PharmaSUG 2026 - Paper SI - 348

Algorithms to align the distribution of follow-up across independently collected cohorts when comparing time to event endpoints using conventional Kaplan Meier and Cox regression methods

Shankar Srinivasan, Bayer Pharmaceuticals;
Yvonne Buttkewitz, Evidenze Germany GmbH;
Regina Uttenreuther, Evidenze Germany GmbH;
Baldeep C Talwar, Syneos Health;
Manjari Dissanayake, Bayer Pharmaceuticals

ABSTRACT

Using data from two independently collected cohorts, such as a historical control from a completed randomized trial and a 1:1 propensity score matched treated cohort from a prospective single arm trial, one can often have differences in the accrual time and follow-up post accrual. Such a context requires special approaches for time-to-event endpoints. We will describe an algorithm and provide SAS® code to implement the algorithm together with its rationale, focusing more on control data.

In our analyses, we matched cohorts when the enrollment of our prospective cohort was completed and before the availability of complete outcome data in the prospective cohort. A team blinded to outcome data from both cohorts conducted the matching. We then tallied events across cohorts while our trial was ongoing till the event thresholds for analyses were reached. To address differential follow-up distributions, an algorithm was used to provide equivalent follow-up in the control to that in the treated arm, both during the study and at study closure.

Conventional methods under proportional hazards could have led to biases through differential accrual and follow-up durations across the groups compared: a simplex paradox where short durations and long durations are associated with hazards of events of differing magnitudes, despite both being associated with the same ratio of hazards across groups. Our protocol pre-specified algorithm randomly paired control patients with similar follow-up distributions in a manner agnostic to censoring or event status, using potential observation times till data-cut off.

1. INTRODUCTION

The ARASEC study is a prospective single arm trial in metastatic hormone sensitive prostate cancer (mHSPC) where patients were treated with darolutamide plus an androgen deprivation therapy (ADT) of the investigator's choice. Patients on this study were compared on time-to-event endpoints and response against a control arm consisting of an ADT alone arm from the CHAARTED randomized controlled trial where the treated arm consisted of ADT with docetaxel. This external comparison was necessary given ethical concerns about the use of a concurrent ADT alone cohort as the standard of care in the United States.

Further details on the ARASEC study [6] are at ClinicalTrials.gov (ID NCT05059236). For the darolutamide product label with efficacy and safety information see [7]. CHAARTED information was used with permission from the National Cancer Institute (NCI) National Clinical Trials Network (NCTN) Data Archive (see acknowledgements) [2,3].

ARASEC was designed with similar inclusion and exclusion criteria as that in the CHAARTED. This design feature and a 1:1 matching of patients across studies facilitated a comparison on study endpoints across studies. A description of these criteria, the study design and the matching are in [4]. The external control data is presented with treated study data only as needed, without any inferential comparisons across cohorts.

Here we describe the SAS® algorithm used to move forward temporally through the external control data to provide equivalent follow-up to that in the treated cohort for the comparison of time to event outcomes across the two matched cohorts. The primary endpoint of the study was progression free survival, defined as the earliest of the occurrence of PSA progression, clinical progression or death. The

PFS endpoint and the secondary endpoint of time to castration resistance prostate cancer endpoint were analyzed using this algorithm, with the radiological progression free survival endpoint and overall survival secondary endpoints analyzed using the simpler truncation rule described in Section 1.2.

1.1 RATIONALE FOR THE USE OF THE ALGORITHM

The algorithm addresses differential follow-up across the two trials. The historical control arm from CHAARTED, has fewer patients with short durations and more with longer durations as it had a longer accrual and follow-up than the prospectively enrolled ARASEC darolutamide-treated study cohort.

In this context the algorithm helped in three ways:

- Subjects in treated/control pairs have the same potential observation time, except for some small differences from the external control due to the granular nature of that data, yielding randomization-like characteristics to follow-up across cohorts.
- It provided equivalent follow-up for event tracking while the treated cohort was ongoing and for analysis at event threshold.
- It allowed for the applicability of conventional proportional hazards-based analyses as it ruled out the possibility of bias due to differing # of patients across cohorts with shorter and longer durations.

The implementation of the algorithm is described in Section 2. Its utility, as listed above, is supported in Sections 4.1 through 4.3.

1.2 A SIMPLER TRUNCATION RULE

In addition to the pre-specified algorithm, we also used a simpler rule which uses events and censors in the CHAARTED control occurring before the largest censor or event duration in ARASEC, with an analytical administrative censoring at that duration for censors and events occurring after. In Section 4 we look at standard Kaplan Meier analysis and median follow-up for the PFS endpoint for the matched cohort using this simpler truncation rule, the raw data and the data on applying the algorithm to align follow-up.

2. METHODS

Note that the follow-up description in the methods section are reported verbatim from the study protocol and statistical analysis plan except for changes in tense, headers, edits for clarity, and the insertion of illustrative Figures 1 to 3. This makes technical descriptions close to what was implemented and consistent with the SAS code provided. The Figures provide more detail on the study enrollment and follow-up, and on the algorithm. The CHAARTED data was provided de-identified with durations provided in lieu of patient dates of enrollment and the dates of events. Our algorithm also provided a work-around for this limitation in assessing and aligning patients across cohorts on study follow-up.

