Master Modélisation Statistique

UFC UNIVERSITÉ DE FRANCHE-COMTÉ

Journée annuelle Biopharmacie & Santé
SFdS

Handling of possibly informative censoring for time to event endpoints & adaptation to treatment switching

Sanofi R&D

Supervisors: Loïc DARCHY – Gautier PAUX
University supervisor: Yacouba BOUBACAR-MAINASSARA

SANOFI

INTRODUCTION

The aim of this internship was to assess how to handle informative censoring with time-to-event endpoints.

Several methods will be evaluated in order to impute the time-to-event for patients who were censored. Methods will be assessed on:

- Simulated data
- Real data

Then they will be adapted to the specific case of treatment switching.

SANOFI

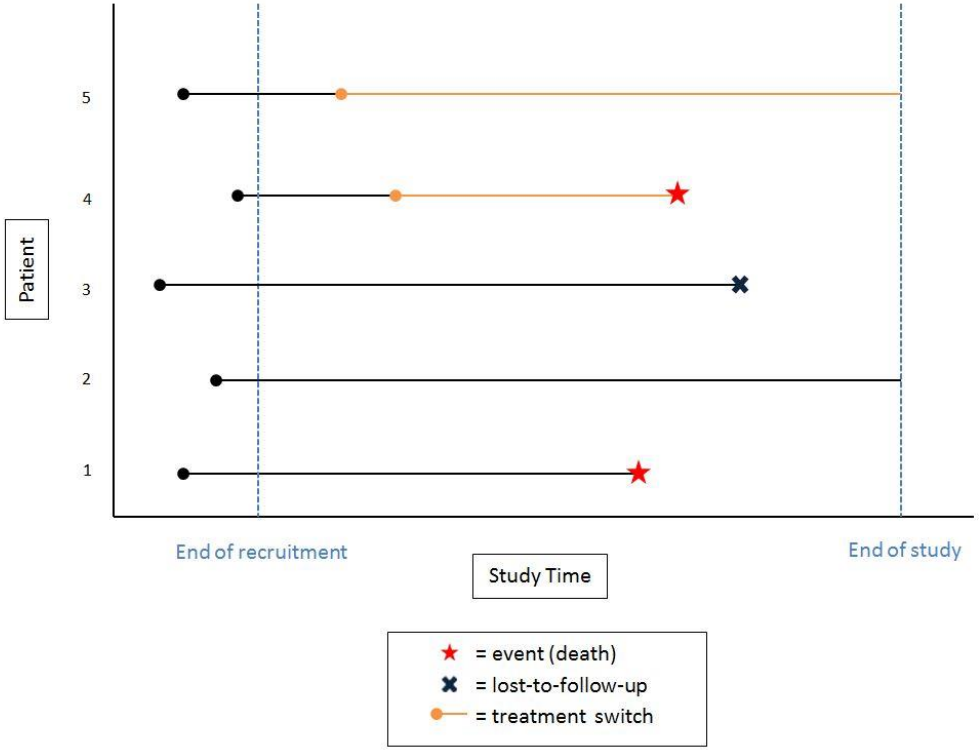
CONTENTS

SANOFI

REMINDER OF THE DEFINITION OF CENSORING

Time-to-event endpoints: primary outcome of interest is the time until the occurrence of a specific event (in our case: death)

Event not observed for patients ➔ censored data



DIFFERENTS FORMS OF CENSORED DATA

- **Administrative censoring:** when a patient has spent the whole study time without experimenting the event of interest
- **Lost-to-follow-up:** when we do not know if patient has experienced the event of interest or not.
- **Non-informative censoring:** lost-to-follow-up without any relation to the treatment and the event of interest.
- **Informative censoring:** reason of censoring is related to the treatment and the event of interest.

  CANNOT BE IGNORED

SANOFI

PRESENTATION OF THE SUBJECT

In presence of informative censoring, classical analyses assuming non-informative censoring lead to biased results.

➔ **MULTIPLE IMPUTATIONS METHODS:** impute event time for subjects who were censored.

ILLUSTRATIVE EXAMPLE

- Event of interest: death
- 2 groups of 1 000 000 subjects: control VS treatment
- Discontinuation time C_i ~ Exp (0.0082)
- Event time control group ~ Weibull (λ_0, k=0.5) with λ_0=0.007
- Before discontinuation: Event time for treatment group ~ Weibull (λ_1, k=0.5) with λ_1=0.00175

  → Hazard ratio: HR= 0.5 → β=log(HR)=-0,693

SANOFI

EXAMPLE

Lets assume conservatively that patients who discontinued from the treatment group have a greater risk of presenting the event after discontinuation while this risk is unchanged in the control group

➔ Modify the survival function by creating a piecewise hazard function:
- Before discontinuation: equal to the classical function knowing the model
- After discontinuation: a penalty e^{γ} is added for patients of the treatment group

$$h(t) = k\lambda_1^k t^{k-1} \times exp\left[\gamma 1(t > C_i)\right]$$

$$= \begin{cases} k\lambda_1^k t^{k-1} & \text{if } t \leq C_i \\ k\lambda_1^k t^{k-1} e^{\gamma} & \text{otherwise.} \end{cases}$$

SANOFI

EXAMPLE

$$h(t) = k\lambda_1^k t^{k-1} \times exp\left[\gamma 1(t > C_i)\right]$$

$$= \begin{cases} k\lambda_1^k t^{k-1} & \text{if } t \leq C_i \\ k\lambda_1^k t^{k-1} e^{\gamma} & \text{otherwise.} \end{cases}$$

The value of γ is chosen in order to have a log(HR) after discontinuation equal to: (1-Φ)β (Φ=0, 0.2, 0.4, … 1) → $\gamma = -\phi\beta$

$$S(t) = \begin{cases} e^{-(\lambda_1 t)^k} & \text{if } t \leq C_i \\ e^{-(\lambda_1 C_i)^k - \lambda_1^k e^{\gamma}\left(t^k - C_i^k\right)} & \text{otherwise.} \end{cases}$$

To impute the event time t, I generated a uniform random variable u and used the inverse function of S.

