Approaches to Assessment of Overall Survival in Oncology Clinical Trials

Guidance for Industry

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 60 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to https://www.regulations.gov. Submit written comments to the Dockets Management Staff (HFA-305), Food and Drug Administration, 5630 Fishers Lane, Rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document, contact (OCE and CDER) Nicole Gormley at OCE-Guidances@fda.hhs.gov or (CBER) Office of Communication, Outreach, and Development, 800-835-4709 or 240-402-8010.

U.S. Department of Health and Human Services
Food and Drug Administration
Oncology Center of Excellence (OCE)
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

August 2025
Clinical/Medical

Approaches to Assessment of Overall Survival in Oncology Clinical Trials

Guidance for Industry

Additional copies are available from:

Office of Communications, Division of Drug Information
Center for Drug Evaluation and Research
Food and Drug Administration
10001 New Hampshire Ave., Hillandale Bldg., 4th Floor
Silver Spring, MD 20993-0002
Phone: 855-543-3784 or 301-796-3400; Fax: 301-431-6353
Email: druginfo@fda.hhs.gov
https://www.fda.gov/drugs/guidance-compliance-regulatory-information/guidances-drugs

and/or

Office of Communication, Outreach and Development
Center for Biologics Evaluation and Research
Food and Drug Administration
Phone: 800-835-4709 or 240-402-8010
Email: industry.biologics@fda.hhs.gov
https://www.fda.gov/vaccines-blood-biologics/guidance-compliance-regulatory-information-biologics/biologics-guidances

U.S. Department of Health and Human Services
Food and Drug Administration
Oncology Center of Excellence (OCE)
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

August 2025
Clinical/Medical

TABLE OF CONTENTS

Approaches to Assessment of Overall Survival in Oncology Clinical Trials
Guidance for Industry[1]

This draft guidance, when finalized, will represent the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA staff responsible for this guidance as listed on the title page.

I. INTRODUCTION

The purpose of this guidance is to provide recommendations to sponsors on the assessment of overall survival in randomized oncology clinical trials conducted to support marketing approval of drugs and biological products, with an emphasis on the analysis of overall survival as a pre-specified safety endpoint.[2] The recommendations in this guidance are based in part on discussions held at the joint FDA, American Association for Cancer Research (AACR), and American Statistical Association (ASA) public workshop: Overall Survival in Oncology Clinical Trials held in July 2023.[3]

Additional information on the related topic of choice of endpoints in oncology clinical trials can be found in guidance for industry *Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics* (December 2018).[4] Pre-specification of a statistical analysis plan (SAP) is a basic statistical principle for clinical trials; refer to guidance for industry *E9 Statistical Principles for Clinical Trials* (September 1998). The estimand framework is useful to describe the plans for evaluation of overall survival and other primary and key secondary endpoints; refer to guidance for industry *E9(R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials* (May 2021).

[1] This guidance has been prepared by the Oncology Center of Excellence (OCE), Center for Drug Evaluation and Research (CDER), and Center of Biologics Evaluation and Research (CBER) at the Food and Drug Administration.

[2] For the purposes of this guidance, references to *drug* or *drugs* include both human drug products and biological products regulated by CDER and CBER, unless otherwise specified.

[3] FDA-AACR-ASA Workshop: Overall Survival in Oncology Clinical Trials. 2023 Jul 18: https://www.aacr.org/professionals/policy-and-advocacy/regulatory-science-and-policy/events/fda-aacr-asa-workshop-overall-survival-in-oncology-clinical-trials/.

[4] We update guidances periodically. For the most recent version of a guidance, check the FDA guidance web page at https://www.fda.gov/regulatory-information/search-fda-guidance-documents.

While the guidance discusses situations in which it is appropriate to consider overall survival for the primary endpoint, this guidance primarily focuses on statistical or design considerations when overall survival is not the primary endpoint. Additionally, this guidance focuses on the assessment of overall survival in randomized trials. In general, time-to-event endpoints (e.g., overall survival, progression-free survival, etc.) are challenging to interpret in single-arm trials due to the lack of a randomized control arm to demonstrate that the experimental treatment results in an improvement in the time-to-event endpoint.

In general, FDA's guidance documents do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidances means that something is suggested or recommended, but not required.

II. BACKGROUND

Overall survival is both an efficacy and a safety endpoint; it can be favorably impacted by the therapeutic benefits of a specific drug and negatively impacted by the drug's toxicity. Overall survival is also an objective, clinically meaningful endpoint that can be measured easily and precisely. It is considered a gold standard endpoint in oncology, as prolongation of life in the setting of a life-threatening disease is of clear inherent value.