The expected ARASEC study duration for the PFS endpoint was 27 months (actuals in Figure 1), assuming participants are enrolled approximately at a rate of 15 per month, and an enrollment ramp-up time of 3 months. An additional follow-up of one year was expected before the assessment of the rPFS endpoint. These durations are longer in the CHAARTED control [2,3], where the accrual period for the study extended from July 2006 to December 2012, a period of 6.5 years. The Datasphere data contains data till 23rd April 2016 on the primary endpoint, yielding more than 3 years of potential follow-up data for the last participant enrolled. Figure 1 provides a side-by-side comparison of enrollment and follow-up across the two studies. To provide a control cohort comparable to the ARASEC cohort at a cut-off point, we will obtain from the matched ARASEC cohort, the maximum potential follow-up time $\boldsymbol{f}$ for the last participant enrolled and the total accrual time $\boldsymbol{a}$. Hence, all participant follow-ups would be distributed between $\boldsymbol{f}$ and $\boldsymbol{f} + \boldsymbol{a}$.

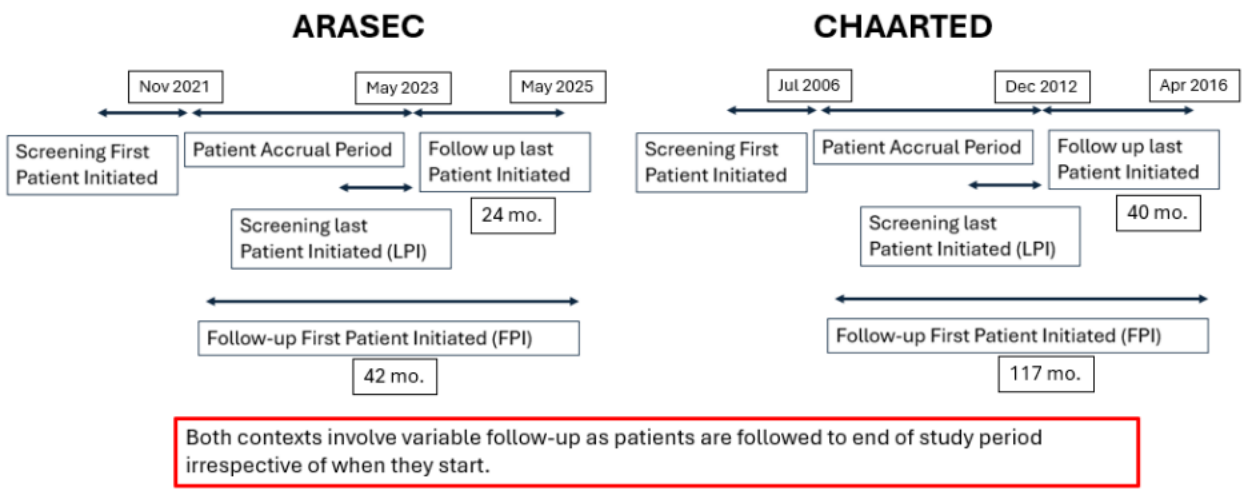

Figure 1. ARASEC vs CHAARTED enrollment and follow-up

Participant-wise truncated follow-ups was sampled from the ARASEC distribution of potential follow-ups using fixed pre-specified seeds and merged into the matched CHAARTED control data. We use the word 'potential' to denote follow-ups attainable (till data cut-off) irrespective of whether the participant was actually censored or had had an event before data cut-off. The duration of time to event or censoring in the available control data was reassessed participant-wise using truncated follow-ups. When this follow-up was less than the time to event or censoring in the control data, this duration was reduced to the truncated follow-up and the participant was deemed censored.

The analytical argument is like that in the computation of expected events rates when computing the sample size in time-to-event trials, such as that in [1]. The expected follow-up is given by

$$\int_{0}^{a} (f + a - t).dF(t) = \int_{f}^{f+a} t.dF^{*}(t)$$

Where f and a are as noted before and $F(t)$ is the distribution of accrual times. This is equivalent to the expression using $F^{*}(t)$, the distribution of potential follow-ups induced in the range from f to $f + a$ through the accrual distribution. Under uniform accrual, this evaluates to $f + 0.5 * a$, half-way through the range. Due to enrollment ramp-up time and possible variation in planned enrollment rates, this expected follow-up may differ. Sampling from the distribution $F^{*}()$ for truncations to follow-up in the matched CHAARTED control, will carve out equivalent follow-up.

Figure 2 provides the potential observation times from enrollment before truncation for the two studies that are referenced above. Following the text in Section 2.1 with this as an aide will help understand the algorithm better.