SANOFI

EXAMPLE

Cox model with treatment group as an exploratory variable on :

- Survival times with only administrative censoring (β^*) representing the true value of β
- Survival times censored at discontinuation times ($\tilde{\beta}$) representing the estimation of β under non-informative censoring

	True values of the HR and the log(HR)		Estimated values of the HR and the log(HR)		Empirical bias
Φ	β^*	HR	$\tilde{\beta}$	HR	$\tilde{\beta} - \beta^*$
0	-0,693	0,500	-0,693	0,500	0
0,2	-0,667	0,513	-0,693	0,500	-0,026
0,4	-0,639	0,528	-0,693	0,500	-0,054
0,6	-0,608	0,544	-0,693	0,500	-0,085
0,8	-0,576	0,562	-0,693	0,500	-0,117
1	-0,541	0,582	-0,693	0,500	-0,152

The use of non-informative censoring methods in presence of informative censoring leads to biased results.

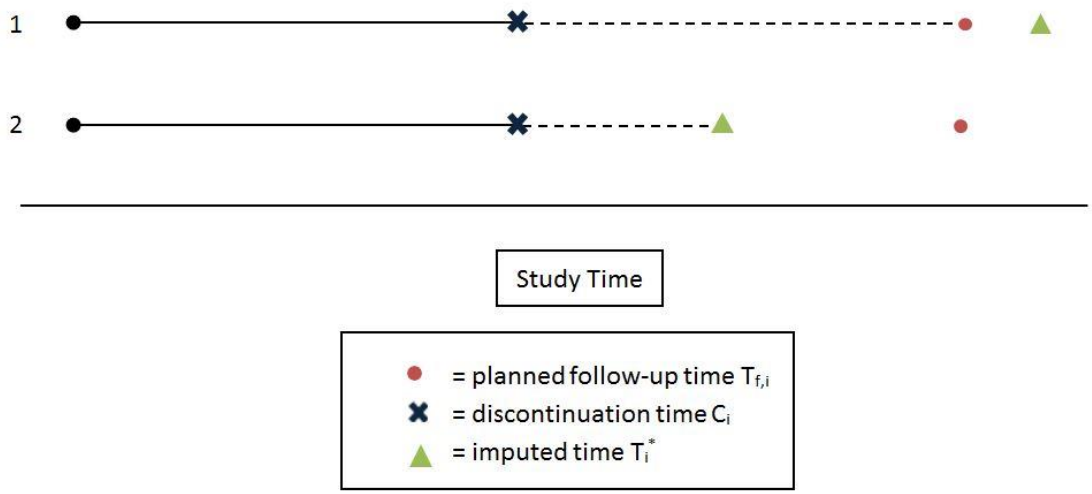
Here it tends to overestimate the treatment effect.

SANOFI

MULTIPLE IMPUTATION METHODS

We will use multiple imputation methods to address time to event for patients who discontinued in order to estimate the log hazard ratio.

The event times will be imputed between the discontinuation time and the planned follow-up time.

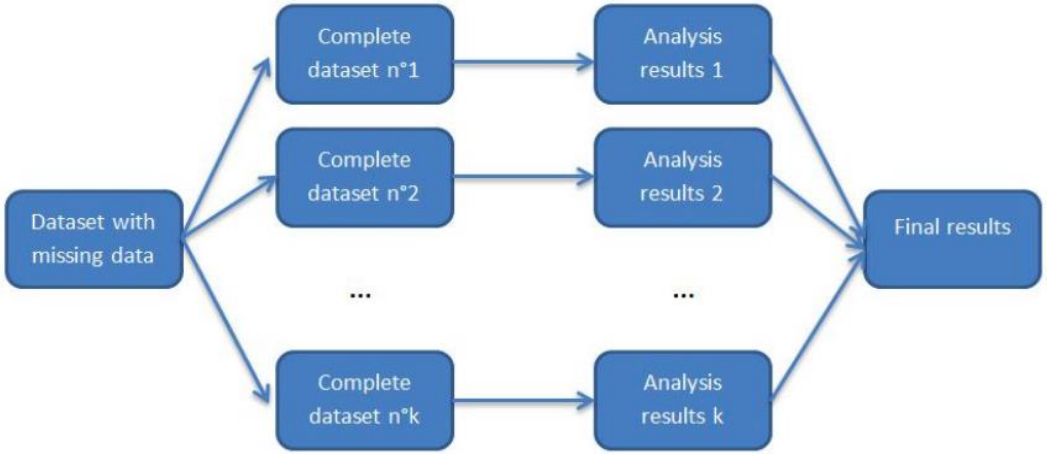


→ 2 methods
→ 2 baseline hazard functions

SANOFI

MULTIPLE IMPUTATION MECHANISM

1. Missing data are imputed k times
2. The k complete datasets are analyzed
3. Results from the k datasets are combined with Rubin's rule.



$$\hat{\theta} = \frac{1}{k}\sum_{i=1}^{k}\hat{\theta}_i \qquad Var(\hat{\theta}) = W + \left(1 + \frac{1}{k}\right)B$$

Where $W = \frac{1}{k}\sum_{i=1}^{k} Var(\hat{\theta}_i)$ and $B = \frac{1}{k-1}\sum_{i=1}^{k}(\hat{\theta}_i - \hat{\theta})^2$

SANOFI

NOTATIONS

- N : number of patients
- X_i: treatment indicator of patient i, 1 for the treatment arm and 0 for the control arm
- $T_{f,i}$: planned follow-up time of patient i
- C_i : discontinuation time of patient i
- T_i^*: event (death) time of patient i

- $T_i=\min(C_i, T_i^*, T_{f,i})$: observed survival time of patient i
- Δ_i : event indicator of patient i, 1 if the event of interest has been observed.

