Comparative Effectiveness Research/Health Technology Assessment (HTA)

Analyzing Overall Survival in Randomized Controlled Trials with Crossover and Implications for Economic Evaluation

Linus Jönsson, MD, PhD1,2*, Rickard Sandin, PhD2, Mattias Ekman, PhD1, Joakim Ramsberg, PhD1, Claudia Charbonneau, PhD3, Xin Huang, PhD4, Bengt Jönsson, PhD5, Milton C. Weinstein, PhD6,7, Michael Drummond, PhD8

1OptumInsight AB, Klarabergsviadukten, Stockholm, Sweden; 2Global Health Economics and Outcomes Research, Pfizer Oncology, Sollentuna, Sweden; 3Global Outcomes Research, Specialty Care BU, Pfizer, Paris, France; 4Oncology Clinical Development, San Diego, CA, USA; 5Stockholm School of Economics, Stockholm, Sweden; 6Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA; 7OptumInsight, Medford, MA, USA; 8Centre for Health Economics, University of York, York, North Yorkshire, UK

A B S T R A C T

Background: Offering patients in oncology trials the opportunity to cross over to active treatment at disease progression is a common strategy to address ethical issues associated with placebo controls but may lead to statistical challenges in the analysis of overall survival and cost-effectiveness because crossover leads to information loss and dilution of comparative clinical efficacy. Objectives: We provide an overview of how to address crossover, implications for risk-effect estimates of survival (hazard ratios) and cost-effectiveness, and how this influences decisions of reimbursement agencies. Two case studies using data from two phase III sunitinib oncology trials are used as illustration. Methods: We reviewed the literature on statistical methods for adjusting for crossover and recent health technology assessment decisions in oncology. Results: We show that for a trial with a high proportion of crossover from the control arm to the investigational arm, the choice of the statistical method greatly affects treatment-effect estimates and cost-effectiveness because the range of relative mortality risk for active treatment versus control is broad. With relatively frequent crossover, one should consider either the inverse probability of censoring weighting or the rank-preserving structural failure time model to minimize potential bias, with choice dependent on crossover characteristics, trial size, and available data. A large proportion of crossover favors the rank-preserving structural failure time model, while large sample size and abundant information about confounding factors favors the inverse probability of censoring weighting model. When crossover is very infrequent, methods yield similar results. Conclusions: Failure to correct for crossover may lead to suboptimal decisions by pricing and reimbursement authorities, thereby limiting an effective drug’s potential. Keywords: cost effectiveness, crossover, oncology, sunitinib, survival.

Introduction

Allowing patients the opportunity to switch to investigational therapy after the primary end point has been reached is a common strategy used to address ethical issues with the use of placebo-controlled randomized trials. This is common practice in oncology trials, often mandated by investigators, patients, and ethics committees [1]. Crossover can also occur when a trial is prematurely unblinded; for example, when an interim analysis shows a significant gain in the primary end point of the investigational treatment or if an active treatment (control or investigational) is less safe than its comparator. In each case, crossover results in loss of information about what the clinical effect would have been in the absence of crossover. A direct consequence of crossover is that standard statistical methods, for example, the intent-to-treat (ITT) analysis, may provide biased estimates of key end points such as overall survival (OS), for instance, and underestimation of the true effect. Furthermore, per protocol


Conflicts of interest: Linus Jönsson, Mattias Ekman, and Joakim Ramsberg were employees of OptumInsight (Stockholm, Sweden), who were paid consultants to Pfizer, Inc, in the development of this manuscript and production of the analyses. Joakim Ramsberg, Milton C. Weinstein, and Michael Drummond have all had an advisory role with Pfizer, Inc. Joakim Ramsberg has received honoraria from Pfizer, Inc, and Michael Drummond has received other remuneration from Pfizer, Inc. Rickard Sandin, Claudia Charbonneau, and Xin Huang are Pfizer, Inc, employees with stock ownership. Bengt Jönsson has no conflicts of interest to disclose.