Overall survival should be prioritized as the primary endpoint when feasible. It is most appropriate for consideration in randomized trials of oncologic diseases with a short natural history (e.g., metastatic pancreatic cancer), late-line disease settings, or in disease settings where there are other therapeutics known to prolong overall survival and it is important to demonstrate retention or improvement of this survival advantage.

In certain oncology settings, it is often not feasible or practical to include overall survival as the primary endpoint. In some indolent diseases or those with extremely efficacious therapeutics that result in long survival times, the follow-up time needed to show superiority of overall survival as the primary endpoint may be impractical. Additionally, as mentioned, overall survival may not be appropriate in single-arm trials. Furthermore, although overall survival is an objective measurement, the results and their interpretation can be impacted by crossover, receipt of subsequent therapy, and other intercurrent events. These additional factors that impact the interpretation of overall survival are discussed in Sections III.A and III.B.

When overall survival is not the primary endpoint and other endpoints such as response rate, progression-free survival, or event-free survival are the primary endpoint in an oncology randomized trial, FDA recommends collection and submission of overall survival data. In some situations, overall survival may be a secondary endpoint included in a pre-specified hypothesis testing plan to permit a formal statistical test for an efficacy evaluation, which includes control of

the overall study-wise Type I error rate.[5,6] When it is not a secondary endpoint with an appropriate formal testing plan, FDA has historically evaluated the overall survival results in a descriptive manner, assessing overall survival as a safety measurement. There have been several drug development programs in which discordant results have been observed between early efficacy endpoints (e.g., response rate, minimal residual disease, or progression-free survival) and overall survival.[7] In some of these instances, the drug demonstrated anti-tumor activity, but the benefits of the anti-tumor activity were outweighed by drug-related toxicity that was evidenced by the observed decrements in overall survival. These examples underscore the importance of evaluation of overall survival as a safety assessment.

There are, however, inherent limitations in the assessment of overall survival in a descriptive manner, even if done as a safety assessment. Without a formal pre-specified plan for evaluation, the interpretation of the overall survival results can be challenging, especially if the study was not designed to collect sufficient overall survival data to provide estimates of treatment effect with a pre-specified level of precision.

III. RECOMMENDATIONS

When overall survival is not an alpha-controlled primary or secondary efficacy endpoint with a formal testing plan in a randomized oncology clinical trial, sponsors should include a pre-specified plan to assess overall survival, as a safety endpoint, with an aim to evaluate for potential harm due to the therapeutic intervention. Sponsors should consult with the FDA review division on their plans to include overall survival analyses in the clinical protocol and SAP.

A. Trial Design Considerations

This section includes an overview of clinical trial design considerations as it pertains to the assessment of overall survival in oncology clinical trials.

- Overall survival is generally defined as the time from randomization until death from any cause.

- All randomized oncology clinical trials should be designed to assess overall survival in order to adequately evaluate the potential for harm (see also Section III.B.2).

[5] Control of the Type I error rate is sometimes referred to as “alpha control”.

[6] See the guidance for industry *Multiple Endpoints in Clinical Trials* (October 2022) for additional information regarding analysis and interpretation of multiple endpoints.

[7] Merino M, Kasamon Y, Theoret M, Pazdur R, Kluetz P, Gormley N, Irreconcilable Differences: The Divorce Between Response Rates, Progression-Free Survival, and Overall Survival. J Clin Oncol. 2023;41(15):2706-12.

  - The assessment of overall survival can be as a primary or secondary efficacy endpoint with a formal testing plan including control of overall study-wise Type I error.

  - Regardless of whether overall survival is included as an efficacy endpoint, an assessment of overall survival should include a pre-specified safety analysis designed to assess for potential harm.

  - Sponsors should consider the amount of overall survival data needed to adequately address key safety and/or efficacy objectives (further details regarding pre-specification of study design features related to the number of overall survival events are included in Section III.B.1 and III.B.2.).

  - Note that use of early or immature overall survival data (i.e., relatively few deaths) can cause high uncertainty in treatment effect estimates.

- Interim analyses of overall survival for futility or harm should be included in the clinical trial protocol and SAP when appropriate.

  - Protocols for large, randomized oncology clinical trials should include interim analyses for futility or harm early enough to limit participant exposure to therapeutics that are potentially harmful.

  - The timing of the interim analyses must balance the need to limit exposure to potentially harmful therapeutics with the uncertainty associated with early information and the potential to prematurely halt investigation of a therapeutic.

  - The timing of the interim and final overall survival analyses should be event-driven rather than based on a pre-specified time period.

  - Sponsors should provide a rationale and justification for the timing of the interim analyses proposed.