2.1 DESCRIPTION OF ALGORITHM

The process of truncation of the follow-up for the CHAARTED matched control participants for Event Tracking and Analysis was as follows:

Step 1a: For the matched ARASEC cohort post-matching, we compute the participant-wise follow-up or observation period for events A_PWFU as follows:

A_PWFU = {Calendar Date} - {Participant Enrollment date}

This is the potential follow-up at calendar date from treatment start (treatment start was mostly at enrollment date with an occasional short delay) irrespective of censoring or event. We used the current calendar date for event tracking and the data cut-off date for analysis.

Step 1b: Then we ordered A_PWFU from smallest to largest and associate it with a sequence number.

Step 2: The estimated participant wise follow-up was not provided by ECOG-ACRIN to maintain participant confidentiality. However, the year of enrollment had been provided. Steps 2 through 5 provide a censoring and event agnostic algorithm for truncating follow-up. For the CHAARTED matched participants we computed follow-up for each participant C_LPIFP as:

C_LPIFP = {Apr 23 2016 cut-off Date} - {Enrollment Date = Dec 31 20XX}

Where 'XX' was the year of enrollment provided.

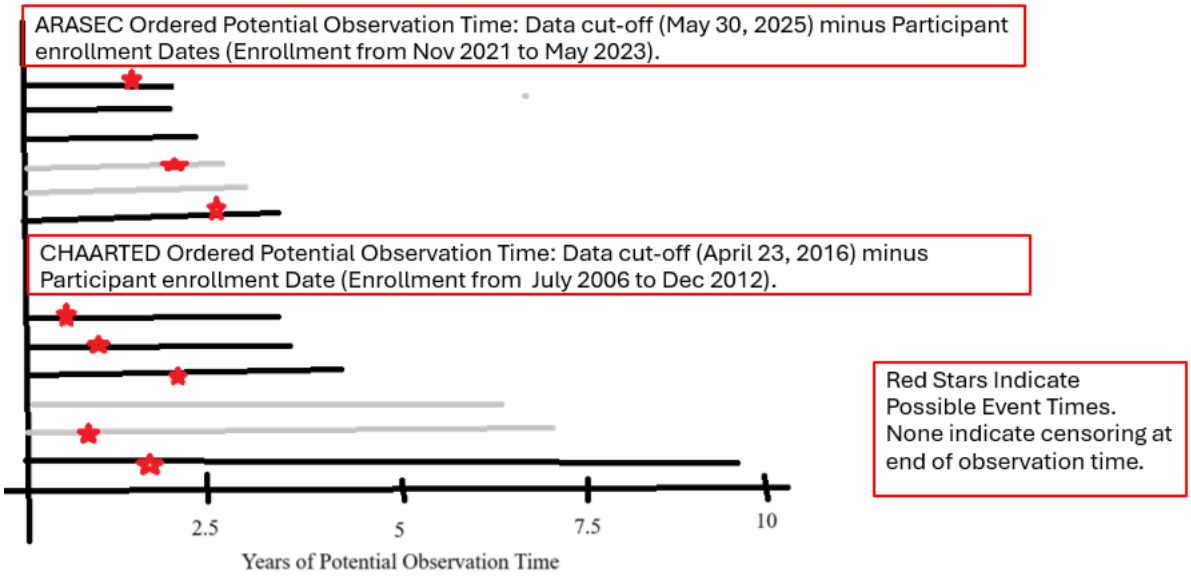

Figure 2. Illustration of Potential Observation Times across cohorts

Step 3: Then we computed the estimated minimal participant wise follow-up for the CHAARTED matched control C_EPWFU as:

C_EPWFU = Max {OS_Dur, C_LPIFP},

where OS_Dur was the max of the time to OS event or time last known alive in CHAARTED (censored due to lost to follow-up or data cut-off) for the participant. Step 4 below involves random selection of ARASEC follow-ups for CHAARTED follow-up truncation to achieve a censor/event agnostic truncation of follow-up as specified in the protocol. Figure 3 illustrates the assessment of CHAARTED based off granular enrollment year data.

Step 4: We randomly re-ordered CHAARTED participant data using a fixed seed and associated with a sequence number.

Step 5: We merged the ARASEC ordered follow-up variable into the CHAARTED data by sequence number and compute the truncated CHAARTED follow-up C_TPWFU as:

C_TPWFU = MIN {A_PWFU, C_EPWFU}

Step 6: For CHAARTED time to event data we computed truncated time to event/censoring T_Duration and the new Censor flag for event/censoring T_Censor as:

T_Duration = MIN {Duration, C_TPWFU},

If T_Duration = Duration, then T_Censor = Censor;

Else T_Censor = 1;

where Duration and Censor are the original participant time to event/censoring and the Censoring flag as reported in the CHAARTED study. (A similar step was used for the ARASEC data based on the truncated follow-up step in Step 5. Only one censored ARASEC patient was affected and was censored earlier.)

Step 7: The differing variable names across the treated and control cohorts were harmonized for duration and censor for the values obtained on truncation.