* Address correspondence to: Linus Jönsson, Karolinska Institutet, SE-171 77 Stockholm, Sweden.

E-mail: linus.jonsson@ki.se.

1098-3015$36.00 – see front matter Copyright © 2014, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

http://dx.doi.org/10.1016/j.jval.2014.06.006
analysis (excluding patients who cross over to investigational treatment) can be subject to selection bias because those who switch from placebo to investigational treatment may not be representative of the entire placebo group. Design solutions such as randomizing crossover can be considered to minimize the impact of crossover [2] but will seldom be feasible and have rarely been implemented in practice. A more common approach is to adjust for crossover effects in the statistical analysis.

Crossover is also of concern in health economics and outcomes research because of its potential to affect estimates of efficacy and cost-effectiveness (CE). If an investigational drug reduces mortality, ITT analysis will underestimate the treatment effect in the presence of crossover and will likely lead to an overestimate of the incremental cost-effectiveness ratio (ICER). As a result, the decisions by pricing and reimbursement agencies regarding access to new therapies may not maximize health outcomes with the available resources if crossover is not corrected for.

In this article, directed toward policymakers, health care providers, health technology assessment agencies, and the pharmaceutical industry, we review available standard and advanced statistical methods for analyzing OS data in the presence of crossover and discuss choice of methodology. We illustrate differences between four methods with two case studies based on clinical trials in two indications for sunitinib (Sutent; Pfizer, Inc., New York, NY), an orally administered, multitargeted tyrosine kinase inhibitor, which is approved in several countries for the treatment of metastatic renal cell carcinoma (mRCC) and imatinib-resistant gastrointestinal stromal tumor (GIST) [3–8]. Both trials showed statistically significant benefit in progression-free survival (PFS) for sunitinib, and each illustrates a different situation regarding crossover. In the mRCC trial, crossover was infrequent and allowed only after overwhelming PFS results were observed [7], whereas in the GIST trial, crossover was frequent and allowed because of the placebo control design [4,9]. As we shall see, the two cases lead to interesting and potentially instructive differences in the use and outcomes of the four methods with respect to estimated treatment benefits and cost-effectiveness.

**Statistical Analysis of Trials with Crossover**

Statistical methods that are used to evaluate OS can be grouped into simple methods, which make no specific attempts to address crossover, and advanced methods based on statistical modeling techniques, which attempt to eliminate or reduce bias due to crossover.

**Simple Methods**

**ITT analysis**

In the standard ITT analysis, data for all randomized subjects are included for the entire period of observation. This method is appropriate as long as the aim is to compare one planned treatment with another irrespective of any subsequent treatment changes. If an investigational drug has a true mortality benefit, however, the ITT analysis will underestimate OS in the presence of crossover, and a cost-effectiveness analysis based on these data will likely overestimate the ICER of the new therapy [10].

Censoring at crossover (on-treatment analysis)

Censoring patients at crossover eliminates observations of patients randomized to the control arm after they receive the investigational treatment. Two disadvantages of the censoring method are selection bias and loss of power. Unless the probability of crossover is random, censoring may introduce bias, because events that result in censoring (e.g., progressive disease) may likely be associated with the final outcome (e.g., death) [2]. Censoring therefore leads to underestimation of gains in OS with the investigational treatment and selective exclusion of patients with a high probability of death. Censoring also lowers the power of the study because of shorter overall observation time and a reduced number of observed events in the control arm.

The potential selection bias induced by censoring can be reduced or eliminated if crossover is determined by randomization or if the entire control group crosses over to the investigational treatment at a prespecified point in time [2]. Even in these situations, however, censoring still reduces the statistical power of the study.