SANOFI

DELTA-ADJUSTED METHOD

The hazard of having an event for subjects of the treatment group who discontinued is multiplicatively increased relative to subjects who continued within this group.

Conservatively, for the control group, hazards with or without discontinuation are similar.

$$\begin{aligned}\mathbb{P}\left(dN_i(t)=1|X_i,N_i(t^-)=0,C_i,\mathbf{C_i}<\mathbf{t}\right) &= \theta\times\mathbb{P}\left(dN_i(t)=1|X_i,N_i(t^-)=0,C_i,\mathbf{C_i}\geq\mathbf{t}\right)\\ &= \theta\times e^{\beta X_i}\Lambda_0(t)\\ &= e^{(\beta+\delta)X_i}\Lambda_0(t) \qquad (6.2)\\ &= \begin{cases}\Lambda_0(t) & \text{if } X_i=0 \text{ (Control group)}\\ e^{\beta+\delta}\Lambda_0(t) & \text{if } X_i=1 \text{ (Treatment group)}\end{cases}\end{aligned}$$

$\theta=e^{\delta}$ with $\delta\geq 0$. The hazard for the dropouts will be multiplicatively increased by e^{δ}

If $\delta=0$ then hazards before and after discontinuation are the same: it represents non informative censoring.

SANOFI

REFERENCE-BASED METHOD

The hazard for subjects of the treatment group who discontinued the treatment lies between the hazard for experimental subjects who continued and the hazard of the control group.

$$\mathbb{P}\left(dN_i(t) = 1 | X_i, N_i(t^-) = 0, C_i, \mathbf{C_i} < \mathbf{t}\right) = e^{(1-\phi)\beta X_i}\Lambda_0(t) \qquad (6.3)$$

$$= \begin{cases} \Lambda_0(t) & \text{if } X_i = 0 \text{ (Control group)} \\ e^{(1-\phi)\beta}\Lambda_0(t) & \text{if } X_i = 1 \text{ (Treatment group)} \end{cases}$$

- Φ=0 represents non-informative censoring, it means that dropouts of the treatment group will have experience after discontinuation similar to the patients who continued within this group.
- Φ=1 implies that dropouts of the treatment group will have experience after discontinuation similar to the control group

SANOFI

BASELINE HAZARD FUNCTIONS

In both methods we need:

- β the log hazard ratio
- Λ_0 the baseline hazard function

We will consider two baseline hazard functions: a piecewise constant function and an unspecified function. The latter one is preferable since it makes no assumption,

2000 imputations for each value of ϕ / δ

Reference-based method	Delta-adjusted method
ϕ=0, 0.2, …, 1	δ=-$\phi\beta$ = 0, 0.136, …

SANOFI

EXAMPLE – SIMULATED DATA

We have simulated the data with the same design as previously with n=400 patients.

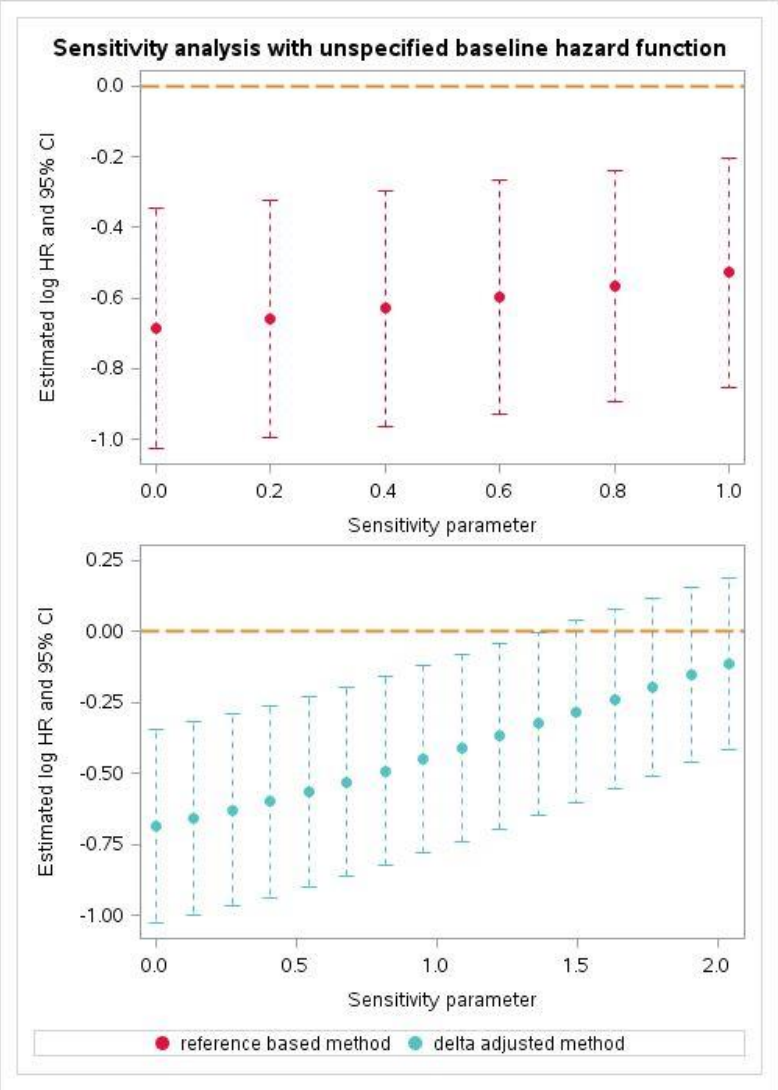
The HR and its 95% CI are equal to: 0.506 and [0.358;0.710]

We will use a tipping point analysis to find the value for which the conclusion of the statistical significance changes.

