**Supplementary material: Details on Data and Measurements**

The data used in this paper were collected from non-Hispanic African American and Caucasian men and women enrolled in the Johnston County Osteoarthritis Project (JoCo OA). JoCo OA is an ongoing, longitudinal population-based prospective study with clearly defined and repeatedly measured radiographic OA, comorbidities, various biomarkers, socio-demographic and physiological variables. The detailed information on data collection, measurement procedures and OA assessment are described elsewhere([1](#_ENREF_1), [2](#_ENREF_2)). The JoCo OA is an example of an epidemiological study that collects both repeated measurements and time-to-event outcomes. In particular, longitudinal measures of different biomarkers are collected together with the repeated assessment of OA status. The simultaneous analysis of these data might help in understanding how these conditions occur through identifying risk factors that can be modified to reduce the burden of disease.

Data for this paper were collected at the baseline and three follow-ups, roughly every 5-6 years. To allow for changes in OA status, we considered the participants who had at least two visits with observed height, weight and knee radiographs for the same knee (either right, left or both). BMI was calculated as mass in kilograms divided by height in meters squared.Radiographic evidence of the presence and severity of knee OA was obtained from weight-bearing anteroposterior (AP) extended radiographs at baseline and fixed-flexion posteroanterior (PA) radiographs at follow-ups and graded using the K-L scale ([3](#_ENREF_3)) by a single musculoskeletal radiologist with high reliability. It was previously shown, that PA and AP views are comparable for the purpose of categorizing joints by K-L grade ([4](#_ENREF_4)). The participants were followed until the first occurrence of K-L worsening, loss to follow-up, or the last follow-up; for this analysis, T3. Our final sample comprised 2,286 participants with 5,325 longitudinal measurements of BMI. In addition, all longitudinal data were excluded from analysis after the K-L grade 4 was reached in both knees and/or one or two knees were replaced - whichever occurred first - as radiographic worsening was no longer possible by definition.

The mean number of longitudinal BMI measurements per individual until event/censoring was 2.3 (SD=0.6) with 1,646 (72%) participants having 2 observations, 527 (23%) having 3 observations, and 113 (5%) having all four observations. At baseline, mean age was 59.7 years (SD=3.9, range 44 to 93.6) and mean BMI was 29.6 (SD=6.2).

One thousand five hundred three participants (65.8%) were women; 701 (30.7%) self-reported African American (AA) race. The mean follow-up time was 8.2 years (range 3.6 to 22.2, SD=3.9). During that time, an increase in K-L grade in at least one knee was observed in 1,543 (67.5%) participants.

**Supplementary material: Details for Joint Modeling**

We fitted several joint models using the R package JM([5](#_ENREF_5)): basic JM, and JM with different parameterizations defined below. The first (basic) model (JM1) consists of the linear mixed-effects model for longitudinal BMI data with normally distributed errors:

*,*

and time-to-event sub-model that specifies the hazard of event as a function of the “true” longitudinal outcome with adjustment for gender and age at the baseline as follows:

.

Here the regression parameters corresponding to fixed effects,  and , represent the mean (“population”) intercept and mean slope;  and  are random individual-specific deviations for intercepts and slopes, respectively;  are independent and normally distributed individual error terms;  is the baseline hazard;  is the vector of baseline explanatory variables (sex and age at first examination) with a corresponding vector of regression coefficients (where *T* denotes mathematical vector transposition). is the “true” longitudinal outcome, which is modeled as:

.

The interpretation of coefficient  is the same as in the CoxPH model with TVCs, i.e.,  is the relative increase in the risk of event with each unit increase of BMI at the same time.

The difference between JM and CoxPH with TVC is that in the former model, the hazard depends on the “true” value of BMI estimated in the longitudinal sub-model, whereas in the latter model the observed value (potentially measured with error) is used. Longitudinal BMI values were logarithmically transformed to satisfy assumptions of normality. The shape of a longitudinal trajectory was determined based on the significance of polynomial terms and the distribution of the mean number of observations in the sample. We included both random intercepts and random slopes in the longitudinal sub-model to consider subject-specific deviations from the average intercept (i.e., mean BMI at the baseline) and slope (mean rate of change). Because the BMI change pattern and K-L worsening may be related to age, we adjusted both time-to-event and longitudinal sub-models for age at baseline. The time-to-event sub-model was also adjusted for sex. The age at the baseline was standardized to improve the model performance. We used hazard ratios (HR) as measures of association between BMI at a particular time (e.g., current level) and the risk for an event at the same time. Because we used a log transformation of BMI measurements, the coefficients for the longitudinal process can be interpreted as a relative change. To provide clinically relevant interpretation, we calculated the HR for a difference of 25% in BMI values which can quantify the risk of the event if a person had 25% higher BMI at the same time.

In the second joint model (JM2), the risk depends on the slope of the “true” trajectory at that time:

,

where is the regression coefficient for the slope and is the derivative of BMI at time/age *t* (i.e., the “slope”). We used JM2 to evaluate the association of change in BMI and onset of worsening in the K-L grade. Here we calculated the HR to compare an annual increase of 10% to an increase of 5%. In the third model (JM3), we assumed that the risk depends on both the current “true” level and the slope of BMI at the same time:

.

In JM3, the interpretation of  is the same as that of in JM1, while  is the regression coefficient for the slope.

**Supplementary material: References**

1. Jordan JM, Linder GF, Renner JB, Fryer JG. The impact of arthritis in rural populations. Arthritis Care and Research. 1995;8(4):242-50.

2. Jordan JM, Helmick CG, Renner JB, Luta G, Dragomir AD, Woodard J, et al. Prevalence of knee symptoms and radiographic and symptomatic knee osteoarthritis in African Americans and Caucasians: The Johnston County Osteoarthritis project. Journal of Rheumatology. 2007;34(1):172-80.

3. Kellgren JH, Lawrence JS. Radiological Assessment of Osteo-Arthrosis. Annals of the Rheumatic Diseases. 1957;16(4):494-502.

4. Nelson AE, Renner JB, Shi XA, Shreffler JH, Schwartz TA, Jordan JM. Cross-Sectional Comparison of Extended Anteroposterior and Posteroanterior Fixed Flexion Positioning to Assess Radiographic Osteoarthritis at the Knee: The Johnston County Osteoarthritis Project. Arthritis Care & Research. 2010;62(9):1342-5.

5. Rizopoulos D. JM: An R Package for the Joint Modelling of Longitudinal and Time-to-Event Data. Journal of Statistical Software. 2010;35(9):1-33.