Paper AS18

From Data Cutoffs to Decision-Making: Programming Futility Analysis in Oncology Clinical Trials

Mani Mudumba, Merck & Co., Inc., Rahway, NJ, USA

ABSTRACT

Interim analyses are increasingly incorporated into oncology clinical trials to improve operational efficiency and support timely decision-making. This paper presents a statistical programming perspective on implementing a randomized Phase 3 study with run in portion and a pre-specified interim futility analysis. We detail key programming strategies for managing interim data cutoffs, deriving futility analysis populations, and supporting both safety and efficacy outputs, particularly when progression-free survival (PFS) is the primary endpoint. The paper explains how the outputs facilitate timely, data-driven decisions by Independent Data Monitoring Committees (IDMCs). Finally, I discuss the programming challenges specific to implementing futility analysis and provide practical solutions to ensure accurate, reliable interim results for decision-making.

INTRODUCTION

Interim analyses are increasingly embedded into oncology clinical trials to support faster decision-making and more efficient drug development. When trials are designed with predefined futility rules, they provide an opportunity to stop early if the investigational treatment is unlikely to demonstrate clinical benefit. This conserves patient resources, reduces operational burden, and redirects development efforts more efficiently.

This paper focuses on the programming implementation of futility analysis in a randomized Phase 3 oncology study that includes an initial Run-In portion, followed by a Phase 3 portion, and a planned interim futility analysis. The main objective is to demonstrate statistical programming strategies for:

- Handling interim data cutoffs
- Deriving populations for safety and efficacy
- Generating PFS-based efficacy outputs
- Enabling timely review by Independent Data Monitoring Committees (IDMC)

STUDY DESIGN

Many oncology studies use a Run-In portion to evaluate the safety, tolerability and early efficacy of the study medication, the data from the run-in portion are typically not included in the primary analysis of the Phase 3 portion. The Phase 3 portion evaluates the treatment's clinical benefit with adequate statistical power. A futility analysis is introduced at a predefined timepoint, driven either by number of events or calendar-based cutoff, to assess whether continuing the trial remains appropriate.

Run-In → Randomized Phase 3 → Interim Futility

RUN-IN PORTION

ADVANTAGES

- Satisfies regulatory requirements
- Enables earlier assessment of safety, tolerability, and efficacy prior to the Phase 3 data readout

CHALLENGES

- Additional operational complexity

- Separate population determination logic
- Additional datasets, flags, and derivations

From a programming perspective, the Run-In phase introduces population transitions, which must be carefully tracked to define who qualifies for randomized analysis.

FUTILITY ANALYSIS

Futility analysis assesses whether continuing the study has a reasonable chance of demonstrating benefit. Stopping early for futility:

- Protects patients from ineffective treatment
- Reduces unnecessary cost and time
- Enables pivot to alternate therapies or trials

DATA CUTOFF

To support interim decision-making, ongoing event accrual is monitored to anticipate when the required number of events will be reached. An event-accrual forecast is generated using currently observed follow-up and events to estimate the expected future event rate. This forecast is then used to identify an operational cutoff date that is most likely to achieve the targeted number of events specified in the interim analysis plan. Because event accrual is influenced by enrollment patterns, varying follow-up times, and changes in risk over time, this estimate is treated as an approximation and is routinely updated as new data become available.

In this paper, the futility decision is based on Progression-Free Survival (PFS), evaluated at a pre-specified interim cutoff date determined from event monitoring. A critical programming responsibility is ensuring that all relevant data are available at the cutoff, as this defines the futility analysis database lock.

POPULATION FLAGS

Population flags were derived using date-based conditions to ensure that only subjects meeting the randomization and data cutoff criteria were included in the futility analysis. Below are the key population flags:

Flag	Label	Derivation	Purpose
RANDFL	Randomized Population Flag	If Date of Randomization is not missing, then RANDFL="Y", otherwise RANDFL="N"	Identifies subjects who were randomized
SAFFL	Safety Population Flag	If Date of First Exposure to Treatment is not missing, then SAFFL="Y", otherwise SAFFL="N".	Identifies subjects included in safety analyses
FUTANAFL	Futility Analysis Population Flag	FUTANAFL = “Y” for randomized subjects included in the prespecified futility analysis population, defined in the SAP as the first N randomized subjects (where N ensures adequate minimum follow-up for post-baseline assessments).	Identifies subjects eligible for the futility analysis at the interim cutoff

EFFICACY OUTPUTS

The primary objective of the interim analysis is to assess whether Experimental Treatment demonstrates adequate evidence of clinical benefit over Control to justify continuation into the confirmatory phase. In this paper, Progression-Free Survival (PFS) is the primary endpoint used for the futility decision.

