

Why patients may not benefit from effective oncology treatments

Blesson Chacko, Jon Moss, Rupert Austin and Joachim Grevel

BAST Inc Ltd., Loughborough, UK

Survival analysis from the patient's perspective:

- A classical survival analysis allows to compare a “drug effect” to a “control effect”. If the drug effect is stronger, **efficacy** has been proven and the road to market approval is wide open.
- The approval process will also assess **safety**. The trade-off between efficacy and safety is carried out by clinical judgement that varies between therapeutic areas. Under life-threatening conditions (e.g. oncology) more concessions are made with regards to safety.
- Yet, a **patient only benefits** when the therapeutic effect is achieved before tolerability is exhausted. From an oncology patient's perspective “dropout” is a terrible waste of time. The same is probably true for all patients, since all are in a race against time.

Why is “waste of time” not listed among the outcomes in clinical trial protocols?

Sub-distribution hazard and cumulative incidence function (CIF)

The cause-specific hazard can be modified to take into account competing risks, and the so-called sub-distribution hazard² for event k at time t is given by:

$$\hat{h}_k(t) = \lim_{\delta t \rightarrow 0} \frac{\Pr(t \leq T < t + \delta t, K = k | T > t \text{ or } T \leq t \text{ and } K \neq k)}{\delta t}.$$

- The sub-distribution hazard $\hat{h}_k(t)$ attempts to evaluate the absolute risk for an individual patient. Thus, it can quantify patient's chance to benefit in the face of competing risks.
- The relationship between the sub-distribution hazard and the CIF for an event k , denoted here as $\hat{F}_k(t)$, is given by
$$\hat{h}_k(t) = -\frac{d}{dt} \log(1 - \hat{F}_k(t)).$$
- The key mathematical difference between the cause-specific and sub-distribution hazards is in the elements of the risk set:
 - For the cause-specific hazard, the risk set decreases each time there is a competing event, since competing events are treated like right-censoring.
 - For the sub-distribution hazard, patients that experience a competing event remain in the risk set.
- The major advantage of calculating the CIF using the sub-distribution hazard function is that the overall covariate effect on the CIF can be easily quantified.

Method: Parametric regression on CIF

- The Fine-Gray method uses a semi-parametric model (just like Cox-PH), and the proportionality condition must be satisfied. This may not always be the case.

- Over the past decade, the principles of the Fine-Gray method have been used to infer parametric CIF models³ which offers a significant advantage over the Fine-Gray method, such as:
 - The proportionality assumption can be relaxed.
 - Simulations can explore different patient populations.
 - Model performance can be assessed using visual predictive check.

- We chose a flexible three-parameter logistic model to characterize the CIF parametrically:

$$\hat{F}_k(t; p_k, b_k, c_k) = \frac{p_k [\exp\{b_k(t - c_k)\} - \exp(-b_k c_k)]}{1 + \exp\{b_k(t - c_k)\}}$$

- Along with a generalized odds-rate link function to include covariate influences:

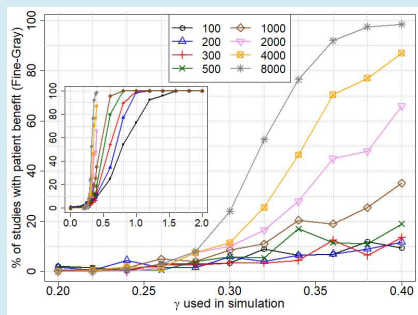
$$g(u; \alpha_k) = \log \left[\frac{(1-u)^{-\alpha_k} - 1}{\alpha_k} \right], \quad \alpha_k > 0.$$

- Note that $g(u; \alpha_k)$ is the logit function when $\alpha_k = 1$ (proportional odds) and tends to log-minus-log function when $\alpha_k \rightarrow 0$ (proportional hazard).