If this value is not relevant (too high) then it means that the conclusion under non-informative censoring is robust to possibly informative censoring.

SANOFI

EXAMPLE – SIMULATED DATA



SANOFI

EXAMPLE – SIMULATED DATA

CONCLUSION:

- Reference-based method: a significant treatment effect is obtained for each value of ϕ
- Delta-adjusted method: when $\delta \geq 1,3$ the results of the primary analysis is overturned. But $\theta = e^{\delta} \geq 3.9$ and in reality such an increase is unlikely to happen.

So we can conclude that the results of the primary analysis are robust to possibly informative censoring.

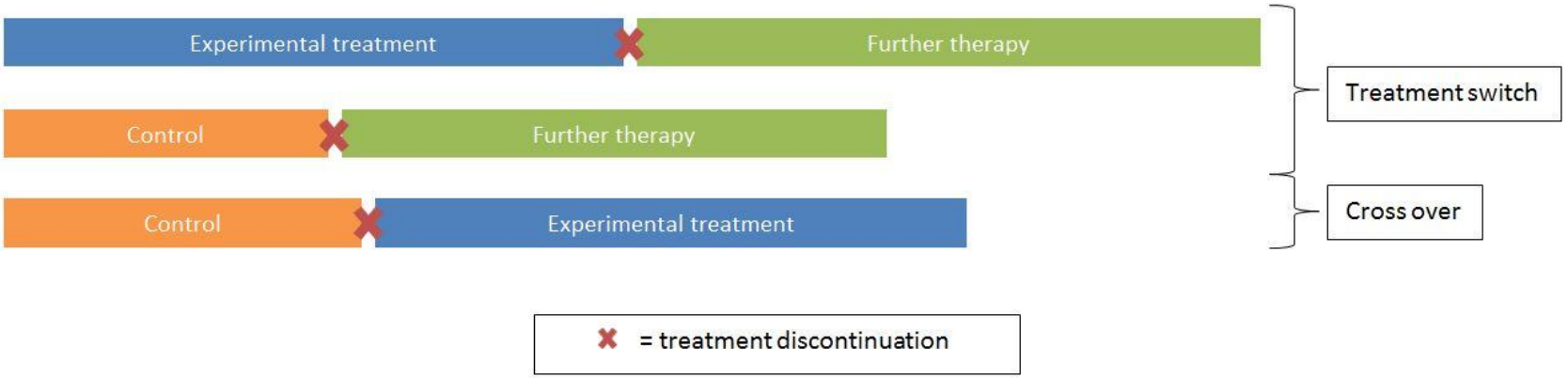
SANOFI

ADAPTATION TO TREATMENT SWITCHING

SANOFI

DEFINITION OF TREATMENT SWITCHING

In a study, patients who discontinued treatment are given the opportunity to receive another therapy called further therapy. Also, sometimes patients of the control group are allowed to switch to the treatment group, that is called a cross-over.



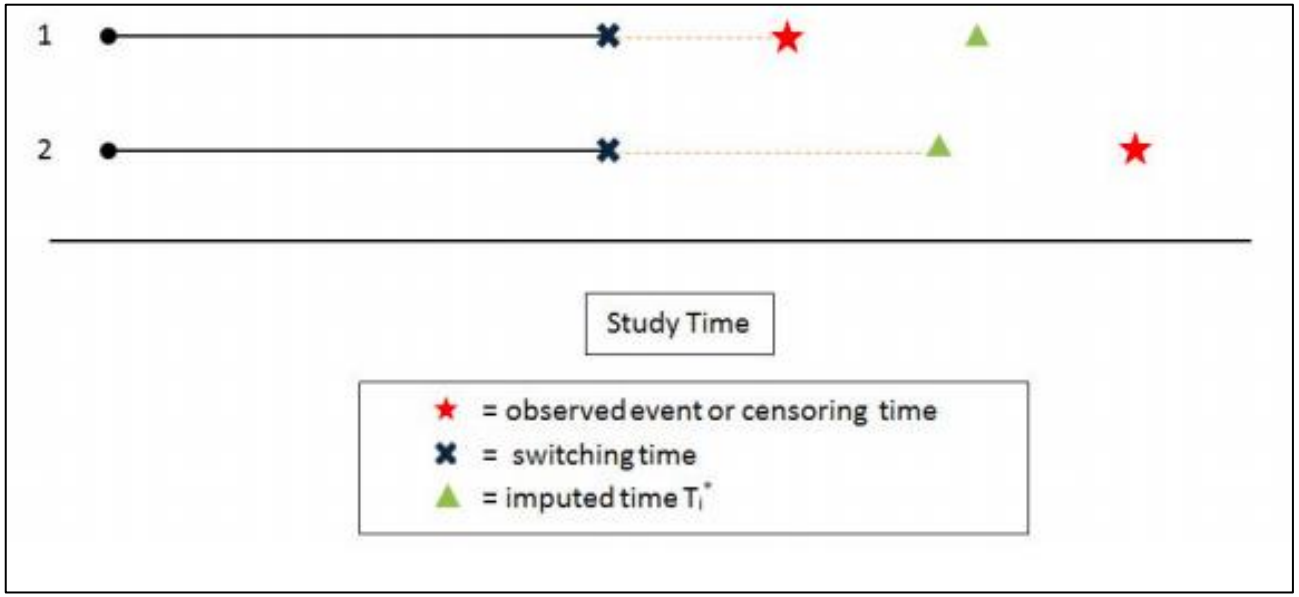
When switching occurs in a study, the event time for switchers takes the switch into account so we ignore what would have happened without switch

Standard ITT analyses are inappropriate to appraise treatment efficacy in the long run
→ Sensitivity analyses.

