

## Practical Exercise: Analysis of prognostic effects (2 – 3 hours for MSc students)

Hester Lingsma, Ewout Steyerberg, Dept of Public Health, Nov 2008

In this practical exercise we will study the effect of prognostic factors in traumatic brain injury. We will assess individual prognostic strength in univariate analyses, and in multivariable analyses, where we aim to adjust for effects of other predictors that may act as confounders.

The data are in SPSS format: "TBI.sav". A description of the dataset is in the attachment.

### 1) Cause of injury

- a) Give the frequencies of cause of injury. What is the most common cause?
- b) Give the frequencies of the outcome (d.gos). What is the most common outcome? Check the categorization in favorable vs unfavorable (d.unfav variable).
- c) Assess the univariable effect of cause of injury on 'poor outcome' (d.unfav) with a logistic regression model. By default SPSS will use the last category ('Other') as the reference category. Which cause gives the highest chance of poor outcome, and which the lowest chance?
- d) Is this what you would expect? Can you think of a confounding effect?  
Hint: What is the mean age for each cause of injury?
- e) Now fit a multivariable model to adjust the effect of cause of injury for age. How does the effect change?

### 2) Age

- a) Give some descriptives of age (mean, median, interquartile range, ...).
- b) Assess the univariable effect of age on outcome (d.unfav) with a logistic regression model. What is the effect of age on the chance of poor outcome? What would be a good way to express the effect?  
Hint: think of recoding age in decades, or a comparison of the risks at the 75<sup>th</sup> to the 25<sup>th</sup> percentiles. How would you do that?
- c) Do you expect the effect of age to be linear? Test this assumption by adding the term age squared to the regression model, what is your conclusion?  
Hint: look at p-value of the squared term.
- d) In some studies, continuous variables are dichotomized into above or below a certain threshold. For the effect of age in TBI, 40 is often used as a cut-off value. Create a new variable for age above or below 40 and assess whether this dichotomization describes the effect better than the continuous term.  
Hint: look at 'model chi-square' of the models.
- e) In the CRASH study, it was found that the effect of age only existed above 30 or 40 years old (see Background). Create new age variables to assess whether these codings of age result in a better fit of the model.  
Hint: look at the 'model chi-square' of the models.
- f) Now fit a multivariable model to adjust the effect of age for motor score and pupillary reactivity. How does the effect change?  
Note that there are missing values in the variable 'd.pupil', which have been filled in with a statistical imputation procedure in 'pupil.i'.  
Perform the analyses with both variants of 'pupils'.
- g) Can we interpret the change in age coefficient from univariate analysis to adjusted analysis if numbers differ?  
Therefore: also perform univariate analyses of age in the patients without missing values for 'd.pupil' to compare the change in age coefficient. Hint: use Select Cases in the Data menu.

**3) Systolic blood pressure**

- a) Give some descriptives of systolic blood pressure.
- b) Assess the univariable effect of systolic blood pressure on outcome (d.unfav) with a logistic regression model. Do you expect the effect of systolic blood pressure to be linear?
- c) Test this assumption by adding the term systolic blood pressure squared to the regression model, what is your conclusion?
- d) To see this in a graph, define the model again and then tick the box Probabilities under Predicted Values under the Save button. Now make a scatterplot of the saved predicted value against systolic blood pressure. Use the Graphs menu, Interactive. How would you describe the relationship?
- e) What would be a good way to express the effect of systolic blood pressure?  
Hint: think of categorizing at the 75<sup>th</sup> and the 25<sup>th</sup> percentiles, and choosing a common reference category. Can you confirm the shape of the graph by the odds ratios?