ADJUSTING OVERALL SURVIVAL FOR TREATMENT SWITCHING: APPLICATION OF COMMONLY USED METHODS TO A PHASE 3 ONCOLOGY TRIAL

Jing Wang, Pfizer Inc.

ABSTRACT

It's not uncommon to see in oncology trials that a higher proportion of patients in the control arm switch over to a subsequent anti-cancer therapy after treatment discontinuation, compared to patients in the experimental arm. The high degree of imbalance in subsequent anti-cancer therapy use between treatment arms will impact the assessment of treatment effect. Sensitivity analyses correcting for this confounding effect would be important. In the EMA document "Question and answer on adjustment for cross-over in estimating effects in oncology trials", several exploratory analysis methods were suggested (such as Inverse Probability of Censoring Weighting (IPCW) and Rank Preserving Structural Failure Time Models (RPSFTM)) . In this poster, we will apply these methods to a Phase 3 oncology trial and discuss the pros and cons of each method.

INTRODUCTION

The Javelin Renal 101 study is a phase 3, multicenter, multinational, randomized, open-label, parallel 2-arm study of avelumab in combination with axitinib versus sunitinib monotherapy as first-line systemic treatment for patients with advanced renal cell carcinoma. The key objectives of this study are to demonstrate that avelumab in combination with axitinib is superior to sunitinib monotherapy in prolonging progression-free survival (PFS) or overall survival (OS) in the first-line treatment of patients with advanced renal cell carcinoma for patients with programmed death ligand-1 (PD-L1) positive tumors and for patients irrespective of PD-L1 expression.

From March 29, 2016, through December 19, 2017, a total of 886 patients were enrolled into the study and randomized to receive avelumab plus axtinib (442 patients) or sunitinib (444 patients). 560 patients (270 randomized to avelumab plus axitinib and 290 randomized to sunitinib) had PD-L1-positive tumors.

At a pre-planned interim analysis based on a cutoff of 28 Jan 2019, avelumab in combination with axitinib demonstrated clinically meaningful and statistically significant benefit in improving PFS in both patients with PD-L1 positive tumors and in all patients irrespective of PD-L1 expression. However, even though the observed HR was in favor of avelumab plus axitinib for OS, the statistical significance was not achieved for OS (observed HR for OS in patients with PD-L1 positive tumors was 0.83 with 95% CI (0.596, 1.151); observed HR for OS in all patients irrespective of PD-L1 expression was 0.80 with 95% CI (0.616, 1.027)). Therefore, the study was continued to follow up on OS.

It was observed in the study that there was a high degree of imbalance in the use of subsequent anti-cancer therapy, especially PD-1/PD-L1 inhibitors, between treatment arms (Table 1). Sensitivity analyses correcting for the confounding of subsequent PD-1/PD-L1 inhibitors were explored.

TABLE 1. FOLLOW-UP ANTI-CANCER DRUG THERAPIES BY CATEGORY

Category	Avelumab + Axitinib (N=442)	Sunitinib (N=444)
Subjects with any follow-up anti-cancer treatments	138 (31.2)	227 (51.1)
Any PD-1 or PD-L1 Inhibitor	33 (7.5)	159 (35.8)
Any VEGF or VEGFR Inhibitor	118 (26.7)	123 (27.7)
Any Other Drug Therapy	46 (10.4)	68 (15.3)

METHODS

In the EMA document "Question and answer on adjustment for cross-over in estimating effects in oncology trials" published in 2018, four methods were mentioned as potential methods for adjusting overall survival for treatment switching:

1. Censoring at time of treatment switching
2. Inverse Probability of Censoring Weighting (IPCW)
3. Rank Preserving Structural Failure Time models (RPSFT)
4. 'Two-Stage' methods

"Censoring at time of treatment switching" method censor a patient at time of switchover to subsequent treatment. This method is easy to implement. However, since switchover is often related to the patient prognosis, this method may lead to informative censoring and a biased estimate of the treatment effect. Thus, this method is in general not recommended.