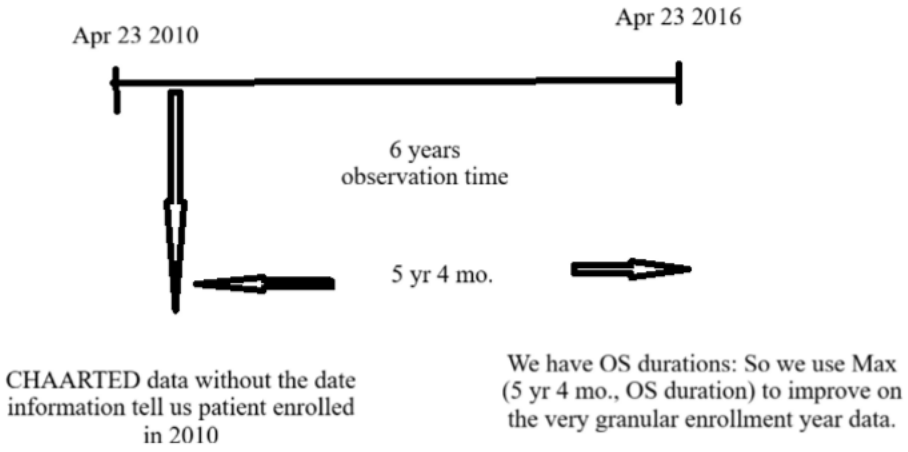

Figure 3. Illustrative example of the observation time computation using OS duration, enrollment year and data cut-off alone

For Event Tracking

We tallied Censor = 0 across the two matched datasets. Blinded event totals were reported to the blinded statistician to assist in timing of the analyses for the primary endpoint. Per protocol, the analysis was to be performed at the earlier of when at least 161 PFS events occurred or two years post enrollment end. We undershot the planned threshold at 157 events at the primary data cut-off, which occurred in May 30th 2025, two years post enrollment.

For Analysis

We compared matched cohorts using the harmonized time to event variables.

3. SAS CODE FOR IMPLEMENTATION

Program 1 shows the SAS Code that was used to compute the truncated PFS duration for ARASEC and CHAARTED.

```
*Datasets used:
*ADSL      - containing matched data
*ARA_TTE   - containing ARASEC RAW data for TTE endpoint
*CHAAR_TTE - containing CHAARTED data for TTE endpoint

*Pull in matched M-FAS (MFASFL  = "Y") ARASEC cohort (COHORTN =1) and compute observation
time for every patient in ARASEC at specified calendar time;
*(CALDT is actual ARASEC data-cutoff date for primary analysis);
DATA ARA_ADSL;
    SET ADSL;
    IF MFASFL  = "Y";
    IF COHORTN = 1;
    CALDT = mdy(5,30,2025);
    FORMAT CALDT YYMMDD10.;
    A_PWFU = CALDT - TRTSDT;
RUN;

*Sort and provide a sequence number;
PROC SORT DATA = ARA_ADSL;
    BY USUBJID;
RUN;

DATA ARA_ADSL_S;
```

```
    SET ARA_ADSL;
    SEQ = _N_;
RUN;

*Pull out CHAARTED M-FAS cohort (COHORTN =2);
DATA CHAAR_ADSL;
    SET ADSL;
    IF MFASFL = "Y";
    IF COHORTN = 2;
RUN;

PROC SORT DATA = CHAAR_ADSL;
    BY USUBJID;
RUN;

*Compute observation time for CHAARTED patients (data in months);
*(23APR2016 is the CHAARTED PFS data cut-off date. Last patient randomization year is
2012 - impute last possible randomization date - 31DEC2012);
DATA CHAAR_TTE2;
    SET CHAAR_TTE;
    IF rand_year EQ '2006/2007' THEN rand_year='2007';
    C_LPIFP = mdy(4,23,2016) - mdy(12,31,rand_year);
    C_EPWFU = max(OS*30.44,C_LPIFP);
RUN;

PROC SORT DATA= CHAAR_TTE2;
    BY USUBJID;
RUN;

*Merge CHAARTED matched data with observation time data;
DATA CHAAR_FUP;
    MERGE CHAAR_ADSL CHAAR_TTE2;
    BY USUBJID;
RUN;

*Selection agnostic to events through fixed seeds starting from 1;
DATA CHAAR_ASEL;
    SET CHAR_FUP;
    Seed = 1 + _N_;
    U = ranuni(Seed);
RUN;

PROC SORT DATA=CHAAR_ASEL;
    BY U;
RUN;

DATA CHAAR_ASEL2;
    SET CHAAR_ASEL;
    SEQ = _N_;
RUN;

*Truncate follow-up in CHAARTED to corresponding follow-up in ARASEC and create duration
to event and censor corresponding to truncated follow-up;
*Drop ARASEC USUBJID before merge as this dataset computes CHAARTED truncated TTE;
*In the CHAARTED dataset events=1 and censor=0, whereas ARASEC uses the opposite.
Note, 'abs(PFS_CRPC_event-1)' is the same as '1-PFS_CRPC_event';
DATA CHAAR_TRUN;
    MERGE ARA_ADSL_S (DROP = USUBJID) CHAAR_ASEL2;
    BY SEQ;
    T_PWFU = min (A_PWFU, C_EPWFU);
    PFS_TRUN = min (T_PWFU, PFS_CRPC*30.44);
    IF PFS_TRUN = PFS_CRPC*30.44 THEN T_Censor = abs(PFS_CRPC_event-1);
    ELSE T_Censor = 1;
    C_EPWFU_M = C_EPWFU/30.44;
    A_PWFU_M  = A_PWFU/30.44;
RUN;
```

```
*Drop CHAARTED USUBJID, Event Duration and Censor data to allow merge of truncated
follow-up alone and computations of truncated TTE for ARASEC;
DATA ARA_TRUN_;
    MERGE CHAAR_TRUN (DROP = USUBJID PFS_TRUN T_Censor) ARA_ADSL_S;
    BY SEQ;
RUN:

*Merge ARASEC TTE raw PFS data with truncated observation time T_PWFU data and create the
truncated PFS endpoint PFS_TRUN;

PROC SORT DATA=ARA_TRUN_;
    BY USUBJID;
RUN;

DATA ARA_TRUN;
    MERGE ARA_TTE ARA_TRUN_;
    BY USUBJID;
    PFS_TRUN = min (T_PWFU, AVAL);
    IF PFS_TRUN = AVAL THEN T_Censor = CNSR;
    ELSE T_Censor = 1;
RUN;

*Combine the PFS_TRUN data from the two cohorts;
DATA BOTH;
    SET ARA_TRUN CHAAR_TRUN;
RUN;
```

