# Using Explanatory Variables

- Now let $X_1$ be an indicator variable, which takes the value zero if an individual is on the standard drug, and unity if an individual is on the new drug.
- If $x_{1i}$ is the value of $X_1$ for the $i^{th}$ individual in the study, i=1, 2,…, n, the hazard function for this individual can be written as

$$h_i(t) = e^{\beta * x_{1i}}\, h_0(t)$$

  - where $x_i=1$ if the $i^{th}$ individual is on the new treatment and $x_i=0$ otherwise.
- This is the proportional hazards model for the comparison of two treatment groups.
- Let $h_0(t)$ *be the hazard function for an* individual for whom the values of $X_1$ zero.
  - The function $h_0(t)$ *is called the baseline hazard function.*

# Using Exploratory Variables

- To get to the Cox Regression model we now generalize to the situation where there also other explanatory variables , such as
  - Gender
  - Age,.....
  - etc…

  These are represented by variables
- Then the hazard of death at a particular time also depends on the values $x_2, \ldots, x_p$ of p explanatory variables, $X_1, X_2, \ldots, X_p$.

  So $$h_i(t) = e^{(\beta_1 * x_{1i} + \beta_2 * x_{2i} + \ldots\ldots)}\, h_0(t)$$
- *Note:*
  - *These variables can be continuous, categorical, Binary.....*
  - *The values of these variables will be assumed to have been recorded at the time origin of the study.*

# The General form of the Proportions Hazard Model

- Since this model can be re-expressed in the format,

$$\text{Log}\,(h_i(t)/h_o(t)) = \text{Log}(\psi) = \beta_1 * x_{1i} + \beta_2 * x_{2i} + \ldots\ldots$$

- The proportional hazards model may also be regarded as a linear model for the logarithm of the hazard ratio.
- It is a generalized linear model with the link function being the log function
- There are other possible forms for $\psi$, ***but the choice $\psi(x_i)=exp(\beta' x_i)$*** leads to the most commonly used model for survival data.
  - Alternatives are the Weibull parametric model, the lognormal model, the Gompertz model
- Notice that there is no constant term in the linear component of the proportional hazards model.
  - If a constant term $\beta_0$, *say, were included, the baseline hazard function could simply* be rescaled by dividing $h_0(t)$ *by* $exp(\beta_0)$, *and the constant term would cancel out.*
  - *Also* we have made no assumptions concerning the actual form of the baseline hazard function $h_0(t)$.