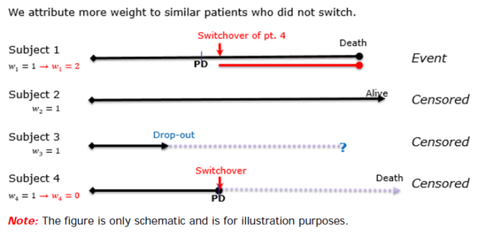
"Inverse Probability of Censoring Weighting" method constructs a pseudo-population that has the same specified (baseline and post-baseline) characteristics as the original population but did not switch over. That is, each time a patient switches over to subsequent treatment, IPCW censors that patient and let the remaining patients in the arm that are similar in terms of specified (baseline and post-baseline) characteristics, count for more patients to replace that patient (See Figure 1 for a simplified illustration of the weighting method). Important assumptions for this method, in order to get unbiased estimates, include 'no unmeasured confounders' assumption and 'positivity' assumption.

"Rank Preserving Structural Failure Time models" estimates the effect in absence of switchover by applying a slow-down factor to switchover patients to recover what event times would have been observed had these patients not switched over. It assumes that the slow-down factor is not dependent on the time of starting subsequent treatment ('common treatment effect' assumption). In our specific case, it also assumes that the effect of different PD-1/PD-L1 inhibitors are the same.

The idea of 'Two-stage' methods is similar to the RPSFT method. The method first estimates the effect specific to switchover (by comparing control patients who do and do not switchover) and use this effect to adjust survival times for patients that did switch over to obtain their counterfactual survival times in absence in switch over. Compared with the RPSFT method, this method uses a different approach in estimating the slow-down factor which assumes no unmeasured confounders.

In this poster, we will use both the IPCW method and the RPSFT method to correct the confounding of subsequent PD-1/PD-L1 inhibitors in the Javelin Renal 101 study.

FIGURE 1. IPCW SENSITIVITY ANALYSIS FOR SWITCHOVER



RESULTS

IPCW method

The IPCW modelling approach consists of a two-stage model fit:

- Fit censoring model: The probability of censoring due to switchover occurring after t for individual with covariate values x_0 and x_t is modelled with a time-dependent Cox hazards:

$$\widehat{S_c}(t;\, x_0, x_t) = [\widehat{S_{c0}}(t)]^{\exp(\widehat{\beta}' [x_0, x_t])}$$

- Compute IPCW weights: The subject-specific IPCW weights are calculated for each switchover time in the data set as inverse of the probability for censoring, i.e.:

$$w_i(t) = \frac{\widehat{S_{KM}}(t)}{\widehat{S_{ci}}(t;\, x_0, x_t)}, i = 1, \ldots., n$$

where $\widehat{S_{KM}}(t)$ is the treatment-specific Kaplan-Meier estimator of the probability of being uncensored by time t. $\widehat{S_{ci}}(t; x_0, x_t)$ is the conditional probability that a subject is uncensored through time t given the covariate values x_0 and x_t predicting the probability for censoring as obtained in the previous step. The multiplication with the Kapan-Meier estimate $\widehat{S_{KM}}(t)$, independent of explanatory variables, leads to stabilized weights and avoids computational problems.

- Estimate the IPCW version of the survival curve: Fit a weighted Cox model to estimate the adapted treatment effect and corresponding confidence intervals.

It is important to specify all the relevant baseline and post-baseline covariates (i.e., x_0 and x_t) in the model in order to obtain unbiased estimates of the treatment effect. We explored how different the treatment effect estimates would be for different sets of covariates included in the model using the Javelin Renal 101 data. Table 2 presents the estimated OS effect of avelumab plus axitinib versus sunitinib after adjusting for the confounding of subsequent PD-1/PD-L1 inhibitors (I/O) using IPCW method.