**Statistical Modeling**

During the last two decades, statistical modeling techniques have been developed to adjust for weaknesses in observational and clinical trial data. There is no criterion standard because the methods have different strengths and weaknesses. Two methods that have recently been used to attempt to correct for crossover in oncology trials are the inverse probability of censoring weighting (IPCW) model and the rank-preserving structural failure time (RPSFT) model. These apply statistical modeling techniques to reconstruct data for the control arm as if crossover had not occurred, with the aim of reducing bias and allowing the treatment effect to be assessed more accurately. Results based on these methods have been considered as relevant by health technology assessment bodies such as the National Institute for Health and Care Excellence (NICE) in the United Kingdom [11] and the Dental and Pharmaceutical Benefits Agency (tandvårdsoch läkemedelsförmånsverket) in Sweden [12].

The IPCW model

The IPCW model is frequently used in epidemiologic research to adjust for nonrepresentative sampling or dropouts. In this method, patients who cross over from control to investigational treatment are censored, while patients remaining in the control arm are weighted to compensate for missing data [13]. The bias introduced by this informative crossover is corrected by weighting each patient by the inverse of his or her predicted probability of not being censored at a given time. The first step in the IPCW analysis is to predict the probability of crossover on the basis of each patient’s baseline characteristics, such as age, sex, race, or biological markers [13–15], often by fitting a logistic regression model. Finally, OS is analyzed with the censored data set and observations weighted by the inverse of the predicted probability of censoring.

The IPCW model assumes that the probability of crossover at a given time depends only on observed covariates and must be independent of the outcome and its timing [13]. If these assumptions hold, then censoring can be made noninformative through the IPCW model. The clinical trial data must contain enough information about the covariates that affect the probability of crossover.

The RPSFT model

The RPSFT model allows a direct comparison of randomization groups by adjusting the OS of patients who cross over so that it reflects the OS had they not received the investigational treatment. The method is related to the accelerated failure time model in OS analysis [16,17], in which prognostic variables measured on the individual level are assumed to act multiplicatively on the time scale, for example, affecting the rate of progression. The first step is to define a causal model relating the observed event time T to the unobserved event time U that would have been observed if crossover had not occurred. This is performed by assuming that T
and \( U \) are related by a constant acceleration factor, \( \exp(\psi) \), which is equal for all patients who cross over:

\[
T_i = T_i^{\text{off}} + T_i^{\text{on}}, \\
U_i = T_i^{\text{off}} + \exp(\psi) T_i^{\text{on}},
\]

where \( T_i \) is the total observed time on control therapy, \( T_i^{\text{on}} \) is the observed time on treatment, \( T_i^{\text{off}} \) is the observed time off treatment, and \( U_i \) is the time that would have been observed if crossover had not occurred. The parameter \( \psi \), the causal effect of treatment on OS, is estimated through G-estimation, whereby \( U \) is computed for a range of possible values of \( \psi \) and finding the value for which a log-rank test of the equality of \( U \) across the two groups has the highest \( P \) value [16]. The second step is to perform the OS analysis on the data set after OS has been adjusted (\( U \) values) for crossover cases in the control group. Event times that fall beyond the time horizon of the study need to be censored, a process referred to as recensoring.

The RPSFT model is rank preserving because a constant factor is used for adjusting the time to event for each patient [16]. Thus, if two patients \( i \) and \( j \) are on the same treatment (either control or crossover), and patient \( i \) fails (dies) before patient \( j \), before adjustment, patient \( i \) will also always fail before patient \( j \) after adjustment: the ranking in failure times is preserved. The model is structural (causal) in the sense that it assumes a defined relationship between the observed event time and the event time that would have been observed if crossover had not occurred [10].

A key assumption of the RPSFT model is that the investigational treatment causes a constant reduction in time to death, assumed equal for all patients before and after progression. This may be a reasonable assumption in some cases but not in others, which may restrict the use of the method to cases in which a constant proportional reduction in the time to event is biologically plausible.

Other methods

There are additional statistical methods that could prove useful in analyzing randomized trials with crossover. Structural nested mean models [19] can be useful for trials in which the primary end point is a measurement (e.g., mm Hg) rather than an event rate. State-transition probability modeling [19] has been used to analyze observational data with time-dependent effects and could potentially be adapted to address the issue of crossover in randomized trials.