SANOFI

ADAPTATION OF THE METHODS

We will modify the methods seen before in order to impute the event time for switchers of both groups between the time of switch and the actual end date of the patient (death or censoring).



SANOFI

REFERENCE-BASED METHOD

The hazard for switchers of the treatment group is bracketed by :

- The hazard for non-switchers of the treatment group
- The hazard of the control group

The hazard for switchers of the control group is multiplicatively increased by $e^{\gamma C}$ relative to that for the patients who did not switch.

$$h(t) = e^{\gamma_C(1-X_i)+(\gamma_C+(1-\phi)\beta)X_i}\Lambda_0(t)$$

$$= \begin{cases} e^{\gamma_C}\Lambda_0(t) & \text{if } X_i = 0 \text{ (control group)} \\ e^{\gamma_C+(1-\phi)\beta}\Lambda_0(t) & \text{if } X_i = 1 \text{ (treatment group)} \end{cases}$$

If Φ=1 the hazard post switch is equal to the hazard post switch for patients of the control group

If Φ=0 the hazard post switch is equal to the hazard prior switch on top of the degradation $e^{\gamma C}$.

SANOFI

DELTA-ADJUSTED METHOD

The hazard of presenting the event after switch is multiplicatively increased relative to that for subjects who did not switch.

The multiplicative parameter depends on the treatment arm: it is equal to $e^{\delta C}$ for the control group and $e^{\delta T}$ for the treatment group, where: $\delta_C \geq 0$ and $\delta_T - \delta_C = \varepsilon \geq 0$

$$h(t) = \begin{cases} e^{\delta_C}\Lambda_0(t) & \text{if } X_i = 0 \text{ (control group)} \\ e^{\delta_T}e^{\beta}\Lambda_0(t) & \text{if } X_i = 1 \text{ (treatment group)} \end{cases}$$
$$= e^{(\beta+\delta_T)X_i+\delta_C(1-X_i)}\Lambda_0(t)$$

SANOFI

EXEMPLE – SIMULATED DATA

This example is used to illustrate the bias obtained when we use non-informative methods in the presence of informative censoring.

Again, we construct the data by considering the same design as previously with n=2 000 000 patients.

For each group, we construct a piecewise hazard function in which the hazard post switch is multiplicatively increased by either $e^{\gamma C}$ or $e^{\gamma T}$ relative to the hazard before switch.

$$h_{ctrl}(t) = k\lambda_0^k t^{k-1} exp\{\gamma_C 1_{t>S_i}\}$$

$$= \begin{cases} k\lambda_0^k t^{k-1} & \text{if } t \leq S_i \\ k\lambda_0^k t^{k-1} e^{\gamma_C} & \text{otherwise.} \end{cases}$$

$$h_{trt}(t) = k\lambda_1^k t^{k-1} exp\{\gamma_T 1_{t>S_i}\}$$

$$= \begin{cases} k\lambda_1^k t^{k-1} & \text{if } t \leq S_i \\ k\lambda_1^k t^{k-1} e^{\gamma_T} & \text{otherwise.} \end{cases}$$

SANOFI

EXEMPLE – SIMULATED DATA

For the reference based imputation, we want the log hazard ratio after switch to be equal to (1-Φ)β with Φ=0, 0.2, 0.4, …, 1.

We set γ_C to some plausible values and then calculate γ_T.

$$\gamma_T = \gamma_C - \phi\beta$$

			True values of the HR and the log(HR)		Estimated values of the HR and the log(HR)		Empirical bias
Φ	γ_C	γ_T	β*	HR	$\tilde{\beta}$	HR	$\tilde{\beta} - \beta^*$
0	0,47	0,47	-0,69	0,502	-0,693	0,500	-0,003
0,2	0,47	0,609	-0,656	0,519	-0,693	0,500	-0,037
0,4	0,47	0,747	-0,62	0,538	-0,693	0,500	-0,073
0,6	0,47	0,886	-0,582	0,559	-0,693	0,500	-0,111
0,8	0,47	1,024	-0,543	0,581	-0,693	0,500	-0,15
1	0,47	1,1631	-0,502	0,605	-0,693	0,500	-0,191

This example is used to illustrate the bias obtained when we use non-informative methods in the presence of informative censoring.

SANOFI

EXEMPLE – REAL DATA

VELOUR is a former Sanofi clinical trial comparing the effect of a treatment (aflibercept) versus a placebo in patients with metastatic colorectal cancer.

1226 patients were included: 614 in the control group and 612 in the treatment group.

The hazard ratio for overall survival (OS) obtained with the ITT analysis is equal to 0.809 and its 95% CI is [0.707;0.925]

In this study, 360 patients in each group had switched to a further therapy.