TABLE 2. OS: AFTER ADJUSTING FOR SUBSEQUENT I/O BY IPCW

Avelumab + Axitinib vs Sunitinib	HR (95% CI)
Model: only included baseline covariates (AIC=2242)	0.91 (0.680, 1.235)
Model: included both baseline and post-baseline covariates (AIC=1939)	0.73 (0.521, 1.028)
Model: included both baseline and post-baseline covariates AND only correct for subsequent I/O in the sunitinib arm (to be comparable to the RPSFT results)	0.71 (0.526, 0.978)

In general, the IPCW method is compute-intensive as the data needs to be split into the count process format (i.e., a separate record at every timepoint a patient in the study switches over). The time-dependent post-baseline covariates also need to be derived for each time interval accordingly.

RPSFT method

The RPSFT method was initially developed to evaluate the impact of crossover (i.e. switching from a randomized treatment allocation to another study treatment) on overall survival.

Under the 'common treatment effect' assumption, the RPSFT model relates the observed outcome (S_i) for patient i to the experimental treatment-free outcome (U_i) through an acceleration factor (ψ) by multiplying survival time on experimental treatment by a ratio $\exp(\psi)$ to the portion of time on treatment.

$$U_i = \int_0^{S_i} \exp[\psi TRT_i(t)]\, dt$$

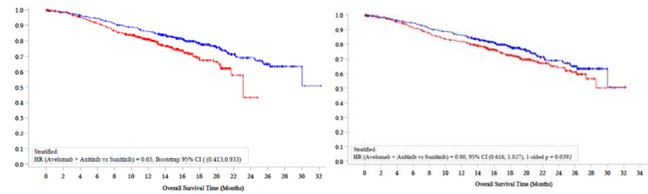
$TRT_i(t)$=binary treatment indicator (1=exp, 0=control) at time t.

ψ could be estimated through grid searching for values (ψ^*) to equalize $U_i(\psi)$ between the two randomized arms.

The observed survival time with crossover (S_i) could then be adjusted to the survival time T_i that would have been observed if there was no crossover. The treatment effect correcting for the confounding of crossover could then be estimated accordingly.

When the RPSFT method is used for adjusting survival for treatment switching to subsequent treatments, further assumption needs to be made. In our specific case in the Javelin Renal 101 study, we assumed that the effect of other subsequent PD-1/PD-L1 inhibitors on OS is the same as the effect of avelumab plus axitinib. The Kaplan-Meier plot of overall survival after adjusting for subsequent PD-1/PD-L1 inhibitors using RPSFT method is presented in Figure 2, compared with the original KM plot of OS without adjustment.

FIGURE 2. KM PLOT OF OS (WITH VS. WITHOUT ADJUSTMENT BY RPSFT)



CONCLUSIONS

This poster presented the commonly used methods for adjusting overall survival for treatment switching and applied the IPCW method and RPSFT method to the Javelin Renal 101 data. Both methods have strong assumptions for which, violations might have a big impact on the estimation of the treatment effect. In practice, the more appropriate method should be that for which the assumptions seem most reasonable or robust to deviations.

REFERENCES AND ACKNOWLEDGMENTS

[1] Motzer, R.J., et al. (2019). Avelumab plus axitinib versus sunitinib for advanced renal-cell carcinoma. *The New England Journal of Medicine.*

[2] European Medicines Agency. (2018). Question and answer on adjustment for cross-over in estimating effects in oncology trials.

[3] Robins, J.M. and Finkelstein, D.M. (2000). Correcting for noncompliance and dependent censoring in an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics.*

[4] Collett, D. (2015). Modelling survival data in medical research. *CRC Press*.

[5] Robins J. M. and Tsiatis A.A. (1991). Correcting for non-compliance in randomized trials using rank preserving structural failure time models. *Commun Stat Theory Methods*.

The author would also like to thank Juliane Manitz, Yiyun Tang and Bo Huang for their help in preparing this poster.