# Analysis of overall mortality in biobank data: challenges of modeling and implications for the estimation of "biological age"

Krista Fischer[1,2], Merli Mändul[1,2], Anastassia Kolde[1,2], Nele Taba[2]

[1] Institute of Mathematics and Statistics, University of Tartu, Tartu, J. Liivi 2, 50409, Estonia;

[2] Estonian Genome Center, Institute of Genomics, University of Tartu, Tartu, Riia 23b, 51010, Estonia;

As several population-based biobanks have already reasonably long follow-up time for the analysis of all-cause mortality, we can fit survival models that combine known predictors with different *omics*-based markers and prognostic scores. We will demonstrate using the Estonian Biobank data, that a polygenic risk score for mortality as well as a risk score based on blood biomarkers from Nuclear Magnetic Resonance Spectroscopy have both extremely strong effects next to the well-known demographic and lifestyle-related predictors.

Next, we will discuss alternative ways of interpreting the results. We point out that traditional overall Kaplan-Meier curves do not account for current age of the individual and can therefore be misleading. A popular alternative way to illustrate biomarker effects on aging is to estimate "biological age" of the individual, where a biomarker profile may indicate that the individual is "biologically older" or "younger" than his/her actual age. We point out that traditional ways of estimating biological age from regressing age on biomarker variables may lead to estimates that are correlated with age, but not with the actual mortality hazard. Instead, we propose an alternative definition of biological age that is directly based on survival analysis and may help to interpret individual predictions based on survival models.