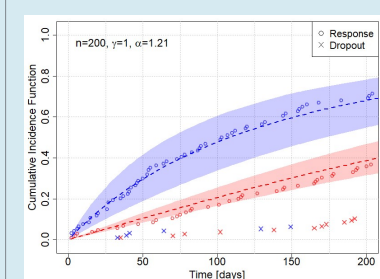
- Finally, we have the constraint that $\sum_{k=1}^m \hat{F}_k = 1$ as $t \rightarrow \infty$.

Results: Simulation of patient benefit (Fine-Gray regression on CIF)

- Note that all study setups with a $\gamma > 0$ would demonstrate efficacy of $COV = 1$, but patient benefit can only be demonstrated with manageable patient numbers (≤ 1000) when the drug effect is rather strong ($\gamma \geq 0.4$)
- Even with very large studies (8000 patients), a patient benefit cannot be shown when the drug effect on response is weak compared to its effect on dropout.
- In general, it takes larger studies to demonstrate patient benefit than to show drug efficacy.



Definition of patient benefit: A study shows a patient benefit when efficacious exposure is significantly ($p < 0.05$) linked to response despite its influence on dropout.



Simulation of a strong effect ($\gamma = 1.0$) of efficacious exposure ($COV = 1$) on response while the effect on dropout is also strong. With only 200 patients in the study, a treatment effect can be clearly demonstrated. See the separation of the blue from the red outcomes. The parametric function describes the data well. At 50 days, about 30% of the patients with efficacious exposure have a response, compared to 10% of the patients with non-efficacious exposure. The strong effect of exposure on response is dominating over the risk of dropout during the entire study.

Is the “HOW” dominating the “WHAT”?

- Cox proportional hazard modelling, which uses the cause-specific hazard approach, (Cox-PH) has been the method of choice in comparative survival analysis for over 40 years. This tool can determine a significant drug effect while “dropout due to lack of tolerability”, “dropout due to lack of efficacy”, and death are combined with “end of observation period” into one type of outcome called “right-censoring”.
- Consequently, the question of **efficacy** has dominated drug development, and a discussion of “WHAT” needs to be established during clinical drug development and before market approval has been limited. (Note: we discussed “method effectiveness” vs “use effectiveness” in the 1990s).
- Credit goes to a presentation at last year's PAGE meeting¹: “Background on estimands and why are they important”.
- For 20 years already, a method is available² to analyse drug effect and dropout simultaneously as competing events. Thus, the “HOW” has changed, shouldn't the “WHAT” change as well to focus on patient benefit?

Method: Regression on CIF

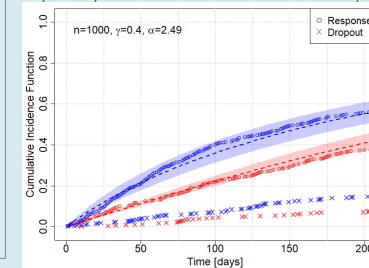
- Fine-Gray (1999)² offer a modification to the Cox proportional hazards model which allows direct regression on the CIF.
- The Fine-Gray method makes use of link functions in order to perform a semi-parametric linear regression, given by
$$g_k\{\hat{F}_k(t; \mathbf{X})\} = g_k\{\hat{F}_k(t; \mathbf{0})\} + \xi_k^T \mathbf{X}.$$
 where, $\hat{F}_k(t; \mathbf{X})$ is the CIF for event k , $\hat{F}_k(t; \mathbf{0})$ is the baseline CIF, ξ_k a vector of coefficients for event k , $\mathbf{X}$ is a vector of covariates and g_k are non-decreasing invertible link functions.
- Fine-Gray uses log-minus-log link function: $g(u) = \log(-\log(1 - u))$.
- This leads to a sub-distribution hazard equation that is analogous to the familiar form of the Cox proportional hazard:
$$\hat{h}_k(t|\mathbf{X}) = \hat{h}_{k0}(t) \cdot \exp(\xi_k \cdot \mathbf{X}).$$
- Fine-Gray models have been criticized on the grounds that their parameter estimates, which describe the relationships between covariates and event incidence, are difficult to interpret.
- This shortcoming can be overcome by visualising the covariate effect on the CIF with a parametric regression.