We have performed another analysis, where switchers were censored at their time of switch :

- HR = 0.837
- 95% CI = [0.686;1.021] : slightly not significant. The increase of the width is due to the decrease in the number of events: (460 and 403 VS 203 and 187)
- P-value = 0.0791 > 0.05

SANOFI

EXEMPLE – REAL DATA

Use the two multiple imputation methods (with the 2 baseline hazard functions) to impute the event time for switchers between the time of switch and the observed event or censoring time.

REFERENCE-BASED METHOD:

- Φ=0, 0.2, 0.4, …, 1
- γ_C=0, 0.2, …, 1.6

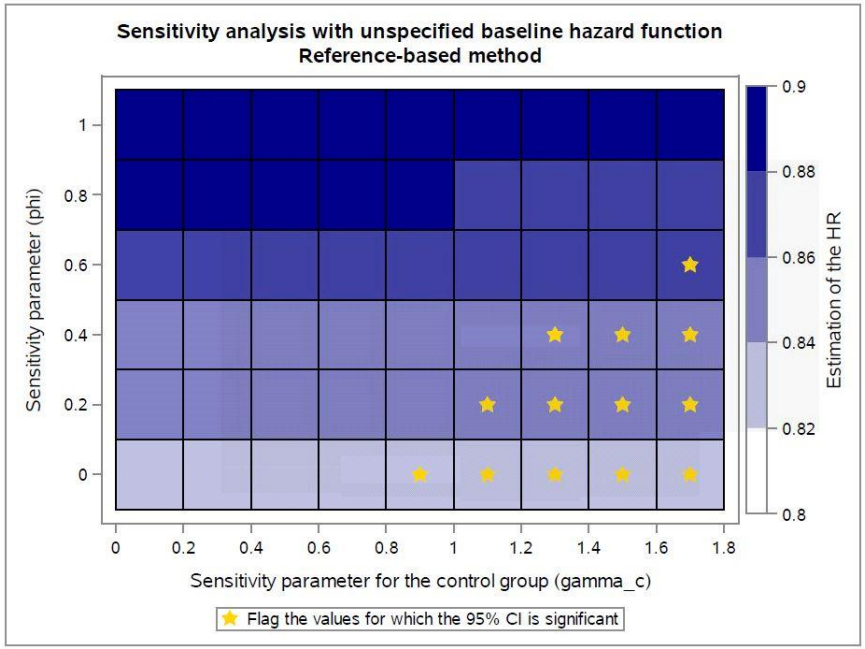
DELTA ADJUSTED METHOD

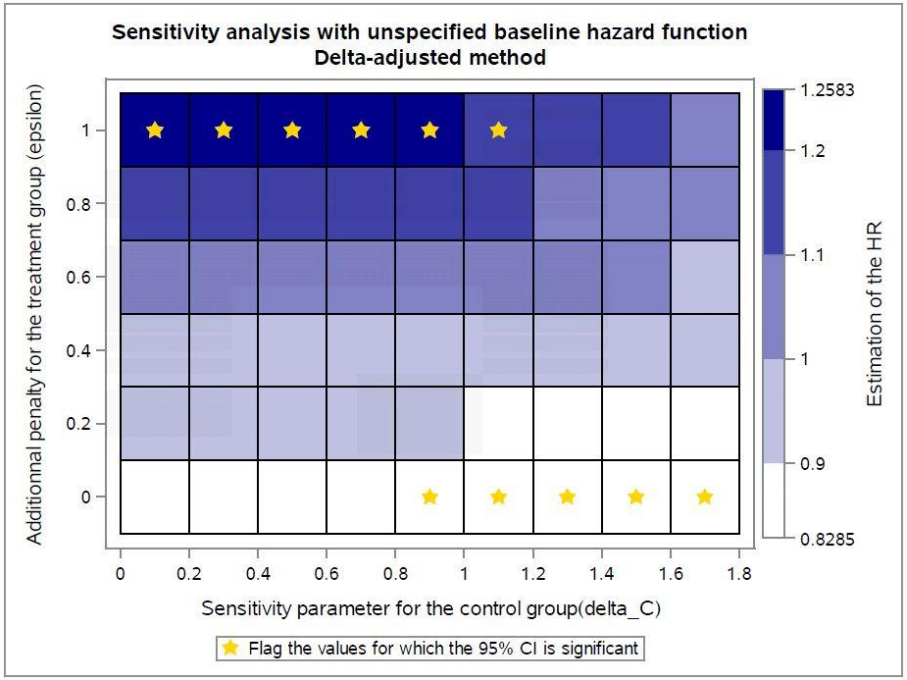
- δ_C=0, 0.2, …, 1.6
- ε=0, 0.2, …, 1

→ 2000 imputations for each couple of parameters

SANOFI

EXEMPLE – REAL DATA





EXEMPLE – REAL DATA

As the analysis censoring at switch (which is the reference analysis for the sensitivity analyses) is not significant, the sensitivity analyses are not relevant. Since these analyses are aimed to penalize the experimental treatment, they deteriorate the HR.

With $\phi=\gamma_C=0$ we get HR= 0.837 and 95% CI is [0,686; 1,020]

The case with $\phi=\gamma_C=0$ is not expected to modify the hazard ratio estimate but the additional events imputed beyond the switch could have improved the precision of the estimation.

But it is not the case : the between imputation variability compensates for the within-imputation precision gain.

SANOFI

CONCLUSION

SANOFI

CONCLUSION – DISCUSSION

We have studied several ways to perform sensitivity analyses that will adjust for the presence of informative censoring or switch within a study. Both MI methods are used to impute event times after switch.

In oncology, when the event of interest is death, by definition, it occurs after switch. Censoring observations at time of switch considerably reduces the number of events → a lot of event times are imputed.

