Practical "Summary"
P10

1. Data *HeadInjury.csv* contains information about persons with head injury who are admitted and operated in a hospital. For each patient the distance to the hospital and the time between injury and surgery are measured, but also verbal responses on the admission to the hospital are recorded.
   What does predict the chance of recovering from a head injury?

   Suggested answer:
   Age, anaesthesia and verbal responses are significant predictors of recovering from a head injury. The chance of recovering is smaller for older persons ($OR = 0.963$, $p < 0.001$), when general anaesthesia is not given during an operation ($OR = 4.569$, $p = 0.001$), and when patients do not respond verbally at all (comparing to oriented patients those who do not respond have 81% smaller odds of recovering; $OR = 0.192$, $p = 0.003$). The distance to the hospital, the time between injury and surgery, patients' sex and alcohol consumption do not significantly influence the recovering from a head injury.

Categorical Variables Codings

			Parameter coding	
		Frequency	(1)	(2)
verbal response on admission to hospital	oriented	66	.000	.000
	confused	51	1.000	.000
	none	29	.000	1.000
alcohol	present	133	.000	
	absent	13	1.000	
general anaesthesia during operation	no	36	.000	
	yes	110	1.000	
sex	male	109	.000	
	female	37	1.000	

   Results of the backward stepwise selection method that uses the likelihood ratio statistics to compare nested models and the probability of entering into the model of 0.10 as the criterion:

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	age	-.038	.011	12.031	1	.001	.963
	sex(1)	-.213	.489	.190	1	.663	.808
	interval	-.069	.175	.155	1	.694	.933
	alcohol(1)	.222	.729	.093	1	.761	1.249
	anaesth(1)	1.704	.490	12.096	1	.001	5.494
	distance	-.014	.013	1.064	1	.302	.986
	verbal_ad			9.979	2	.007	
	verbal_ad(1)	-.712	.482	2.177	1	.140	.491
	verbal_ad(2)	-1.839	.582	9.967	1	.002	.159
	Constant	1.444	.678	4.540	1	.033	4.238
Step 2[a]	age	-.037	.011	12.196	1	.000	.963
	sex(1)	-.227	.487	.219	1	.640	.797
	interval	-.066	.175	.141	1	.707	.936
	anaesth(1)	1.713	.489	12.257	1	.000	5.544
	distance	-.014	.013	1.139	1	.286	.986
	verbal_ad			10.082	2	.006	
	verbal_ad(1)	-.695	.479	2.105	1	.147	.499
	verbal_ad(2)	-1.801	.568	10.066	1	.002	.165
	Constant	1.425	.673	4.477	1	.034	4.158
Step 3[a]	age	-.037	.011	12.108	1	.001	.963
	sex(1)	-.225	.486	.215	1	.643	.798
	anaesth(1)	1.692	.484	12.192	1	.000	5.428
	distance	-.016	.011	2.099	1	.147	.984
	verbal_ad			9.978	2	.007	
	verbal_ad(1)	-.675	.475	2.022	1	.155	.509
	verbal_ad(2)	-1.773	.562	9.961	1	.002	.170
	Constant	1.274	.539	5.590	1	.018	3.577
Step 4[a]	age	-.037	.011	12.088	1	.001	.964
	anaesth(1)	1.671	.481	12.070	1	.001	5.315
	distance	-.017	.011	2.290	1	.130	.983
	verbal_ad			9.976	2	.007	
	verbal_ad(1)	-.681	.474	2.063	1	.151	.506
	verbal_ad(2)	-1.771	.561	9.952	1	.002	.170
	Constant	1.240	.534	5.398	1	.020	3.454
Step 5[a]	age	-.037	.011	12.300	1	.000	.963
	anaesth(1)	1.519	.465	10.687	1	.001	4.569
	verbal_ad			9.260	2	.010	
	verbal_ad(1)	-.784	.465	2.840	1	.092	.457
	verbal_ad(2)	-1.649	.546	9.112	1	.003	.192
	Constant	1.074	.516	4.337	1	.037	2.928

a. Variable(s) entered on step 1: age, sex, interval, alcohol, anaesth, distance, verbal_ad.

2. Dataset *QoL.sav* includes summary of a study on the effect of cosmetic surgery on quality of life (the smaller score of variables $PostQoL$, $BaseQoL$ the better quality of life).
   Does the quality of life change after the surgery? Does the change differ when patients undergo surgery only and when they have extra meetings with a psychologist?