Program 1. Truncation algorithm

4. UTILITIES OF THE ALGORITHM

In this section, we discuss the three functions this algorithm serves that were listed earlier in Section 1.1. In randomized trials with time-to-event endpoints, analyses are associated with several features well understood by both statisticians and clinical reviewers. We will see how these are retained, using the algorithm, in the Kaplan Meier (KM) plot in Figure 4, the number of patients at risk usually displayed below these plots and reported median follow-up (in Section 4.2).

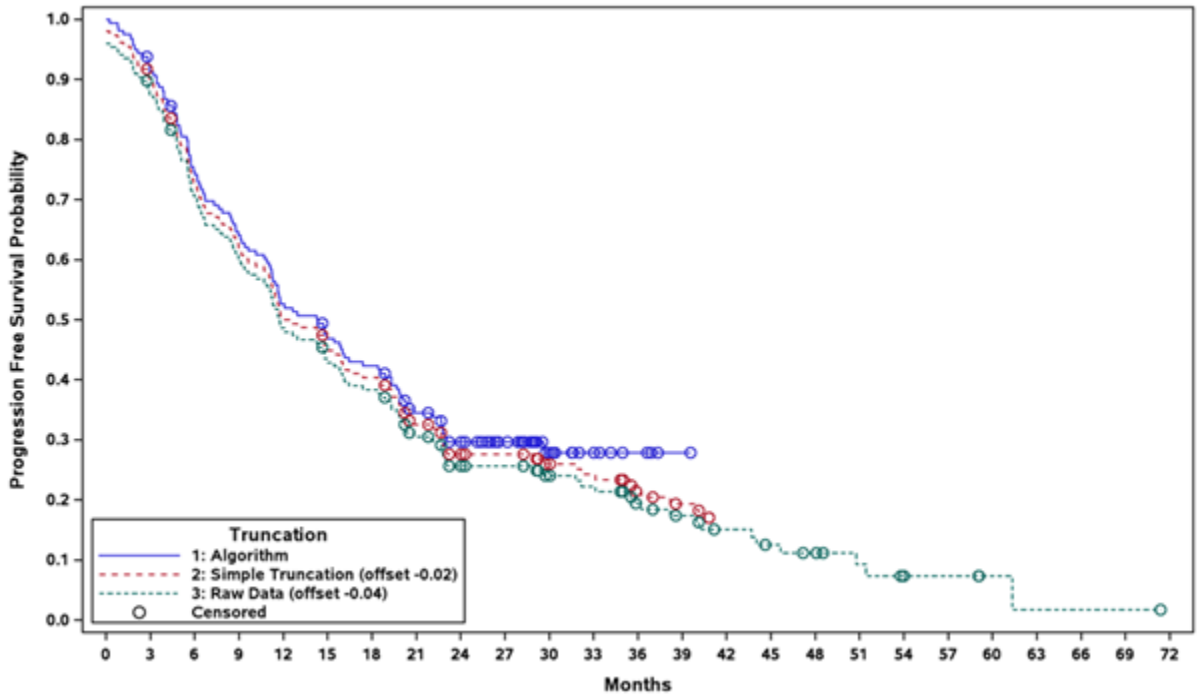

Figure 4. Offset PFS Kaplan Meier Curves for External Control by Data Handling Rule

4.1 RANDOMIZATION-LIKE CHARACTERISTICS TO FOLLOW-UP ACROSS COHORTS

The KM plot for the external control is in Figure 4, and the numbers at risk for both the control and treated cohorts are in Table 1. Both displays are by three data handling rules: the algorithm, a simple truncation as in Section 1.2 and the raw untruncated data.