- Develop these methods in other contexts than oncology (less loss of events)
- Study other methods available like the Rank-Preserving Structural Failure Time (RPSFT) model

SANOFI

THANK YOU FOR YOUR ATTENTION.

Any questions ?

BIBLIOGRAPHY

[1] D Collet. Modelling survival data in medical research. Chapman and Hall.

[2] Nicholas LATIMER and Keith ABRAMS. NICE DSU Technical support document: adjusting survival time estimates in the presence of treatment switching, jul 2014.

[3] Kaifeng LU, Dayong LI, and Gary G.KOCH. Comparison between two controlled multiple imputation methods for sensitivity analyses of time-to-event data with possibly informative censoring. Statistics in biopharmaceutical research, Vol. 7, No.3, aug 2015.

[4] James M.ROBINS and Dianne M.FINKELSTEIN. Correcting for noncompliance and dependent censoring in an aids clinical trial with inverse probability of censoring weighted log-rank tests. BIOMETRICS 56: 779-788, sep 2000.

[6] Claire WATKINS, Xin HUANG, Nicholas LATIMER, Yiyun TANG, and Elaine J.WRIGHT. Adjusting overall survival for treatment switches: commonly used methods and practical application. Pharmaceutical statistics, oct 2013.

SANOFI

APPENDIX

PIECEWISE CONSTANT BASELINE HAZARD FUNCTION

Let 0=a0 < a1 <…<aj=∞ be the partition of the time axis, chosen in order to:

- Have at least one event in each interval
- Have approximatively the same number of events in each interval

$$\lambda_0(t) = \sum_{j=1}^{J} \lambda_j 1(a_{j-1} \leq t < a_j)$$

λ are drawn from a gamma distribution with the following parameters:

- **Shape parameter:** dj is the number of event in interval j

$$d_j = \sum_{i=1}^{n} \Delta_i 1(a_{j-1} \leq t < a_j)$$

- **Inverse scale parameter**: $\sum_{i=1}^{n} S_j(T_i) e^{\tilde{\beta} X_i}$
  where Sj is the exposure time of patient i in interval j

SANOFI

PIECEWISE CONSTANT BASELINE HAZARD FUNCTION

The posterior draw of Λ_0 is given by: $$\tilde{\Lambda}_0(t) = \int_0^t \tilde{\lambda}_0(u)du$$

SANOFI

UNSPECIFIED BASELINE HAZARD FUNCTION

To construct the unspecified baseline hazard function we need to modify the Breslow estimate of $\Lambda_0(t)$ to allow the imputed event time to take any values on the time axis.

Let $t_1, \ldots, t_M$ be the M distinct event times.

The Breslow estimate of $\Lambda_0(t)$ is a step function with jumps at observed event times. Jumps and slopes at t_j are given by:

$$\alpha_j = \frac{d_j}{\sum_{i=1}^{n} 1(T_i \geq t_j) \times e^{\hat{\beta} X_i}} \text{ where } d_j = \sum_{i=1}^{n} 1(T_i = t_j, \Delta_i = 1)$$

$$\lambda_j = \frac{\alpha_j}{t_j - t_{j-1}}$$

As the imputed time can be greater than the last observed event time, we extrapolate the hazard function beyond this point :

$$\lambda_{M+1} = \frac{\sum_{j=M-f+1}^{M} \alpha_j}{t_M - t_{M-f}}$$

SANOFI

UNSPECIFIED BASELINE HAZARD FUNCTION

Finally, we construct the following piecewise constant baseline hazard function:

$$\lambda_0(t) = \sum_{j=1}^{M} \lambda_j 1(t_{j-1} \leq t < t_j) + \lambda_{M+1} 1(t \geq t_M),$$

An obtain the modified Breslow estimate of $\Lambda_0(t)$:

$$\hat{\Lambda}_0(t) = \int_0^t \lambda_0(u) du$$

Bayesian version :

- consider $\tilde{\beta}$ the posterior draw of β
- assume a non-informative prior for α
- draw α_j from a Gamma distribution with d_j as shape parameter and the following inverse scale parameter:

$$\sum_{i=1}^{n} Y_i(T_j) e^{\tilde{\beta} X_i}$$

SANOFI

EXAMPLE – SIMULATED DATA

Sensitivity parameters		Reference based method		Delta adjusted method	
ϕ	δ	Piecewise	Unspecified	Piecewise	Unspecified
0	0	-0.693	-0.686	-0.693	-0.686
0.2	0.136	-0.665	-0.659	-0.665	-0.658
0.4	0.272	-0.636	-0.629	-0.637	-0.629
0.6	0.408	-0.603	-0.598	-0.605	-0.598
0.8	0.544	-0.568	-0.564	-0.571	-0.564
1	0.680	-0.530	-0.527	-0.535	-0.529
	0.816			-0.498	-0.491
	0.952			-0.458	-0.451
	1.088			-0.416	-0.410
	1.224			-0.374	-0.367
	1.361			-0.330	-0.325
	1.497			-0.286	-0.282
	1.633			-0.243	-0.238
	1.769			-0.2	-0.195
	1.905			-0.157	-0.153
	2.041			-0.117	-0.113

SANOFI

IPCW

SANOFI

Inverse probability of Censoring Weighting

IPCW is another method used to estimate the treatment effect in presence of informative censoring.