   Suggested answer:
   Overall, the quality of life is significantly improved after cosmetic surgery and the aver-

age improvement is 3.95 unit ($t(275) = -8.524$, $p < 0.001$). What is more, this improvement differs when patients undergo surgery only and when they have extra meetings with a psychologist ($t(274) = 2.422$, $p = 0.016$). The latter group of patients score their quality of life about 5.12 units better than before surgery, while patients who undergo surgery only score they quality of life about 2.89 units better than before surgery. However, the two groups of patients differ with respect to age ($t(271.162) = -4.641$, $p < 0.001$), gender ($\chi^2(1) = 24.941$, $p < 0.001$) and reasons of surgery ($\chi^2(1) = 4.600$, $p = 0.032$). Patients who have surgery only are on average 36 years of age, 71% of them are female and 41% undergo surgery due to physical reasons, while patients who have extra meetings with a psychologist are on average 42 years of age, 41% of them are female and 29% undergo surgery due to physical reasons. Adjusting for these three patient characteristics shows that there is a significant difference in quality of life improvement of 2.3 units between the two types of treatments ($\beta = -2.339$, $t(274) = -2.429$, $p = 0.016$).

Overall improvement of the quality of life:

One-Sample Test

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
diff_QoL	-8.524	275	.000	-3.94819	-4.8600	-3.0364

Comparison of the improvement of the quality of life between the two types of treatment:

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
diff_QoL	Equal variances assumed	1.143	.286	2.422	274	.016	2.22729	.91945	.41720	4.03737
	Equal variances not assumed			2.417	268.031	.016	2.22729	.92167	.41265	4.04193

Report

diff_QoL

Cosmetic Surgery	Mean	N	Std. Deviation
cosmetic surgery	-2.8910	145	7.45358
cosmetic surgery + meeting with psychologist	-5.1183	131	7.81603
Total	-3.9482	276	7.69491

Comparison of average age between the two types of treatment:

Group Statistics

	Cosmetic Surgery	N	Mean	Std. Deviation	Std. Error Mean
Age	cosmetic surgery	145	36.35	11.684	.970
	cosmetic surgery + meetings with psychologist	131	42.28	9.516	.831

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Age	Equal variances assumed	8.020	.005	-4.594	274	.000	-5.931	1.291	-8.472	-3.389
	Equal variances not assumed			-4.641	271.162	.000	-5.931	1.278	-8.446	-3.415

Comparison of proportions of males/females between the two types of treatment:

Cosmetic Surgery * Gender Crosstabulation

Count

		Gender		
		Female	Male	Total
Cosmetic Surgery	cosmetic surgery	103	42	145
	cosmetic surgery + meetings with psychologist	54	77	131
Total		157	119	276

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	24.941[a]	1	.000		
Continuity Correction[b]	23.740	1	.000		
Likelihood Ratio	25.289	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	24.851	1	.000		
N of Valid Cases	276				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 56.48.

b. Computed only for a 2x2 table

Comparison of proportions of physical reason/change appearance reason between the two types of treatment:

Cosmetic Surgery * Reason for Surgery Crosstabulation

Count

		Reason for Surgery		
		Physical reason	Change Appearance	Total
Cosmetic Surgery	cosmetic surgery	60	85	145
	cosmetic surgery + meetings with psychologist	38	93	131
Total		98	178	276

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	4.600[a]	1	.032		
Continuity Correction[b]	4.076	1	.044		
Likelihood Ratio	4.630	1	.031		
Fisher's Exact Test				.033	.022
Linear-by-Linear Association	4.583	1	.032		
N of Valid Cases	276				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 46.51.

