Marginal modeling is a term used for an approach where the effect of explanatory factors is estimated based on considering the marginal distributions. The dependence is not the interesting aspect and is not considered in detail. Afterwards, the variability of the regression coefficient estimators is determined by a procedure that accounts for the dependence between the observations. For parallel data, there are in practice two versions of this general idea. The coordinate-wise (CW) approach considers each marginal separately and estimates the regression coefficients in each marginal. The covariance matrix of these estimates is estimated and used for combining the estimates from the coordinates by means of a weighted average. The estimated covariance matrix is further used to evaluate the variance of the combined estimate. The second version of the approach is the independence working model (IWM) approach, where the estimate is found under the (incorrect) assumption of independence between the coordinates. This yields directly the final estimate of the regression coefficients. The uncertainty of the regression coefficient estimate is evaluated by means of an estimator that accounts for the dependence between the coordinates. This is typically done by a "sandwich estimator" (see below). The independence working model approach is closely related to the so-called generalized estimating equations. For recurrent events, there is an approach similar in concept to the IWM approach.

The marginal modeling approaches are model-free regarding dependence assumptions. It can be seen as an advantage that we do not have to specify and rely on a specific model; but on the other hand, it is unknown whether the approaches are particularly good or bad for specific models.
In terms of purpose of the analysis these approaches may be useful for finding the effect of covariates, but do not make sense for assessment of dependence, goodness-of-fit, or for prediction. In fact, it is further limited, as the coordinate-wise approach cannot handle matched pairs covariates.

The CW and IWM approaches are derived in a frame of parallel data for several individuals. They are less good for recurrent events data, as such data show an order restriction $T_1 < T_2$, which means that the observations are concentrated on a subset of $(0, \infty)^2$. The marginal modeling approach does not use this fact, but is based on some asymptotic normal approximations. The approaches furthermore have difficulties handling the varying number of observations. I think, these approaches should be used only when the observations vary in a product set and when the marginal model includes the hypothesis of independence between the observations. The approaches do not appear relevant for event-related dependence and multi-state models, because they do not recognize this type of dependence. None of the methods should be used for longitudinal data on different events, where one of the events is death, first because the methods do not accept the process-dependent censoring owing to death; and second, because there is no reason to suggest that the covariate should have the same effect for all events. Furthermore, it seems relevant only to consider continuous data, ruling out data showing instantaneous dependence. The coordinate-wise approach is described in detail in Section 13.1. The independence working model approach is described in Section 13.2. The two approaches are discussed and compared in Section 13.3. A similar approach to studying recurrent events is described in Section 13.4.

A somewhat complementary approach to the marginal modeling is the concept of copulas (Section 13.5). The purpose of this approach is purely to study the dependence, removing all effects of the distribution as function of time and covariates. This is obtained by assuming that the marginal distributions are known. For standardization a uniform distribution on the unit interval is chosen. If the distribution is another continuous distribution, it can be transformed to the uniform case. Then the dependence is evaluated by specifying a family of distributions for the bivariate observations. It can, however, be discussed to which extent this approach is useful as the marginal distributions are rarely known. Typically, there will be some parameters also in the marginal distributions and then we do need a larger model for the analysis. In the terminology for purposes introduced in Section 1.14 this approach can be used for assessment of the dependence and for evaluating the goodness-of-fit of specific models. Furthermore, the copula approach is relevant for probability evaluations made for understanding the dependence pattern.

However, the marginal and copula approaches can be combined so that the effect of the covariates and the time scale is modeled by means of the marginal hazard functions and the dependence is modeled by the copula approach (Section 13.6). In this way, we get the advantages of both pro-