Both the algorithm and the simple truncation, result in a KM graphic with an axis for duration extending to about 40 months in line with that for the treated cohort (not pictured) rather than that for the raw control data extending to about 70 months. There are more censor markings in the curve corresponding to the algorithm than the simple truncation. The algorithm introduces additional analytical administrative censoring in the control arm consistent with that in the treated arm, where patients enrolled later in the trial are administratively censored at the last assessments before data cut-off. The three offset curves in the figure validate the use of the algorithm as the estimates of progression free survival over time are unchanged using the algorithm, except for more variation through larger stairs and steps in the tail - due to fewer patients at risk, as in KMs in the randomized context.

Cohort	Rule	Months												
		1	3	…	24	27	30	33	36	39	42	…	69	72
Treated	Algorithm	160	156		94	54	33	16	6	1	0		0	0
Control	Algorithm	160	145		41	30	14	8	4	1	0		0	0
Treated	Simple Truncation	160	156		94	54	33	16	6	1	0		0	0
Control	Simple Truncation	160	145		41	39	33	30	23	19	0		0	0
Treated	Raw Data	160	156		94	54	33	16	6	1	0		0	0
Control	Raw Data	160	145		41	39	33	30	23	19	15		1	0

Table 1. Number of patients at risk by cohort and data handling rule

In Table 1, we see an anomaly for both the simple and the raw data handling approaches, where the number at risk in the control arm is larger than that in the treated arm. In a randomized study, with treatment having effect, the ongoing number at risk drops over time due to both events and censoring. While censoring tends to be similar across cohorts, the events tend to occur more frequently in the control resulting in the expectation that the number at risk for control would drop more rapidly over time for the control. The number at risk decreases in the treated arm due to administrative censoring for later enrollees with their assessment for outcome ending earlier due to the data cut-off. For the external control the data cut-off resulted in administrative censoring occurring much later resulting in a larger number of patients still at risk at comparable follow-up when using the simpler rule or the raw data (see the data in the black rectangles in Table 1). The algorithm aligns follow up and corrects this anomaly in a similar manner to that in a randomized oncology trial as in the blue rectangle.

4.2 EQUIVALENCE OF FOLLOW-UP

A second utility(advantage) of the algorithm is that it creates equivalent follow-up. We use the computations as described in [5] to characterize the follow-up for the algorithm, the simple truncation, and the raw data from the external control.

The first method is often used and provides the median of durations to both events and censoring, and the median is reduced when many events occur, such as that in a control arm. In Table 2, this is 13.6 months for all three data handling rules for the control as 50% of the patients had been censored or had had an event at this juncture. For the treated cohort, all three rules report the same median follow-up. As noted earlier, the treated cohort was minimally impacted by the truncation rules, with only one patient censored earlier. Hence, the median follow-ups are the same across data handling rules but do differ across the methods used to assess follow-up.

Methods	**Treated: All**	**External Control: Algorithm**	**External Control: Simple**	**External Control: Raw Data**
Duration to event or censoring	24.8	13.6	13.6	13.6
Censoring Duration	26.2	28.3	35.9	34.8
Median time-to-censoring	26.9	28.7	40.8	41.2

Table 2. Follow-up medians (mo.)

The second method reports the median duration for censored patients alone, and this reflects the potential follow-up that were used in Sections 2 and 3 when most of the censoring is administrative i.e. the last collected data pertinent to the endpoint prior to data cut-off. This is also the follow-up used at study planning to determine event thresholds for analysis, sample size and accrual rates. Though it could be lower than the potential follow-up if there is early loss to follow-up for other reasons. With this method, the medians for the external control are closer to that for the treated for the algorithm (26.2 vs 28.3 mo.) and much larger for the simple truncation and the raw data. The external control had its usual administrative censoring much later compared to that in the treated arm leading to this discordance in follow-up across cohorts for the simple truncation and the raw data. The simple truncation median is larger than that for raw data (35.9 and 34.8). Six patients who had an event in the raw data were now censored at month 40.8 in the simple truncation, leading to the median follow-up for the simple rule being called higher up in the ordered list of censored durations.

The third method computes the median using the Kaplan Meier approach which inverts events and censoring. Again, the median follow-up for the external control are closer to that for the treated for the algorithm (26.2 vs 28.7 mo.). For the raw data and the simple truncation, the median is larger for the control reflecting the longer study and a clustering of censors at the maximal treated follow-up respectively. The median time to censoring is 41.2 months for the raw data versus 40.8 using the simple algorithm.

4.3 APPLICABILITY OF CONVENTIONAL PROPORTIONAL HAZARDS-BASED ANALYSES

The log-rank test and the Kaplan Meier analysis, used to test and assess differences between the two treatment groups, as well as the Cox Regression providing the hazard ratios and confidence intervals, are non-parametric methods. They are often used, as they allow for the pre-specification of analyses for time-to-event endpoints, as in this treated versus external control comparison.