Case Studies

Case 1. Sunitinib in mRCC

In an international phase III trial, 750 patients with mRCC were randomized to receive either sunitinib (\( n = 375 \)) or interferon-alfa (IFN-\( \alpha \); \( n = 375 \)) [6,7]. Crossover was allowed only after an interim analysis had concluded a significant gain in the primary end point PFS. Twenty-five patients (7%) in the IFN-\( \alpha \) group crossed over to sunitinib after an average of 70.8 weeks. Although relatively low, the amount of crossover could still potentially affect results.

There were 390 deaths in total: 190 in the sunitinib and 200 in the IFN-\( \alpha \) arm, respectively. All events were included in the ITT analysis, while censoring at the time of crossover led to the exclusion of five deaths in the IFN-\( \alpha \) arm occurring after the switch to sunitinib. The IPCW model used progressive disease, male gender, young age, and nephrectomy as covariates for estimating the propensity scores for crossover. All four covariates significantly increased the risk of crossover. In the RPSFT model, the estimated value for the acceleration parameter \( \psi \) using a grid search method was -0.244, corresponding to a decrease in the OS time by \( 1 - \exp(\psi) = 22\% \) with IFN-\( \alpha \) than with sunitinib.

The hazard ratios (HRs) for OS between the experimental treatment and control for the four different methods of handling crossover ranged between 0.807 and 0.821 (Fig. 1), their similarity explained by limited crossover in the trial. As expected, the ITT method produced the highest HR because the effect of sunitinib on OS was included in both arms. Compared with censoring at crossover, the IPCW resulted in a slightly lower HR estimate, but a somewhat wider confidence interval (CI) due to the additional uncertainty introduced by the estimation of inverse probability weights. The slightly lower HR with the RPSFT indicates that OS after crossover was slightly better than “expected” for patients in the IFN-\( \alpha \) arm, but the difference was small.

The CE of sunitinib in mRCC was evaluated by NICE [20]. Applying the manufacturer’s model, the ITT method yielded an expected OS of 163 versus 136 weeks for sunitinib versus IFN-\( \alpha \) and an ICER of £72,000/quality-adjusted life-year (QALY) (Table 1). Censoring at crossover provided an ICER of £71,760/QALY [21]. Structural models were not used for the evaluation. Applying the RPSFT adjustment changed the incremental OS gain with sunitinib only marginally (by <1 week) and similarly reduced the ICER only marginally to £71,850/QALY. NICE finally based its

![Fig. 1 - Hazard ratios (with 95% confidence intervals) for overall survival with sunitinib compared with interferon-alfa in patients with metastatic renal cell carcinoma using four different statistical methods to account for crossover; in this example, there was a low rate of crossover (7%) [7]. The structural method estimates were not adjusted for poststudy treatment, and 95% confidence intervals were estimated using nonparametric bootstrapping. No recensoring was performed in the RPSFT analysis. IPCW, inverse probability of censoring weighting; ITT, intent-to-treat; RPSFT, rank-preserving structural failure time.](image-url)
recommendation for general reimbursement on an OS analysis for a subset of the population that did not receive any poststudy treatment (including crossover) with estimated ICER below the £50,000 threshold.

**Case 2. Sunitinib in GIST**

An international phase III study randomized 361 patients with GIST in a 2:1 ratio to sunitinib (n = 243) or placebo with best supportive care (n = 118) [4,9]. Crossover was allowed in the protocol at documented progression, and additional crossover was allowed as a result of early recommended unblinding due to longer time to progression for sunitinib.

A total of 103 patients (87%) in the placebo group crossed over to sunitinib. The median time to progression for the ITT placebo population was 6.4 weeks (95% CI 4.4–10.0). Patients crossed over on average 4.1 weeks after progression (range –1 day to 32 weeks; the negative lower value is due to potential discrepancies in the date of progression as determined by central reading and the investigator). A total of 266 patients died: 90 in the placebo arm and 176 in the sunitinib arm (2:1 randomization).

