also be used as training samples to predict hours of sleep in groups and individuals. The data file has to be recoded for the purpose.

4   Data Example

The underneath study assesses whether a sleeping pill is more efficaceous than a placebo. The hours of sleep are the outcome values.

Outcome 1	Outcome 2	predictor
6,10	5,20	79,00
7,00	7,90	55,00
8,20	3,90	78,00
7,60	4,70	53,00
6,50	5,30	85,00
8,40	5,40	85,00
6,90	4,20	77,00
8,70	6,10	66,00
7,40	3,80	34,00
5,80	6,30	67,00

outcome = hours of sleep
predictor = years of age

5   Recoding the Data File

After recoding the data file is adequate for a generalized linear analysis.

Outcome		predictor	pat. no.	treatment
Outcome 1	6,10	79,00	1,00	1,00
outcome 2	5,20	79,00	1,00	2,00
outcome 1	7,00	55,00	2,00	1,00
outcome 2	7,90	55,00	2,00	2,00
outcome 1	8,20	78,00	3,00	1,00
outcome 2	3,90	78,00	3,00	2,00
outcome 1	7,60	53,00	4,00	1,00
outcome 2	4,70	53,00	4,00	2,00
outcome 1	6,50	85,00	5,00	1,00
outcome 2	5,30	85,00	5,00	2,00
outcome 1	8,40	85,00	6,00	1,00
outcome 2	5,40	85,00	6,00	2,00

the outcomes 1 and 2 are paired observations in one patient
predictor = patient age
treatment = treatment modality (1 or 2)

Note that in the lower one of the above two tables each patient has two, instead of the usual one, row.

6   Analysis: Generalized Linear Models

The module Generalized Linear Modeling includes pretty sophisticated analysis of variance methods with so called link functions. The data file is in extras.springer. com, and is entitled "chapter4generalizedlmpairedcontinuous". SPSS is used for analysis, with the help of an XML (Extended Markup Language) file for future predictive testing from this model. Start by opening the data file in SPSS.

For analysis the module Generalized Linear Models is required. It consists of two submodules: Generalized Linear Models and Generalized Estimation Models. The first submodule covers many statistical models like gamma regression (Chap. 30), Tweedie regression (Chap. 31), Poisson regression (Chaps. 21 and 47), and the analysis of paired outcomes with predictors (current Chap.). The second is for analyzing binary outcomes (Chap. 42). We will use the linear model with age and treatment and as predictors. We will start with allowing SPSS to prepare an export file for making predictions from novel data.

Command:
Click Transform. . ..click Random Number Generators. . ..click Set Starting Point . . ..click Fixed Value (2000000). . ..click OK. . ..click Analyze. . ..Generalized Linear Models. . ..again click Generalized Linear models. . ..click Type of Model. . ..click Linear. . ..click Response. . ..Dependent Variable: enter Outcome. . ..Scale Weight Variable: enter patientid. . ..click Predictors. . ..Factors: enter treatment. . .. Covariates: enter age. . ..click Model: Model: enter treatment and age. . ..click Save: mark Predicted value of linear predictor. . .. click Export. . ..click Browse. . ..File name: enter "exportpairedcontinuous". . .. click Save. . ..click Continue. . ..click OK.

Parameter estimates

| Parameter | B | Std. Error | 95% Wald confidence interval | | Hypothesis test | | |
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	6,178	,5171	5,165	7,191	142,763	1	,000
[treatment=1,00]	2,003	,2089	1,593	2,412	91,895	1	,000
[treatment=2,00]	0[a]						
age	−,014	,0075	−,029	,001	3,418	1	,064
(Scale)	27,825[b]	3,9351	21,089	36,713			

Dependent variable: outcome
Model: (Intercept), treatment, age
[a]Set to zero because this parameter is redundant.
[b]Maximum likelihood estimate.

The output sheets show that both treatment and age are significant predictors at $p < 0.10$. Returning to the data file we will observe that SPSS has computed predicted values of hours of sleep, and has given them in a novel variable entitled XBPredicted (predicted values of linear predictor). The saved XML file entitled "exportpairedcontinuous" will now be used to compute the predicted hours of sleep in five novel patients with the following characteristics. For convenience the XML file is given in extras.springer.com.

Age	pat no.	Treatment (1 = sleeping pill, 2 = placebo)
79,00	1,00	1,00
55,00	2,00	1,00
78,00	3,00	1,00
53,00	4,00	2,00
85,00	5,00	1,00

Enter the above data in a new SPSS data file.

Command:
Utilities....click Scoring Wizard....click Browse....click Select....Folder: enter the exportpairedcontinuous.xml file....click Select....in Scoring Wizard click Next ....click Use value substitution....click Next....click Finish.

The above data file now gives individually predicted hours of sleep as computed by the linear model with the help of the XML file. Enter the above data in a new SPSS data file.

Age	pat no.	Treatment	Predicted values of hours of sleep in individual patient
79,00	1,00	1,00	7,09
55,00	2,00	1,00	7,42
78,00	3,00	1,00	7,10
53,00	4,00	2,00	5,44
85,00	5,00	1,00	7,00

7 Conclusion

The module Generalized Linear Models can be readily trained to predict from paired observations hours of sleep in future groups, and, with the help of an XML file, in individual future patients. The module can simultaneously adjust the data for patient characteristics other than their treatment modality, e.g., their age.

We should add, that, alternatively, repeated-measures analysis of variance (ANOVA) with age as between-subject variable can be used for the analysis of data files with paired outcomes and predictor variables. Just like in the current model statistically significant treatment and age effects will be observed. In addition, interaction between treatment and age will be assessed. The repeated-measures ANOVA does, however, not allow for predictive modeling with the help of XML files. Repeated-measures ANOVA is in the module General Linear Models, and will be reviewed in the Chaps. 9 and 10.

8 Note

Also *binary* paired outcome data with additional predictors can be analyzed with Generalized Linear Models. However, the submodule Generalized Estimating Equations should be applied for the purpose (see Chap. 42).