Parametric methods exist which use distributions such as exponential, log normal, Weibull etc. for the time-to event variables. These involve fitting and choosing an appropriate analysis in a data dependent manner which is difficult to pre-specify. Further, when even the proportional hazards assumption fails, reported hazard ratios are interpretable as an average HR over follow-up and the non-parametric methods continue to be seen as relevant when there is a strong observed effect.

These non-parametric methods, however, do entail a proportional hazards assumption to be strictly valid [see 8 for instance]. A hazard is a probability of an event in a small interval, such as a month, and can vary over the duration of follow-up. With the proportional assumption, the ratio of the

hazards of an event across groups should ideally be about the same over time, without requiring that the hazards in each group themselves be constant. An assessment of proportional hazards can be conducted by visual inspection (see Allison PD, 2010) using a plot of the negative natural log of event free survival (S) versus the duration (t) to event.

Under exponential assumptions, the survival at time t

$S(t) = \mathrm{Exp}(-\mathrm{Hazard} * t)$ and implies $-\mathrm{LN}(S(t)) = \mathrm{Hazard} * t$.

Hence, the expectation in this -LN survival vs. duration plot is two lines like spokes extending from bottom left to top right with slopes being the hazards for events in the two groups. We would have two straight lines if we had constant hazards over time in each group or non-straight but proportional slopes under the less restrictive proportional hazards assumption.

This is plotted in Figure 5 for the external control with simple truncation. It reflects the possibly of two hazard pieces, one for durations shorter than 23 months and another for longer durations with reduced hazards. The line for the treated cohort locates as a spoke below that for the external control (not displayed pending clinical publication) and provides good support to the proportional hazard assumption with a similar reduction in the slope for the longer duration.

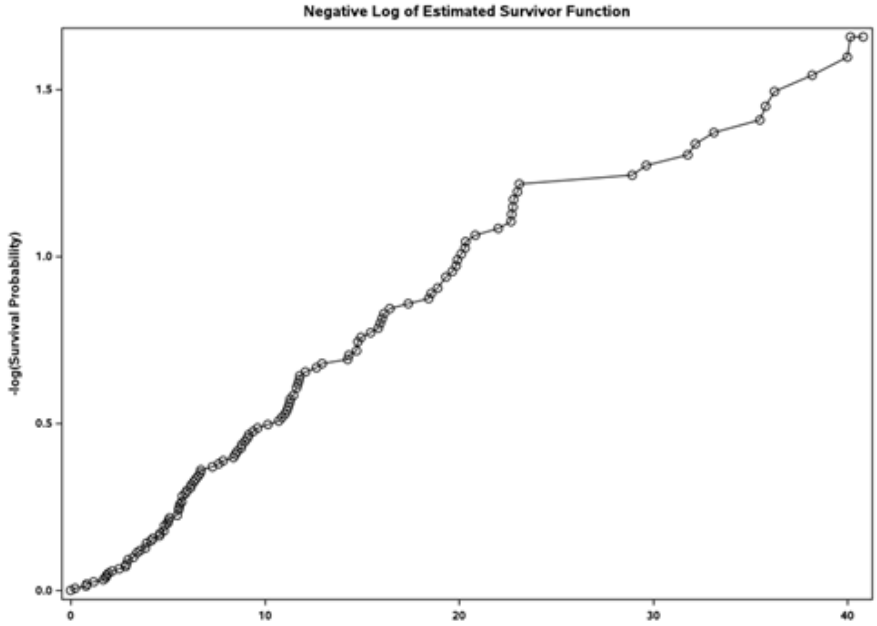

Figure 5. Negative LN Survival by Duration (mo.) Plot for the External Control

In this proportional but differing hazard context, leads to the third utility of the algorithm - removal of any bias due to unbalanced patient representation in hazard pieces across follow-up, such as long versus short durations, by aligning the distribution of follow-up across cohorts. To see why that could have been an issue see the hypothetical example in Table 3. In this hypothetical example, the control has fewer patients with short durations and more with long durations (a hypothetical 60 and 100 vs 100 and 60). This is like our raw data for the external control vs the treated study cohort. In both the short durations and the long durations, the ratio of hazards (HR) is both 0.5 meeting the proportional hazard assumption of standard time to event analyses, with the long durations having lower hazards of events per unit period in the control of 0.02 or 2% vs 4% respectively, with a 50% reduction for the treated cohort.

One would expect a net HR of 0.5 for the overall data but the combined HR would be biased upwards at 0.59 due to the differences on short and long durations. Our algorithm removed the possibility of such biases, despite proportional hazards, by analytically pairing patients across propensity score matched cohorts on potential follow-up. A feature of randomization. Only when there is proportional as well as constant hazards within groups over time would such biases not be an issue in the simple truncation context.

Hazard by follow-up post index	Control	Treated	Hazard Ratio
Short Duration Interval			
Hazard Short	0.04	0.02	0.5
N Short Durations	60	100	
Long Duration Interval			
Hazard Long	0.02	0.01	0.5
N Long Durations	100	60	
Net hazard	0.0275	0.01625	0.59

Table 3. Simplex Paradox when groups have differing follow-up durations

5. CONCLUSION

We used an algorithm based on potential follow-up (patient durations from treatment start till respective data-cutoffs) agnostic to event or censoring to align follow-up distributions across the compared cohorts. Patients were paired off across matched cohorts in a random manner, akin to a randomization, and then a truncation of follow-up was applied based on the shorter duration in the pair. This provides an additional tool to facilitate comparisons to external controls to those such as matching and inverse probability weighting, by borrowing this other strong feature of randomized clinical trials in oncology.