The HRs for OS between experimental treatment and control for the three different methods of handling crossover ranged from 0.505 to 0.876 (Fig. 2). Censoring at the time of crossover led to the exclusion of 77 events in the placebo arm, with only 13 events remaining for the OS analysis. Calculating a Cox regression based on only 13 events leads to considerable uncertainty, as reflected by the 95% CI stretching from 0.454 to 1.499 (HR = 0.825).

There was a high proportion of crossover in the trial; out of 73 patients in the placebo group whose disease progressed, 69 crossed over to sunitinib. Because disease progression was strongly correlated with crossover (P < 0.001), and also with the final outcome of death, the simple censored analysis leads to bias. In the RPSFT model, the estimated value for the acceleration parameter $\psi$, using a grid search method [10], was equal to –0.656. The resulting HR for OS comparing sunitinib with placebo after recensoring was estimated at 0.505 [22]. The IPCW model could not be applied in this case because the number of remaining events after censoring was too few (2 in the placebo group).

Again, the ITT method produced the most conservative efficacy estimate. Censoring at crossover gave a lower HR with wider CIs because all the information on OS after crossover was discarded. The RPSFT model yielded the most favorable HR, although it fell short of statistical significance. The substantial difference in the point estimate compared with the ITT analysis indicates that OS was improved after crossover; however, the uncertainty introduced by the RPSFT model was illustrated by wider CIs.

The CE of sunitinib in GIST was evaluated by NICE [23]. The base-case analysis used the RPSFT model, which yielded an expected OS of 73 weeks versus 39 weeks for sunitinib compared with placebo with best supportive care and an ICER of £31,800/QALY (Table 1). In contrast, the ITT method showed an ICER of £90,500/QALY with an expected OS of 65 weeks for patients on placebo with best supportive care without adjusting for crossover. Thus, the RPSFT adjustment changed the incremental OS gain with sunitinib from 8 to 34 weeks and substantially reduced the CE. Based on the results from the RPSFT model, NICE recommended sunitinib for use in GIST.

**Other examples**

Other recent examples of economic evaluations based on trials with substantial crossover showed similar differences between estimated OS gains and CE when crossover was corrected for with the use of structural models. In the pivotal trial VEG105192 of pazopanib versus placebo in mRCC, 40 of 78 (51%) patients in the pazopanib arm crossed over to placebo at the time of the final analysis [24,25]. The economic evaluation undertaken for NICE, a development of the RPSFT model, which used weighted log-rank tests rather than the standard unweighted test to

---

**Table 1 – Methods for crossover adjustment and corresponding incremental cost-effectiveness ratios (ICERs) in GBP.**

<table>
<thead>
<tr>
<th>Method</th>
<th>mRCC</th>
<th>GIST</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITT (no crossover adjustment)</td>
<td>27</td>
<td>8</td>
</tr>
<tr>
<td>Cost per QALY (£)</td>
<td>72,000</td>
<td>90,500</td>
</tr>
<tr>
<td>RPSFT</td>
<td>28</td>
<td>34</td>
</tr>
<tr>
<td>Cost per QALY (£)</td>
<td>71,850</td>
<td>31,800</td>
</tr>
<tr>
<td>NICE estimate</td>
<td>&lt;50,000</td>
<td>31,800</td>
</tr>
</tbody>
</table>

GIST, gastrointestinal stromal tumor; ITT, intent-to-treat; mRCC, metastatic renal cell carcinoma; NICE, National Institute for Health and Care Excellence; QALY, quality-adjusted life-year; RPSFT, rank-preserving structural failure time.

---

Fig. 2 – Hazard ratios (with 95% confidence intervals) for overall survival with sunitinib compared with placebo in patients with gastrointestinal stromal tumor using three different statistical methods to account for crossover; in this example, there was a high rate of crossover (87%) [9]. The structural method estimates were not adjusted for poststudy treatment, and 95% confidence intervals were estimated using nonparametric bootstrapping. ITT, intent-to-treat; RPSFT, rank-preserving structural failure time.
Discussion

The presence of crossover in clinical trials can invalidate standard calculations of gains in life expectancy and incremental CE. Because ethical reasons for permitting crossover can take precedence over statistical considerations, valid and accepted methods, such as the RPSFT and IPCW models, are needed to handle the statistical issues raised by crossover and nonrandom differences in postprogression treatment.

