**Exercises - Day 1 - Workshop on Dose-Response Meta-Analysis**

**Authors: Nicola Orsini and Alessio Crippa**

**Group 3: Survival outcome**

**Intended Learning Outcomes:**

* Read and describe individual patient data from multiple studies
* Read and describe aggregated data from multiple studies
* Estimate and interpret a linear trend model in each study
* Perform and interpret a meta-analysis of linear trends
* Conduct statistical inference (test, confidence) about the summary dose-response
* Present graphically the estimated summary dose-response function

**Part I: Meta-analysis of linear trend estimated on individual data**

Let’s consider 12 observational cohort studies investigating the association between alcohol consumption and cardiovascular disease. Smoking status is considered an important confounding factor to adjust for. The dataset contains stacked individual data for the 12 studies with the following variables:

* *id*: identifier of the study
* *x*: alcohol consumtpion measured in drinks/week
* *time*: follow-up time.
* *y*: event status indicator (0 = censored, 1 = had a cardiovascular event)
* *c*: binary variable for smoking status (1 = smoker vs 0 = non-smoker)
* *z*: study-level characteristic (a = general population, b = only women, c = older men (possible survivorship bias))

**Questions**

1. Describe the main features of the individual data, in particular the exposure and outcome distributions.

1. Create a table with the study-specific linear trend coefficients and standard error for the exposure-outcome association suitable for conducting a meta-analysis.

1. Present the study-specific estimates of the linear trends using a forest plot and/or the estimated summary dose-response function. Write the findings for the result section of an hypothetical review paper.

**Part II: Meta-analysis of linear trend estimated on aggregated data**

Let’s consider aggregated dose-response data derived from the individual patient data of the 12 studies analyzed in part I of the lab. The dataset contains the following variables:

* *id*: identifier of the study
* *xcat*: categories of alcohol consumtpion (drinks/day)
* *dose*: assigned alcohol consumtpion (drinks/day) value representative of the exposure category
* *cases*: number of events (fractures) for each exposure category
* *n*: total number of subjects for each exposure category
* *e\_b*: estimated log(HR)
* *e\_se*: estimated standard error of log(HR)
* *e\_lb*: lower 95% confidence interval bound of log(HR)
* *e\_ub*: upper 95% confidence interval bound of log(HR)
* *type*: type of outcome type (ir or 2 = cumulative incidence)
* *z*: study-level characteristic (a = general population, b = only women, c = older men (possible survivorship bias))

**Questions**

1. Describe the main features of the aggregated data, in particular th exposure and outcome distribution.

1. Create a table with study-specific estimated regression coefficients and estimated standard error suitable for conducting a meta-analysis.

1. Present the study-specific estimates of the linear trends using a forest plot and/or the estimated summary dose-response function. Write the findings for the result section of an hypothetical review paper.
2. Comment main differences between the main results obtained from the individual and the aggregated data.