We used an algorithm to align follow-up across the cohorts for the PFS endpoint and compared this to the use of a simple truncation of follow-up at the maximal follow-up in the treated cohort. With the latter rule, the treated patients are censored administratively at the last endpoint data collected before data cut-off with the # at risk dropping off consequently. For the external control that administrative censoring occurred much later resulting in an anomaly of a larger number of patients who are still at risk at comparable follow-up when using the simpler rule. This anomaly of having more at risk in a control arm is corrected using an additional analytical administrative censoring through the algorithm.

The algorithm allows for the tracking of event counts as in randomized time-to-event trials by providing an algorithm to move forward temporally through a historical control in parallel to follow-up in the treated arm. The equivalence of observation time induced by the algorithm allows for the conventional reporting of median follow-up times in contexts where the cohorts have differing study lengths and enrollment durations.

Finally, having an aligned follow-up provides additional robustness of conventional time-to-event analysis tools under proportional hazards.

Limitations of this algorithm include applicability to contexts where the distribution of potential follow-up for two cohorts have overlap even when one tends to have patients with longer follow-up. In such a context, the random pairing followed by truncation would reduce the rate of new events tallied across cohorts with increasing data maturity, over the follow-up aligned treated and control arms. Long waits for event thresholds in randomized time-to-event trials due to reduced patients at risk over time is already an issue.

This and some added complexity in implementation can be justifiable depending on clinical utility, the availability of statistical analysis planning time and resources and where randomized controlled trials are not feasible.

REFERENCES

1. Collett D. 2003. Modelling Survival Data in Medical Research. Second Edition. New York, New York: Chapman & Hall/CRC.

2. Sweeney CJ, Chen YH, Carducci M, Liu G, Jarrard DF, Eisenberger M, et al. 2015. Chemohormonal Therapy in Metastatic Hormone-Sensitive Prostate Cancer. N Engl J Med. Aug 20;373(8):737-46.
3. Kyriakopoulos CE et al. 2018. Chemohormonal Therapy in Metastatic Hormone-Sensitive Prostate Cancer: Long-Term Survival Analysis of the Randomized Phase III E3805 CHAARTED Trial. Journal of clinical oncology. April;36(11): 1080-87.
4. McKay RR, Ross AE, Preston MA, et al. 2025. Darolutamide plus androgen-deprivation therapy: Propensity score matching of ARASEC and historic clinical trial patients. Future Oncol.;21(11):1365–1375. doi:10.1080/14796694.2025.2482360.
5. Betensky RA. 2015. Measures of follow-up in time-to-event studies: Why provide them and what should they be? Clin Trials. 12(4):403-408. doi:10.1177/1740774515586176.
6. Open-label study of Androgen Receptor Inhibition with dArolutamide plus Androgen Deprivation Therapy (ADT) Versus ADT in men with Metastatic Hormone-Sensitive Prostate Cancer Using an External Control Arm (ARASEC) ClinicalTrials.gov ID NCT05059236. https://clinicaltrials.gov/study/NCT05059236.
7. Nubeqa® (darolutamide) product label. https://www.accessdata.fda.gov
8. Hosmer D, Lemeshow, S and May S. 2008. Applied Survival Analysis: Regression Modelling of Time-to-Event Data. Second Edition. Hoboken, NJ: John Wiley and Sons, Inc.
9. Allison PD. 2010. Survival Analysis Using SAS®: A Practical Guide, Second Edition. Cary, NC: SAS Institute Inc.

ACKNOWLEDGMENTS

The authors are thankful to patients and investigators of the ARASEC study, the study steering committee and the Bayer and Parexel clinical and operations team. Additional support in implementation was provided for biostatistics by Bernd Niebank, MSc equivalent, of Parexel International GmbH, and for statistical programming support by Jana Wree, MSc equivalent, and Firdevs Topcu-Alici, MSc, both of Parexel International GmbH.

This manuscript was prepared using data from Datasets NCT00309985-D3, -D5, -D7, -D8, -D10, and -D11 from the NCTN Data Archive of the National Cancer Institute's (NCI's) National Clinical Trials Network (NCTN). Data were originally collected from clinical trial NCT number E3805 (CHAARTED). All analyses and conclusions in this manuscript are the sole responsibility of the authors and do not necessarily reflect the opinions or views of the CHAARTED clinical trial investigators, the NCTN, or the NCI.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the first or last author at:

Shankar S Srinivasan Bayer Pharmaceuticals, Whippany, NJ Shankar.srinivasan@bayer.com	Manjari Dissanayake Bayer Pharmaceuticals, Whippany, NJ Manjari.dissanayake@bayer.com