PROGRESSION-FREE SURVIVAL (PFS)

PFS was derived as the time from randomization to first documented disease progression or death, whichever occurred first. Subjects without an event by the data cutoff date were censored at their last valid tumor assessment date on or before cutoff.

The following are the key Variables required for PFS calculations in ADTTE prior to the interim cutoff:

PDDT	Date of tumor progression	Used for the event indicator and PFS Duration
DTHDT	Date of Death	Used for event indicator if death occurs before progression
TUMASDT	Last Adequate Tumor Assessment Date	Used for censoring
RANDDT	Date of Randomization	Used for Time zero
DCUTDT	Data Cutoff Date	Used for data cutoff

DERIVATION

```
/* Derive PFS time and event flag */
if (pddt <= dcutdt and pddt ne .) or
   (dthdt <= dcutdt and dthdt ne . ) then do;
    /* Subjects with PD or death before cutoff */
    aval = (min(pddt, dthdt) - randdt + 1) / 30.4375;
    cnsr = 0;
end;
else do;
    /* Censored at last adequate tumor assessment before cutoff */
    aval = (min(tumasdt, dcutdt) - randdt + 1) / 30.4375;
    cnsr = 1;
end;
```

This paper explains the development of statistical programming deliverables, which include an PFS Summary table and KM Plot outputs, as well as our lessons learned. An illustrative scenario based on a SAS dummy ADTTE dataset with each row represents one subject's PFS observation will be used throughout the paper to help with understanding the concept. The snippet of the dummy ADTTE is shown below:

USUBJID	TRT01P	TRT01PN	PARAM	PARAMCD	PARAMN	STARTDT	ADT	AVAL	CNSR
Study_100000003	Control	2	Progression_Free_Survival	PFS	1	10JAN2022	17MAR2022	66	1
Study_100000004	Experimental	1	Progression_Free_Survival	PFS	1	20NOV2021	17APR2022	148	0
Study_100000002	Experimental	1	Progression_Free_Survival	PFS	1	05JUL2021	14DEC2021	162	1
Study_100000005	Control	2	Progression_Free_Survival	PFS	1	15MAY2021	14DEC2021	213	0
Study_100000006	Experimental	1	Progression_Free_Survival	PFS	1	01FEB2022	21OCT2022	263	1
Study_100000001	Control	2	Progression_Free_Survival	PFS	1	12MAR2021	21JUN2022	456	0

The Cox proportional hazards model was fitted using PROC PHREG to estimate the hazard ratio and corresponding p-value. Kaplan–Meier survival curves were generated using PROC LIFETEST to visualize time-to-event outcomes. Sample code is provided below:

```
/*PHREG to Get HR and P-Value*/

ods output ParameterEstimates = hr_table;
proc phreg data = adtte alpha=0.05;
    where trt01pn in (1,2);
    class trt01pn(ref='2');
    model aval*CNSR(0) = trt01pn/ rl  TIES=EFRON;
    ods output ParameterEstimates=stat1;
run;
ods output close;
```

```
/*KM Plot*/
ods graphics on;
proc lifetest  data= adtte method=KM alpha=0.05 timelist=6 outsurv=stat
reduceout;
      time aval*CNSR(1);
      strata trt01pn/order=internal;
      format trt01pn trtfmt.;
      label trt01pn=Planned Treatment for Period 01;
      ods output quartiles=medpfs;
run;
ods graphics off;
```

PFS SUMMARY TABLE

The actual rule depends on the trial design but for the interpretation purpose we are using the HR Futility Boundary as 1. Futility conclusion based on observed HR is shown below:

Observed HR	Conclusion
HR < 1.0 (Futility Boundary) favoring Experimental Treatment	Positive trend → continue
HR ≥ Futility boundary	Insufficient benefit, Consider early stop

Futility Boundary: Stop if HR ≥ 1.0

Table 1
PFS Summary Table
(HR < 1.0, Continue Study)

Stats	Experimental	Control
N	283	282
Events (%)	133 (46.9)	200 (70.9)
Median PFS (Months)	10.5	6.7
Hazard Ratio	0.527 (0.42, 0.65)	
p-value	0.0010	

Table 2
PFS Summary Table
(HR > 1.0, Consider Early Stop)