  - Independent data monitoring committees should be provided with the results of planned interim analyses for futility or harm. Any decision rules based on the interim analysis results should be specified in the data monitoring committee charter. Study personnel should not have access to unblinded interim data in order to maintain study integrity. Trial integrity is best protected when statisticians preparing unblinded data are external to and independent from the sponsor. If a sponsor's statistician independent from the study personnel is used, a firewall should be implemented and should be well-described in the charter.[8]

[8] See the guidance for industry *Establishment and Operation of Clinical Trial Data Monitoring Committees* (March 2006) and the draft guidance for industry *Use of Data Monitoring Committees in Clinical Trials* (February 2024) for additional information related to maintaining trial integrity and the data monitoring committee charter. When final, this guidance will represent FDA's current thinking on this topic.

  - In most trials, interim analyses of overall survival would be appropriate to include; however, interim analyses of overall survival may not be appropriate in some small randomized clinical trials or clinical trials in diseases with a low overall survival event rate (i.e., anticipated long overall survival times).

- Crossover

  - Protocol-specified crossover, is when trial participants who have disease progression on one arm are allowed to initiate the therapeutic intervention of another arm. Protocol-specified crossover has been included in clinical trials.

  - Inclusion of crossover in clinical trials can improve enrollment and help with retention of enrolled participants.

  - Crossover can impact the interpretation of overall survival results because it can confound treatment effect estimates. In general, its use should be limited but it would be most appropriate to include in clinical trials of disease settings that have no other or very limited therapeutic options.

- Unequal randomization scheme (e.g., 2:1 randomization)

  - Sponsors should consider the utility of using unequal randomization in the protocol, as this may reduce the precision in the assessment of the difference in overall survival between arms and the ability to adequately contextualize safety findings.

- Sponsors should include in their clinical trial protocols adequate plans for collecting data to minimize the amount of missing data for the assessment of overall survival and other long-term outcomes, including details on how data will be collected with regards to capturing long-term safety and efficacy information.

  - In general, treatment discontinuation should be distinguished from study withdrawal, and all randomized participants, including those who discontinue treatment and/or use subsequent therapies, should continue to be followed for overall survival through the end of the controlled period.

  - Participants and investigators should be counseled regarding the importance of obtaining adequate follow-up information, even after discontinuation of study drug treatment or after conclusion of active trial participation.

  - Diverse approaches to minimize loss to follow-up may be specified, such as increased efforts to maintain contact.

  - Participants should be adequately followed after any intercurrent events to capture important safety and efficacy information. If there is subsequent therapy use,

reasons for initiating subsequent therapy and time to next therapy should also be captured.

- If non-proportional hazards are anticipated, then additional considerations for adequate overall survival assessment should be pre-specified, i.e., the study should be designed for long-term data collection, and sample size and power calculations should be based on justifiable assumptions.

B. Statistical Analysis Considerations

This section includes considerations for pre-specification of statistical analyses, assessments for harm, and post hoc analyses of overall survival. Sponsors should, a priori, specify the statistical analyses of overall survival, prior to knowledge of any study results or data unblinding, in a clinical trial protocol or accompanying separate document (e.g., SAP). If the protocol and SAP do not adequately characterize plans for assessing, measuring, or analyzing overall survival, evaluations of overall survival may be considered post hoc. FDA also considers additional analyses conducted on overall survival outside the scope of those pre-specified in the SAP to be post hoc and for exploratory purposes (e.g., for hypothesis generation).

1. General Considerations for Specification of Statistical Analyses

- The SAP should prospectively detail assessment timing and methods for analyzing overall survival.

- The primary analysis (or main analysis) of overall survival should be specified in the study protocol in order to define the primary metric for decision-making based on overall survival, even if overall survival is not a primary endpoint. The choice of primary analysis (and primary estimand) for overall survival should be based on the stated study objectives.

  - The primary analysis for overall survival should generally be based on the intention-to-treat (ITT) principle, that is, all randomized participants with comparisons based on randomized treatments.

  - FDA generally recommends the hazard ratio (HR) obtained from a Cox proportional hazard model as a summary measure, along with its 95% confidence interval. In some cases, pre-specified covariates may be justified to be included in the Cox proportional hazards model.[9]

  - FDA generally recommends the log-rank test for the primary analysis of overall survival when there is alpha allocation, and either a stratified or unstratified test should be pre-specified.

[9] See the guidance for industry *Adjusting for Covariates in Randomized Clinical Trials for Drugs and Biological Products* (May 2023).

  - Sponsors should generate a Kaplan-Meier plot of overall survival information for the primary analysis population and as part of other supplementary analyses, such as for any subgroups of interest.

- In addition to the analyses specified for the primary evaluation of overall survival, the SAP for overall survival should also include at minimum the following details:

  - Pre-specified analyses to assess any assumptions for the methods used (such as proportional hazards). Note that if non-proportional hazards are observed, early overall survival data may not represent the true overall survival effect, and additional analyses may be needed to adequately characterize overall survival (see the last bullet in this section below regarding pre-specified sensitivity/supplementary analyses and Section III.B.3 regarding post-hoc analyses).