b. Computed only for a 2x2 table

Results of the multiple linear regression:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.232[a]	.054	.040	7.54003

a. Predictors: (Constant), Reason for Surgery, Age, Cosmetic Surgery, Gender

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	876.283	4	219.071	3.853	.005[b]
	Residual	15406.906	271	56.852		
	Total	16283.189	275			

a. Dependent Variable: diff_QoL

b. Predictors: (Constant), Reason for Surgery, Age, Cosmetic Surgery, Gender

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-8.469	2.217		-3.820	.000
	Cosmetic Surgery	-2.339	.963	-.152	-2.429	.016
	Age	.160	.055	.231	2.899	.004
	Gender	-3.320	1.341	-.214	-2.476	.014
	Reason for Surgery	1.216	1.144	.076	1.062	.289

a. Dependent Variable: diff_QoL

3. Dataset *LevamisoleColonCancer.sav* shows information of a randomized trial for adjuvant therapy of resected colon carcinoma. A national intergroup trial was conducted in the 1980s to study drugs levamisole and fluorouracil. Patients were randomly assigned to observation (control group), levamisole alone, or levamisole combined with fluorouracil and the time to cancer recurrence was considered as the most important outcome measure.
   What can be concluded from this dataset? Did randomization produce balanced experimental groups, i.e. did patients from the different groups have similar baseline char-

acteristics? Which of the adjuvant therapies was the most effective to prolong cancer progression?

Suggested answer:
Patients from the three experimental groups had similar distribution of sex ($\chi^2(2) = 6.027$, $p = 0.049$), bowel obstruction ($\chi^2(2) = 1.501$, $p = 0.472$), colon perforation ($\chi^2(2) = 1.343$, $p = 0.511$), age ($F(2, 1169) = 0.053$, $p = 0.948$) and number of positive lymph nodes ($F(2, 1138) = 0.335$, $p = 0.715$). The 2-degree of freedom logrank test of homogeneity of survival curves across the three treatment groups rejects the null hypothesis at $p < 0.001$. The KM plot shows that combination of fluorouracil and levamisole was the most effective in prolonging cancer progression. Less than 50% of patients who received both therapies experienced a progression during the follow-up time of about 13 years (median follow-up was not reached), while in the other two treatment groups 50% of patients experienced a progression after 3 and 3.5 years in observation and levamisole groups, respectively. In total 160, 227 and 213 patients out of 385, 391 and 396 experienced a progression in fluorouracil and levamisole, observation and levamisole groups. Moreover, based on the Cox proportional hazards models one can say that combination of fluorouracil and levamisole decreased the chance of experiencing a cancer progression for about 40% when compared to observation group (unadjusted $HR = 0.61$, $p < 0.001$; adjusted $HR = 0.61$, $p < 0.001$) and the effect of levamisole alone is similar to the effect of no treatment at all (unadjusted $HR = 0.92$, $p = 0.396$; adjusted $HR = 0.89$, $p = 0.252$).

Comparison of proportions of males/females between the three types of treatment:

treatment * sex Crosstabulation

Count

		sex		
		female	male	Total
treatment	observation	192	204	396
	levamisole	176	215	391
	fluorouracil + levamisole	207	178	385
Total		575	597	1172

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	6.027[a]	2	.049
Likelihood Ratio	6.034	2	.049
Linear-by-Linear Association	2.130	1	.144
N of Valid Cases	1172		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 188.89.

Comparison of proportions of occurrence/no occurrence of obstruction between the three types of treatment:

treatment * obstruction Crosstabulation

Count

		obstruction		
		no	yes	Total
treatment	observation	305	91	396
	levamisole	305	86	391
	fluorouracil + levamisole	310	75	385
Total		920	252	1172

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1.501[a]	2	.472
Likelihood Ratio	1.515	2	.469
Linear-by-Linear Association	1.409	1	.235
N of Valid Cases	1172		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 82.78.

Comparison of proportions of occurrence/no occurrence of perforation between the three types of treatment:

treatment * perforation Crosstabulation

Count

		perforation		
		no	yes	Total
treatment	observation	383	13	396
	levamisole	372	16	388
	fluorouracil + levamisole	371	10	381
Total		1126	39	1165

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1.343[a]	2	.511
Likelihood Ratio	1.343	2	.511
Linear-by-Linear Association	.248	1	.618
N of Valid Cases	1165		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 12.75.

