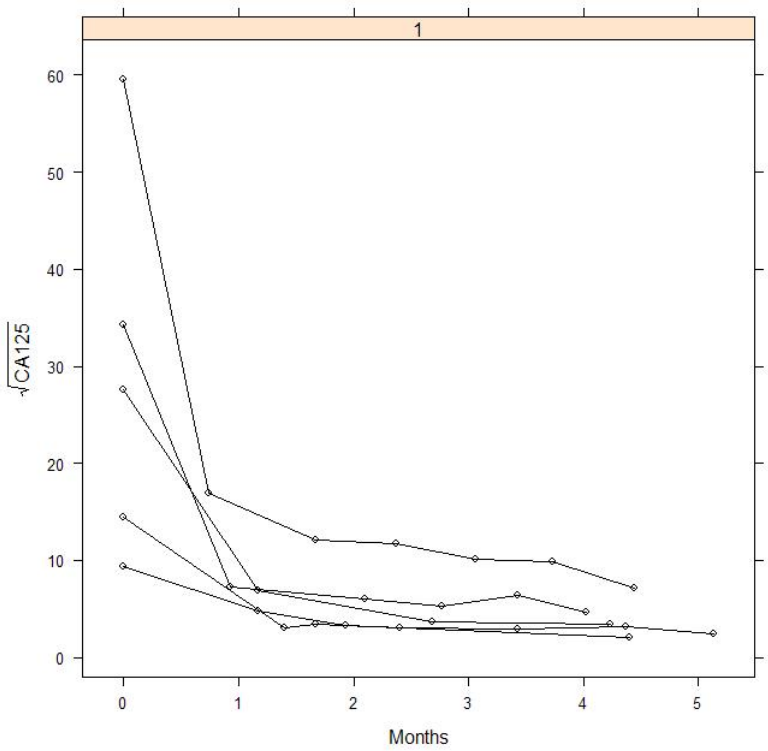
1 Joint Model

For each individual, we observe his/her survival data and covariates. Denote the ith individual's event time (e.g. survival time or progression-free survival time) and censoring time by T_i and C_i, respectively. We thus observe $X_i = min(T_i, C_i)$ and the failure indicator δ_i, which is equal to 1 if the failure is observed ($T_i < C_i$) and 0 otherwise. We assume that C_i is independent of T_i and covariates. To simplify the model, here we only consider a single covariate which is repeatedly measured such as the longitudinal CA125. Denote by $y_i(t)$ the value of the longitudinal outcome at time point t for the ith individual. In practice, we cannot observe $y_i(t)$ at all time points; instead we only observe at several time points $t_{i1}, \ldots, t_{in_i}$. In vector notation, we denote $\boldsymbol{y}_i = (y_i(t_{i1}), \ldots, y_i(t_{ij}))$ and $\boldsymbol{t}_i = (t_{i1}, \ldots, t_{in_i})$. The observed longitudinal CA125 thus consist of the measurements $y_{ij} = y_i(t_{ij}), j = 1, \ldots, n_i$, $y_i(t)$ is measured with error; hence we also denote the true (uncontaminated with error) and unobserved longitudinal outcome at time point t by $m_i(t)$. Other covariates are time-independent, and let $\boldsymbol{u}_i = (u_{i1}, \ldots, u_{ik})^T$ denote as the time-independent covariates for the ith subject.

The CA125 distribution at baseline (before surgery) was skewed; therefore, for the reminder of this paper we will work with the square root of CA125 values and refer to the transformed covariate as y_{ij}.

As shown in figure 1, there was clearly a lot of heterogeneity in the baseline CA125 values and the slopes of the CA125 profiles between individuals.


1
60
50
40
30
20
10
0
√CA125
0
1
2
3
4
5
Months

Therefore, it is reasonable to use the mixed-effects model to fit the CA125 trajectories. Moreover, the CA125 trajectories seem stable after a cut point (around 2 months, roughly after 1 or 2 rounds of chemotherapy) and before the end of the first-line chemotherapy. Therefore, it is reasonable to model the observed CA125 values as follows:

$$y_{ij} = m_i(t_{ij}) + \epsilon_{ij} = (a_0 + b_{0i}) + (a_1 + b_{1i})t_{ij} + a_2(t_{ij} - c)I_{t_{ij}-c} + \epsilon_{ij} \quad (1)$$

where c is the cut point, I_x is the indicator function which is equal to 1 if $x \geq 0$, and 0 otherwise; $m_i(t_{ij})$ is the true and unobserved CA125 values at time t_{ij} for the ith subject; ϵ_{ij} is the error with the distribution $N(0, \sigma^2)$ (ϵ_{ij_1} and ϵ_{ij_2} are independent if $j_1 \neq j_2$); a_0 and a_1 are the population intercept and slope before the cut point, respectively; a_2 is the difference of the population slopes between the curves before and after the cut point; and b_{0i} and b_{1i}, accounting for the individual's heterogenity of the intercepts and slopes before the cut point, respectively, are assumed to be a bivariate normal random variable, i.e.,

$$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{00} & \sigma_{01} \\ \sigma_{01} & \sigma_{11} \end{pmatrix}\right)$$