Method: A simulated study setup to assess patient benefit

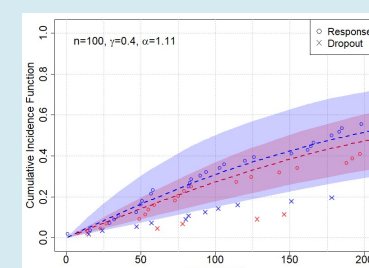
- A simplified oncology study is simulated where “response” and “dropout” are two competing events and right-censoring stems from the end of the observation period. Response signifies the efficacy the treatment. Dropout is considered to be linked to tolerability.
- A Weibull distribution was used to generate response and dropout event times for various setups, each consisting of 200 virtual studies:
 - Response hazard (varying between setups): $h(t) = 0.0015 \cdot t^{0.15} \cdot \exp(\gamma \cdot COV)$
 - Dropout hazard (fixed in all setups): $h(t) = 0.0002 \cdot t^{0.25} \cdot \exp(COV)$
 - ...where COV was sampled from either 0 or 1 with equal probability to represent a dichotomous drug exposure covariate (1 for above and 0 for below an efficacy threshold)
 - ...where different value of γ (from 0, 0.2, ..., 1.8, 2) signify the influence of exposure (COV) on the response (from “weak” to “strong”).
 - ...where influence of efficacious exposure ($COV=1$) on dropout is always strong (poor tolerability).
 - Study setups also differ by the number of patients (from 100, 200, ..., 8000).
- Simulated event times are created by sampling a number ρ from a uniform distribution between 0 and 1, and calculating $S^{-1}(\rho)$, where $S(t)$ is the survival function derived from the cause-specific hazard $h(t)$.
- A right-censoring time is generated for each patient ($\in$ study $\in$ setup) by sampling from a uniform distribution between 60 and 730 days.
- The observation time for each patient is the first of either a response, a dropout or right-censoring.

Results: Visualising the covariate effect on response and dropout (Parametric regression on CIF)

The outcome of a single study within a simulation setup is shown as the cumulative incidence of a treatment response while at risk of dropout. Note, that three different simulation setups are displayed. Compare count data (circles) with the fitted function (dashed line as median and shaded area as 95% prediction interval) to achieve a visual predictive check. The colours blue and red indicate the outcomes for patients with $COV = 1$ and $COV = 0$, respectively. The crosses count the incidence of dropout.



Simulation of a moderate effect ($\gamma = 0.4$) of efficacious exposure ($COV = 1$) on response while the effect on dropout is strong. With 1000 patients in the study, a treatment effect can be clearly demonstrated. See the separation of the blue from the red outcomes. The parametric function describes the data well. At 50 days, about 20% of the patients with efficacious exposure have a response, compared to 13% of the patients with non-efficacious exposure. At the same time 5% and 2% of patients with and without efficacious exposure have dropped out, respectively. Thus, patients are about 4 times more likely to have a response than to dropout.



Simulation of a moderate effect ($\gamma = 0.4$) of efficacious exposure ($COV = 1$) on response while the effect on dropout is strong. With 100 patients in the study, a treatment effect cannot be demonstrated. See the overlap of the blue and the red outcomes. The parametric function describes the data well. At 50 days, about an equal percentage (about 15%) of the patients with efficacious exposure and with non-efficacious exposure have a response. The chance of a patient to have a response is still higher than his risk to dropout. With the small study size, the difference is hard to quantify.

Conclusion

- We have demonstrated that weak treatment effects and poor tolerability result in no benefit.
- Benefit to the patient can be analysed by simultaneously modelling responses and dropout.
- When using parametric models for the covariate regression on the CIF, model results can be quantified with visual predictive checks

References:

- PAGE 27 (2018) Abstr 8792 [www.page-meeting.org/?abstract=8792].
- Fine, J. P., & Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American statistical association*, 94(446), 496-509.
- Shi, H., Cheng, Y., & Jeong, J. H. (2013). Constrained parametric model for simultaneous inference of two cumulative incidence functions. *Biometrical Journal*, 55(1), 82-96.