Stats	Experimental	Control
N	280	281
Events (%)	169 (60.3)	156 (55.5)
Median PFS (Months)	5.79	6.58
Hazard Ratio	1.127 (0.90, 1.40)	
p-value	0.2805	

PFS summary tables illustrate the treatment effect (HR < 1.0 or HR > 1.0), indicating continuation or discontinuation of the study according to the futility boundary rules

To gain a deeper understanding of the analysis results, we generated the PFS summary tables above illustrate two contrasting scenarios in the context of futility analysis. Table 1 shows a scenario where the HR is below the futility boundary (HR < 1.0), favoring the experimental treatment. Here, the experimental arm demonstrates superior efficacy compared to the control arm, with fewer events (133/283, 46.9% vs. 200/282, 70.9%) and a longer median PFS (10.5 months vs. 6.7 months). The hazard ratio of 0.527 (95% CI: 0.42, 0.65) and a significant p-value of 0.0010 indicate a 47% reduction in the risk of progression or death, supporting continuation of the study.

In contrast, Table 2 represents a scenario where the HR exceeds the futility boundary (HR > 1.0), suggesting insufficient benefit of the experimental treatment. In this case, the experimental arm has a higher proportion of events (169/280, 60.3%) compared to the control arm (156/281, 55.5%) and a shorter median PFS (5.79 months vs. 6.58 months). The HR of 1.127 (95% CI: 0.90, 1.40) indicates a 13% higher risk of progression in the experimental arm, and the non-significant p-value of 0.2805 demonstrates that the observed difference is not statistically meaningful. This pattern of results, where the experimental treatment shows weak or no efficacy relative to control, would support consideration of early termination of the study for futility.

Together, these tables demonstrate how the futility boundary and hazard ratio can guide interim decisions in a clinical trial, ensuring timely assessment of experimental treatment efficacy while minimizing patient exposure to ineffective therapy.

KAPLAN–MEIER ANALYSIS OF PFS

Figure 1: Kaplan–Meier Analysis of PFS (Experimental Treatment Favorable) (Study to Continue)

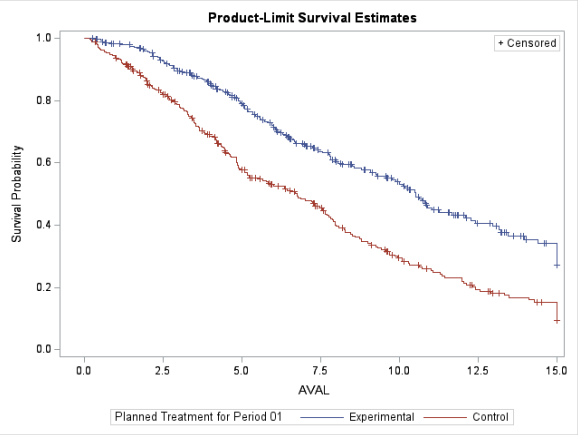
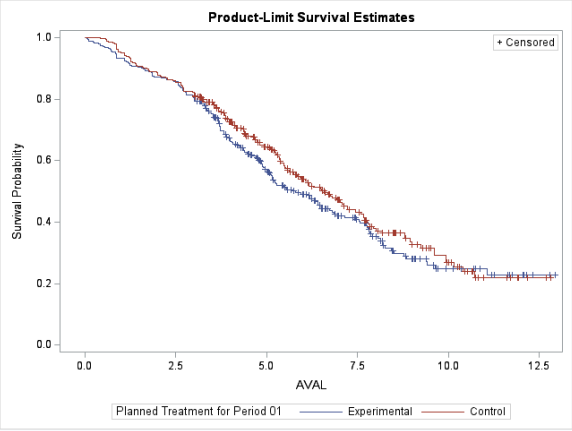


Figure 2: Kaplan–Meier Analysis of PFS (Experimental Treatment Not Favorable) (Consider Early Stop)



Kaplan–Meier plots show progression-free survival for the experimental and control treatments, supporting the IDMC's decision to continue or stop the study

To gain a deeper understanding of the analysis results, we also generated The Kaplan–Meier plots which illustrate progression-free survival (PFS) for the experimental and control arms under two scenarios. In Figure 1, the experimental arm curve remains consistently above the control arm throughout the study period, showing clear separation between the treatments. This indicates a strong efficacy signal for the experimental therapy, supporting continuation of the study.

