

# Assignment II for

## EP16: Missing Values in Clinical Research

### Multiple Imputation

*13 – 17 May, 2019*

## Data

The **MIdata1** data comprise information of 488 participants of an observational study on the effects of testosterone on heart rate.

The dataset contains the following variables:

variable	explanation
ID	subject identifier
Gender	participant's gender
Income	Household income
HR	heart rate
sleep	hours of sleep per night
PA	number of days per week on which physical activity is performed
alc	number of alcoholic drinks consumed per week
smoke	smoking status
PC	number of hours spent using a computer per day
TV	number of hours spent watching television per day
height	participant's height in cm
weight	participant's weight in kg
BMI	participant's BMI
Age	participant's Age
Testosterone	participant's testosterone level in ng/dL
PhysActive	Does the participant perform physical activity regularly?

## Analysis model of interest

The analysis model of interest is a linear regression with outcome HR and covariates **Testosterone**, **Gender**, **Age**, **smoke**, **sleep**, **PA**, **PC**, **BMI** and **alc**.

We assume that there are interaction effects between **Testosterone** and **Gender** as well as between **Age** and **smoke**.