- patients are censored at the time of switch
- remaining observations are weighted in order to remove selection bias due to the relation between the prognosis of the patient and his/her decision to switch
- weight depends on the probability to be censored
- weight are included in the calculation of the number of patients at risk

- Difficult to implement and to interpret for non-experts
- Limitation: no unmeasured confounders assumption

SANOFI

IPCW - Analyses

3 different analyses:

- Subjects censored by the minimum of time to loss to follow-up and time to treatment switch **for any reason**
- Subjects censored by the minimum of time to loss to follow-up and time to treatment switch **for progression only**
- Subjects censored by minimum of time to loss to follow-up and time to treatment switch **for medically related reasons** : PD or AE

SANOFI

IPCW - Example

Application on VELOUR study: 1226 patients

- 863 events (death) and 363 patients censored
  - 390 occurred before subjects switched to further therapy.
  - 411 occur after the subject switched for progression or adverse events
  - 337 for PD and 74 for AE

	Estimate of β	HR	95% CI	Number of events	Length of the 95% CI
ITT analysis	-0.212	0.809	[0.707 ; 0.925]	863	0.218
Analysis censoring at switch	-0.178	0.837	[0.686 ; 1.021]	390	0.335
First analysis	-0.172	0.842	[0.691 ; 1.026]	390	0.335
Second analysis	-0.232	0.793	[0.668 ; 0.940]	526	0.272
Third analysis	-0.274	0.760	[0.633 ; 0.914]	452	0.281

- First analysis: HR is the one that would have been observed if switches were prohibited.
- Second analysis: represents the HR if all patients had remained on their assigned treatment unless switching was for a progression disease
- Third analysis:, HR is the one that would have been observed if patients had to stay on the assigned treatment except for AE or PD.

SANOFI

IPCW - Example

	Estimate of β	HR	95% CI	Number of events	Length of the 95% CI
ITT analysis	-0.212	0.809	[0.707 ; 0.925]	863	0.218
Analysis censoring at switch	-0.178	0.837	[0.686 ; 1.021]	390	0.335
First analysis	-0.172	0.842	[0.691 ; 1.026]	390	0.335
Second analysis	-0.232	0.793	[0.668 ; 0.940]	526	0.272
Third analysis	-0.274	0.760	[0.633 ; 0.914]	452	0.281

The second analysis improves the estimation of the HR relative to the one obtained by censoring at switch (0,793 VS 0,837) and this results is significant → Progression certainly involves a higher risk of death.

The third analysis (when switches for AE are added) further improves the HR, this is unexpected since there are more switches for AE in the aflibercept arm (95) than in the placebo (36).

SANOFI

IPCW

	Placebo / Folfiri (N=614)	Alfibercept / Folfiri (N=612)	**Total**
Reason of switching:			
Adverse events	36 (10.0%)	95 (26.39%)	131
Disease progression	264 (73.33%)	188 (52.22%)	452
Other reason	57 (15.83%)	76 (21.11%)	133
Poor compliance to protocol	2 (0.56%)	1 (0.28%)	3
Subject lost to follow-up	1 (0.28%)	0	1
Total	360	360	720

SANOFI

IPCW – Calculation of weights

IPCW assigns a weight $W_i(t)$ to subject I, calculated as follows:

$$W_i(t) = \frac{K_i^0(t)}{K_i(t)}$$

$K_i(t)$ is the probability to remain uncensored until t for subject i and it is estimated with the Kaplan-Meier estimator based on the following model:

$$\lambda_C(t|X, \overline{V}^*(t), T^* > t) = \lambda_0(t)exp\left[\alpha_X^T V^*(t)\right]$$

With:

- X the treatment group
- $V^*(t)$ the values of the covariates included in the model
- T^* the censored event time

K_i^0 is the probability of being uncensored by time t.

SANOFI

IPCW – Calculation of weights

K_i is estimated by $\widehat{K}_i$ using the Kaplan-Meier estimator based on the previous model:

$$\hat{K}_i(t) = \prod_{j,U_j \leq t, \Delta_j = 0, X_j = X_i} \left[1 - \frac{(1-\Delta_j) exp(\hat{\alpha}_{X_i}^T V_i^*(U_j))}{\sum_{k=1}^n Y_k(U_j) exp(\hat{\alpha}_{X_i}^T V_k^*(U_j)) I(X_j = X_i)} \right]$$

$\widehat{K_i^0}$ is estimated by the usual KM estimator

$$\hat{K}_i^0(t) = \prod_{j; U_j \leq t, X_j = X_i} \left[1 - \frac{(1-\Delta_j)}{\sum_{k=1}^n Y_k(U_j)} \right]$$

Finally, the subject specific weights are estimated by:

$$\hat{W}_i(t) = \frac{\hat{K}_i^0(t)}{\hat{K}_i(t)}$$

SANOFI

IPCW - Example

1. Model the censoring time by applying a Cox model where the event of interest is censoring. For each one of the 3 definitions of censoring.

```
proc phreg ;
model (start, stop)*censorC(0)= <covariates>
run;
```

Time interval : one for each change in the value of time dependent variable

Flag for censoring: equal 1 if the patient is censored at time t=stop

SANOFI

IPCW - Example

1. Model the censoring time by applying a Cox model where the event of interest is censoring. For each one of the 3 definitions of censoring.

2. calculation of $\hat{K}_i(t)$ and $\hat{K}_i^0(t)$ in order to calculate $\widehat{W}_i(t)$

SANOFI

IPCW - Example

1. Model the censoring time by applying a Cox model where the event of interest is censoring. For each one of the 3 definitions of censoring.
2. calculation of $\hat{K}_i(t)$ and $\hat{K}_i^0(t)$ in order to calculate $\widehat{W}_i(t)$
3. Apply another Cox model, with the event as the event of interest

```
proc phreg ;
model (start, stop)*censorE(0) = group / rl;
weight w;
run;
```

Weight patients by $\widehat{W}_i(t)$

Flag for the event: equal 1 if the patient died at time t=stop

SANOFI