In Figure 2, the curves for the experimental and control arms are much closer and overlap at multiple time points. At certain intervals, the control curve is slightly above the experimental curve, and the lines move closer over time. This pattern indicates weak efficacy and a lack of consistent benefit for the experimental treatment. Such overlapping and occasionally crossing curves suggest that the study may be considered for early termination due to futility, as the experimental therapy does not demonstrate a clear advantage over control.

SAFETY OUTPUTS

In addition to the efficacy assessments, safety datasets were also simulated to reflect standard programming practices for interim analyses. The simulated safety summaries showed an overall balanced distribution of adverse events between treatment arms, supporting the expectation that no major safety concerns would influence the interim decision. However, as the primary focus of this paper is on the implementation of futility analysis using PFS, detailed safety results are not discussed here. Instead, safety analyses are acknowledged to illustrate how they are typically included alongside efficacy outputs to provide a complete safety and efficacy profile for IDMC review, while avoiding disclosure of any real clinical trial data.

INDEPENDENT DATA MONITORING COMMITTEE (IDMC) INTERPRETATION

The PFS summary tables and Kaplan–Meier plots provide critical input for IDMC decision-making regarding interim futility. In scenarios where the experimental treatment demonstrates a consistent and meaningful benefit over control, with clear separation of survival curves, the IDMC would recommend continuation of the study. Conversely, when the curves overlap heavily, cross frequently, or the experimental arm shows weak or inconsistent efficacy, these patterns signal insufficient treatment benefit, and the IDMC may consider early termination for futility. This approach ensures that decisions are guided by both the magnitude and consistency of treatment effect, balancing patient safety and efficient use of resources.

LESSONS LEARNED / DISCUSSION

Implementing interim futility analysis in oncology trials requires careful coordination of programming, data management, and statistical methodology. Key lessons include:

- **Accurate population derivation:** Proper use of randomization, safety, and futility flags ensures that only eligible subjects are included in interim analyses. Misclassification can significantly affect HR estimates and IDMC decisions.
- **Data cutoff management:** Maintaining correct data availability at the interim cutoff is essential to avoid bias in efficacy outputs and to lock the futility analysis dataset reliably.
- **Time-to-event derivations:** Correct handling of censoring, progression, and death dates ensures accurate PFS calculations and hazard ratio estimates.
- **Visualization and interpretation:** PFS Summary tables and Kaplan–Meier plots are invaluable for IDMC review. Clear separation of survival curves supports continuation, while overlapping or crossing curves highlight weak efficacy, signaling potential early stop.

These lessons emphasize the importance of combining precise programming, rigorous quality checks, and informative visualization to facilitate robust, data-driven interim decisions.

CONCLUSION

Futility analysis provides a structured framework to assess early whether an experimental treatment is unlikely to demonstrate clinical benefit. By integrating precise SAS programming, accurate population derivations, and clear visualization of PFS outcomes, interim analyses can support IDMCs in making timely, evidence-based decisions. These strategies enhance operational efficiency, safeguard patient welfare, and optimize the overall drug development process.

ACKNOWLEDGMENTS

The authors would like to thank Amy Gillespie, Naveen Kommuru, Sai Govind Chenna, Deepak Grewal, and Xiang Peng for reviewing the paper and providing valuable suggestions.

REFERENCES

1. Jitlal M, Khan I, Lee SM, Hackshaw A. Stopping clinical trials early for futility: retrospective analysis of several randomised clinical studies. Trials. 2011;12(Suppl 1):A53. https://trialsjournal.biomedcentral.com/articles/10.1186/1745-6215-12-S1-A53
2. Mehta HB, Mehta V, Girling A, et al. Interim futility analysis with intermediate endpoints. J Clin Oncol. 2008;26(17):?. https://pubmed.ncbi.nlm.nih.gov/18283075
3. Herson J, Dodd L, Schuetz M, et al. Statistical controversies in clinical research: futility analyses in oncology – lessons on potential pitfalls from a randomized controlled trial. Ann Oncol. 2017;28(7):1419-1426. https://www.sciencedirect.com/science/article/pii/S0923753419322446
4. Meurer WJ, Tolles J. Guidance on interim analysis methods in clinical trials. J Clin Transl Sci. 2023;7(1):?. https://www.cambridge.org/core/journals/journal-of-clinical-and-translational-science/article/guidance-on-interim-analysis-methods-in-clinical-trials/5051FDCF5284970B3DB01FE609AAA4C2

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Mani Mudumba

Merck & Co., Inc., Rahway, NJ, USA

Email: mani.mudumba@merck.com