Practical "Introduction to SPSS"
P0

1. Dataset *trial_rt.sav* contains general information from a trial with prostate cancer patients, while dataset *trial_acuretox.sav* contains toxicity information of these patients. Open both datasets and sort the records in both datasets on the patient study number, from the lowest to the highest numbers (*Data* → *Sort Cases*). Merge the file *trial_rt.sav* with the file *trial_acuretox.sav* using the patient study number as the key variable (*Data* → *Merge Files* → *Add Variables*).

2. Save the new combined dataset under a different name in SPSS format (.sav) and Excel format (.xls) (*File* → *Save As*).

3. For patients with Grade 3 rectal toxicity, report how many of them were randomized to "Conventional 68 Gy" and "Experimental 78 Gy" treatment arms (*Data* → *Select Cases* and choose the *If condition is satisfied* option to specify "maxarect=3"; *Analyze* → *Descriptive Statistics* → *Frequencies*). Save the list of commands for this exercise and all the following ones in a syntax file.

4. Find distributions of grade acute bladder toxicity for smokers and non-smokers (*Data* → *Split File* and choose the *Organize output by groups* option; *Analyze* → *Descriptive Statistics* → *Frequencies*).

5. For all patients explore the variable "psa" which contains information about the prostate specific antigen level (*Analyze* → *Descriptive Statistics* → *Explore*). What are the average values of "psa" in the different radiotherapy treatment groups (*Analyze* → *Descriptive Statistics* → *Explore* and in the factor list include the "rtgroup" variable)?

6. Plot a histogram of "psa" for all patients and separately for patients in the "Conventional 68 Gy" and "Experimental 78 Gy" treatment arms (*Graphs* → *Legacy Dialogs* → *Histogram* and in the *Rows* box include the "arm" variable).

7. Recode the variable "psa" into categorical variable "psa_level" with low level ($\leq$ 10), intermediate level (10-20), high level ($\geq$ 20) (*Transform* → *Recode into Different Variables* and choose the *Old and New Values* option or *Transform* → *Compute* and then *If* option).

8. Compare percentages of patients in the "Conventional 68 Gy" and "Experimental 78 Gy" treatment arms within each level of "psa" (*Analyze* → *Crosstabs* and choose the *Cells* option to specify row/column percentages).

9. Compute a new variable which indicates whether or not the received dose ("dose" variable) was according to the planned dose ("arm" variable). Create a second new variable that indicates the dose difference between planned and received dose.

10. Save the dataset under a different name and with only the "studnr" variable and the new variables you have created (*File* → *Save as* and then used the *Variables* option).