The example of sunitinib in GIST showed that crossover and the method to adjust for this problem can have a substantial impact on estimates of incremental OS gains and on estimates of CE. The ITT method underestimates the OS benefit of treatment in the presence of crossover, and CE is thereby typically underestimated, although other factors such as the utility and costs associated with end-of-life health states, the specific model design, and other assumptions will affect the overall impact on the ICER. The variation in the HR for OS by method of analysis was more pronounced in the GIST trial than in the mRCC trial, which was expected given the greater frequency of crossover in the GIST trial.

There are few other options available to assess the impact on OS by a new therapy than to make best use of the randomized clinical trial data, even though it is limited by crossover. Extrapolating from intermediate end points such as PFS is problematic; the relationship between PFS and OS may be unknown or variable depending on the availability and efficacy of downstream therapies. Conducting additional placebo-controlled studies is usually not possible from an ethical perspective. Prospective observational studies, use of data from earlier trials [26], and retrospective database studies offer possible alternatives but have other limitations, and data may not be available at the time that reimbursement decisions are made.

There is limited experience with how reimbursement agencies view crossover; agencies have yet to issue specific methodologic guidance. In the NICE appraisal of sunitinib for GIST, the RPSFT model was accepted as appropriate, although this was also identified as an area of concern. The consistency of the HR for OS during the period before crossover (0.49) with the RPSFT estimate (0.505) strengthened the confidence in this method [22]. Decisions on pazopanib and everolimus show an increased reliance on applying the crossover methods by NICE over time. Similarly, the Swedish reimbursement agency (tandvårds-och läkemedelsförvaltningsverket) has acknowledged structural methods, although little information is available regarding its specific views on these methods [12].

Some reimbursement agencies, however, seem more skeptical. The pan-Canadian Oncology Drug Review issued a recommendation regarding pazopanib in mRCC, stating that “The Committee discussed that it would be difficult to obtain statistically significant OS results given the high rate of cross-over in the placebo arm of the trial” [27]. No reference was made to the RPSFT estimate provided in the funding application, which may indicate that the committee did not consider this as relevant evidence. In the reviews of the Australian Pharmaceutical Benefits Advisory Committee (PBAC) of sunitinib in the GIST trial, the OS estimate for the placebo group was considered highly uncertain because it represented a modeled estimate [28]. Reviewing everolimus for mRCC, the PBAC considered that the main area of uncertainty was in the estimation of OS within the study using the IPCW model, in particular regarding the reweighting of individuals in the placebo arm. The PBAC also found that estimates using IPCW and RPSFT models were confounded by design limitations of the RECORD-1 study, which permitted early and extensive crossover [29].

The choice of method for handling crossover is important and depends on several factors related to the treatments under study, the trial design, the frequency and timing of crossover (gradually, such as on patient progression, vs. all at once, such as after early unblinding), the OS time after progression, and whether assumptions behind statistical modeling methods are fulfilled.

Ignoring crossover or using simple methods such as censoring at crossover likely produces biased estimates and can be acceptable only with very low frequency of crossover, and when combined with a statistical modeling method as a sensitivity analysis. If crossover is frequent, a statistical modeling method is recommended as the base-case analysis. However, no single method completely solves the missing data problem, and there is no consensus or guidance on the choice of the method to analyze OS data from trials with crossover, although various alternatives have been used in previously published studies [10,13–15,30,31]. The models are complex and allow variability in implementation, such as the regression specification in the IPCW model and the approach taken to RPSFT modeling [32]. The performance of the methods can be tested by using Monte-Carlo simulation: the predictive performance of the RPSFT and IPCW models might, for example, be tested by randomly excluding a number of patients from both arms of a trial and seeing which results correspond most closely to true values [33].