Comparison of average age between the three types of treatment:

ANOVA

age

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	15.409	2	7.704	.053	.948
Within Groups	169367.250	1169	144.882		
Total	169382.659	1171			

Comparison of average number of positive lymph nodes between the three types of treatment:

ANOVA

Number of positive nodes

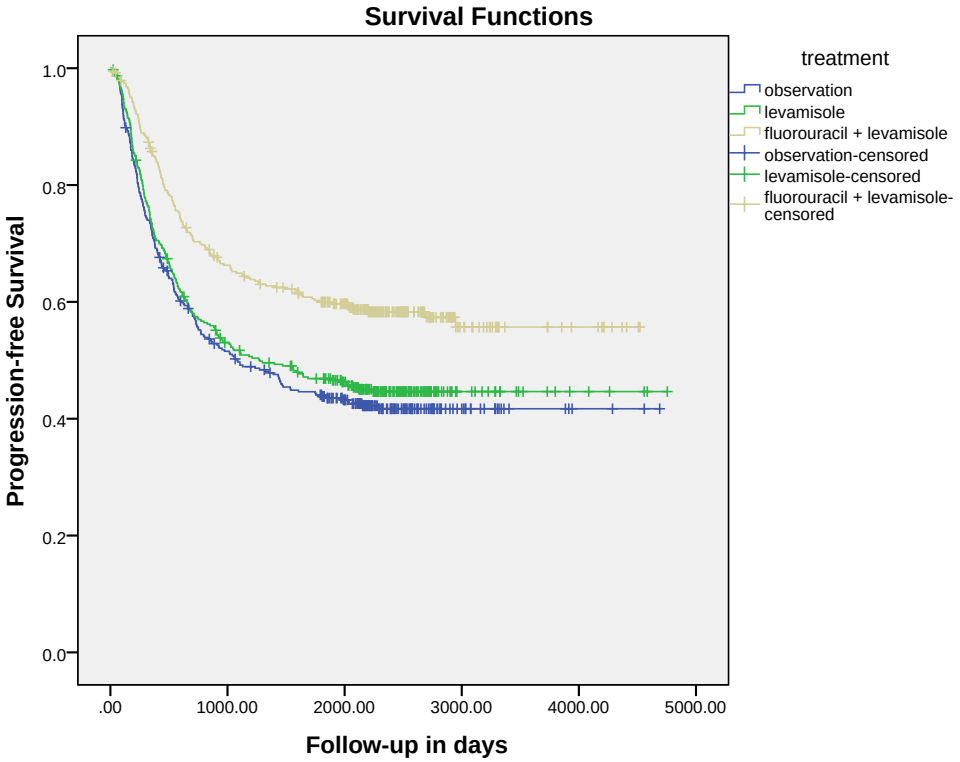
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	8.478	2	4.239	.335	.715
Within Groups	14392.724	1138	12.647		
Total	14401.202	1140			

Comparison of the progression-free survival curves for the three types of treatment:

Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	25.407	2	.000

Test of equality of survival distributions for the different levels of treatment.



Descriptives of the progression-free survival times:

Means and Medians for Survival Time

	Mean[a]				Median			
			95% Confidence Interval				95% Confidence Interval	
treatment	Estimate	Std. Error	Lower Bound	Upper Bound	Estimate	Std. Error	Lower Bound	Upper Bound
observation	2275.521	106.311	2067.153	2483.890	1081.000	193.489	701.762	1460.238
levamisole	2425.460	108.875	2212.065	2638.854	1275.000	325.241	637.527	1912.473
fluorouracil + levamisole	2856.027	103.586	2652.997	3059.056	.	.	.	.
Overall	2566.688	63.695	2441.845	2691.531	1876.000	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

treatment * progression_status Crosstabulation

Count

		progression_status		
		no	yes	Total
treatment	observation	169	227	396
	levamisole	178	213	391
	fluorouracil + levamisole	225	160	385
Total		572	600	1172

Categorical Variable Codings[a]

		Frequency	(1)	(2)
treatment[b]	1=observation	393	0	0
	2=levamisole	389	1	0
	3=fluorouracil + levamisole	382	0	1

a. Category variable: treatment

b. Indicator Parameter Coding

Results of the univariate Cox model:

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)
treatment			24.946	2	.000	
treatment(1)	-.081	.096	.720	1	.396	.922
treatment(2)	-.498	.104	22.867	1	.000	.608

Results of the multiple Cox model:

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)
treatment			22.402	2	.000	
treatment(1)	-.113	.098	1.311	1	.252	.894
treatment(2)	-.493	.107	21.401	1	.000	.611
sex	-.135	.085	2.546	1	.111	.874
age	-.004	.004	1.404	1	.236	.996
obstruction	.260	.099	6.924	1	.009	1.297
perforation	.445	.204	4.783	1	.029	1.561
pos_nodes	.090	.008	115.332	1	.000	1.094