Denote $\mathbf{\Sigma} = \begin{pmatrix} \sigma_{00} & \sigma_{01} \\ \sigma_{01} & \sigma_{11} \end{pmatrix}$.
Note that given the cut point c, (1) is the usual linear mixed-effects model. We proposed 2 methods to estimate the cut-point, which will be introduced

in the next section. For simplicity, we consider each trajactory of the CA125 curve has the same cut point. The model can then be easily extended to allow different cut points for different trajactories.

The aim is to link the true value of the longitudinal outcome (CA125) at time t, $m_i(t)$, with the survival time T_i. We model the hazard of failure using the Cox PH model, where the hazard depends on the longitudinal covariate only through its true current value (without error).Therefore, the hazard is modeled as follows:

$$h(t|\boldsymbol{u}_i, \boldsymbol{y}_i, \boldsymbol{t}_i) = h_0(t) \exp\{\boldsymbol{u}_i\boldsymbol{\alpha} + \beta((a_0 + b_{0i}) + (a_1 + b_{1i})t + a_2(t-c)I_{t-c})\}, \quad (2)$$

where $h_0(t)$ is the baseline hazard at time t, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k)^T$ is the coefficient vector, $\boldsymbol{u}_i = (u_{1i}, \ldots, u_{ki})$ is the ith individual's time-independent covariate vector, and β is the coefficient for the longitudinal covariate. Or equavalently,

$$h(t|\boldsymbol{u}_i, m_i(t)) = h_0(t) \exp\{\boldsymbol{\alpha}\boldsymbol{u}_i + \beta m_i(t).\} \quad (3)$$

The joint model is composed of (1) and (3). The likelihood for the observed data is given by

$$\prod_{i=1}^{n} \left[\int_{-\infty}^{\infty} \left\{ \prod_{j=1}^{m_i} f(y_{ij}|c, \boldsymbol{a}, \boldsymbol{b}_i, \sigma^2) \right\} f(\boldsymbol{b}_i|\boldsymbol{\Sigma}) f(X_i, \delta_i|\boldsymbol{b}_i, h_0, \beta, \boldsymbol{\alpha}) \right], \quad (4)$$

where

$$f(y_{ij}|c,\boldsymbol{a},\boldsymbol{b}_i,\sigma^2) = (2\pi\sigma^2)^{-1/2}\exp\{-\left(y_{ij}-a_0-b_{0i}-a_1t_{ij}-b_{1i}t_{ij}-a_2(t_{ij}-c)I_{t_{ij}-c}\right)^2/2\sigma^2\} \quad (5)$$

$$f(\boldsymbol{b}_i|\boldsymbol{\Sigma}) = (2\pi|\boldsymbol{\Sigma}|)^{-1/2}\exp\{-(\boldsymbol{b}_i)^T\boldsymbol{\Sigma}^{-1}\boldsymbol{b}_i/2\}, \quad (6)$$

and

$$\begin{aligned} & f(X_i,\delta_i|c,\boldsymbol{b}_i,h_0,\beta,\boldsymbol{\alpha}) \\ = \quad & [h_0(t)\exp\{\beta((a_0+b_{0i})+(a_1+b_{1i})X_i+a_2(X_i-c)I_{X_i-c})\}]^{\delta_i} \\ + \quad & \exp[-\int_0^{X_i} h_0(u)\exp\{\beta((a_0+b_{0i})+(a_1+b_{1i})u+a_2(u-c)I_{u-c})\}du] \end{aligned}$$

In the next section, we will introduce how to estimate the cut point and other parameters.

2 Estimating parameters

2.1 Estimating the cut point

Below we provide two possible methods to estimate the cut point.

(1) empirical method: By observing all the patients' CA125 profiles or from previous experience, sometimes we have knowledge about where the change point is.

(2) maximum likelihood method: The second method is based on the likelihood (5). By maximizing this likelihood function, we can obtain an estimate

of the cut point. Note that given the cut point c, 1 is a linear mixed-effects model. Therefore, given the cut point c, it is easy to obtain the maximum likelihood estimate using the R commend lme for other parameters. Therefore, given c, it is easy to obtain the maximum likelihood value (for that c). By searching for c that gives the largest maximum likelihood value, we can obtain the maximum likelihood estimator for c.

In the 2nd version of the codes for our joint model, we provide the code computing the maximum likelihood estimate for the change point.