Advantages of the RPSFT model include using the complete data set of patients in the trial and that ranking of the observed time-to-event data is preserved after adjustment. Limitations include the fact that the method does not use information on patient covariates, which may affect the probability of crossover, although this can be amended in more complex RPSFT models. Finally, the assumption that mortality decreases constantly during the time that the investigational drug is received may not reflect reality. The IPCW model, however, uses only those patients who have not been censored by crossover, which may lead to an unacceptable loss of power, as in the GIST example. The method is data-demanding if it is to give reliable results because adjustment is based entirely on observed covariates. For these reasons, the IPCW model may be most appropriate in trials with a relatively large sample size, in which only a moderate proportion of patients cross over and which have sufficient information regarding potential confounding factors. The RPSFT model would be preferable for smaller trials with relatively little information on covariates, and is also suitable for trials in which a large proportion of patients cross over, when the assumption about proportionality of the treatment effect on OS can be justified.

Table 2 provides a summary of considerations when choosing the appropriate method to adjust for crossover under different circumstances regarding degree of crossover, trial size, and availability of data on confounders. This is based largely on theoretical considerations and is supported by only limited empirical evidence; further research is needed to provide recommendations regarding the choice of analysis in different
situations. When crossover is very infrequent, methods can be expected to yield similar results. In this case, the argument can be made for retaining an ITT analysis.

Finally, the analysis method(s) to adjust for crossover should be prespecified in the statistical analysis plan, including an algorithm for method selection and inclusion of covariates, to avoid bias associated with post hoc analyses. We recommend applying several methods, with one prespecified as the main option and the others as sensitivity analyses.

Crossover is one of several problems that limit the analysis of OS in trials with PFS as the primary end point. Other issues include insufficient follow-up time leading to low power, changes in supportive care over time, and protocol-driven changes in postprogression treatment. We have presented two of the main methods to address the issue of crossover, each using available data from a trial to make predictions about the true effect on OS. We recommend that these methods be used routinely, both to report trial outcomes and as building blocks in subsequent economic evaluations based on these trials. By reducing bias and loss of power, the methods can inform pricing and reimbursement decisions and ultimately help optimize patient access to innovative therapies.

Acknowledgments

Editorial assistance was provided by Andy Gannon at ACUMED (New York, NY) with funding from Pfizer, Inc. Beata Korytowsky (formerly of Pfizer, Inc.) provided insights on crossover methodology. Suzanna Campbell and Donna Lawrence provided the review of decisions by the Australian and Canadian reimbursement agencies.

Source of financial support: This study was sponsored by Pfizer, Inc.

REFERENCES


[27] Fun-Canadian Oncology Drug Review. PCCOR Expert Review Committee (pEPC) Final Recommendation: pazopanib hydro-

Table 2 – Considerations for selecting method to analyze overall survival in the presence of crossover according to trial type and availability of data.

<table>
<thead>
<tr>
<th>Consideration</th>
<th>Crossover at random*</th>
<th>Crossover not at random</th>
</tr>
</thead>
<tbody>
<tr>
<td>Few patients cross over</td>
<td>ITT</td>
<td>IPCW</td>
</tr>
<tr>
<td>Many patients cross over</td>
<td>ITT or RPSFT</td>
<td>RPSFT</td>
</tr>
<tr>
<td>Small trial</td>
<td>ITT or RPSFT</td>
<td>IPCW</td>
</tr>
<tr>
<td>Large trial</td>
<td>ITT or RPSFT</td>
<td>RPSFT</td>
</tr>
<tr>
<td>Little information on confounding factors</td>
<td>ITT or RPSFT</td>
<td>IPCW</td>
</tr>
<tr>
<td>Abundant information on confounding factors</td>
<td>ITT or RPSFT</td>
<td>IPCW</td>
</tr>
</tbody>
</table>

* Crossover at random means that crossover is independent of patient characteristics and prognostic factors that are correlated with survival.