  - Planned length of long-term follow-up, planned frequency of follow-up, specified final analysis timing at an appropriate number of events to provide adequate survival information, and number/percentage of participants expected to be lost to follow-up.

  - Strategies for censoring and handling incomplete and/or missing data and intercurrent events.

  - Key assumptions and justification for overall survival analyses at each analysis time point, including the following, where applicable: the number of events, corresponding information fraction, and boundaries (on the p-value scale and translated to the estimated treatment effect) for efficacy, futility, and harm, with statements specifying whether boundaries are binding.

  - In studies with hypothesis testing of overall survival and multiple endpoints, a statistical plan for multiple testing should be pre-specified to control the overall study-wise Type I error rate, including a table of all relevant testing/alpha spending scenarios listing details for each scenario at each analysis time point.

- Intercurrent events, such as crossover or subsequent therapy use, can impact the hazard rate over time, affecting the shape of the survival curve. Sponsors should prespecify the handling and interpretation of intercurrent events in the SAP. Sponsors should consider the following:

  - Assessment of the timing, pattern, and reason for intercurrent events is important to determine the potential impact on the assessment of overall survival.

  - The baseline characteristics of participants experiencing and not experiencing an intercurrent event should also be assessed for potential subgroups of participants more at risk for such events.

289 – All-cause mortality should be considered for the primary evaluation of overall
290 survival, rather than an endpoint excluding specific causes of death.
291
292 – The primary analysis should include all survival information, including survival
293 times and deaths after subsequent therapy (i.e., a treatment policy approach within
294 the estimand framework).
295
296 – Alternative strategies to handle certain intercurrent events may be justified.
297 Note that strategies to handle intercurrent events should be supported by a
298 rationale for the plausibility of the underlying assumptions (such as non-
299 informative censoring).
300
301 • Sponsors should provide supplementary and/or sensitivity analyses based on different
302 methods and/or assumptions to inform the evaluation of overall survival and the
303 robustness of results, particularly if the intercurrent event rate is high, uncertainty is high,
304 or results are marginal. Appropriate sensitivity analyses should evaluate the robustness of
305 results, and tipping point analyses could help evaluate how sensitive the observed results
306 are to various assumptions. Supplementary analyses should target different estimands and
307 can provide supportive information regarding overall survival.
308
309 – If non-proportional hazards are anticipated, sponsors should specify in the SAP
310 additional sensitivity analyses to evaluate the robustness of the overall survival
311 results. Supplementary analyses, such as landmark survival rates, restricted mean
312 survival time, or other metrics, should be specified in the SAP and may provide
313 supplementary information to further characterize overall survival.
314
315 – Sponsors should evaluate narratives and toxicity data to understand the causality
316 of deaths, such as if deaths are due to progression or toxicity. Sensitivity analyses
317 that account for attribution of the death to treatment can be performed to assess
318 the impact of deaths potentially due to causes other than treatment, or other types
319 of competing risk analysis can be performed. For example, COVID-related
320 mortality may be a concern for some studies. Participants with COVID-19 could
321 be more vulnerable to drug risks, and risk of death could be increased on
322 treatment. For this reason, COVID-19 deaths should be included in all-cause
323 mortality.[10] However, a supplementary analysis may be performed where
324 COVID-19 deaths are censored.
325
326 – Using different approaches to handling subsequent therapy or crossover can
327 evaluate different estimands via supplementary analyses. Common approaches
328 include evaluating restricted mean survival time or utilizing causal models and
329 methods to adjust for treatment switching. However, many of these methods rely
330 on unverifiable assumptions and generally are not suitable as the sole or primary
331 evaluation of overall survival.

[10] See the guidance for industry, investigators, and IRBs *Considerations for the Conduct of Clinical Trials of Medical Products During Major Disruptions Due to Disasters and Public Health Emergencies* (September 2023).

2. *Assessment of Harm*

- To adequately inform regulatory decision-making, any trial should be designed a priori to adequately capture the number of events needed to rule out harm based on a specified threshold(s). Sponsors should specify and justify in the SAP a threshold or a range of thresholds to indicate the potential for harm for the overall survival summary measure.

  - Trials should be designed to collect sufficient overall survival data in order to rule out a clinically relevant degree of harm with a pre-specified degree of precision (e.g., based on the number of expected overall survival events, the power to rule out a pre-specified threshold of harm can be provided). Different scenarios based on a variety of plausible assumptions (e.g., varying number of events, overall survival effect size, or thresholds for harm) can provide support that the study is adequately designed to rule out harm.

- When determining the pre-specified threshold(s) for harm and proposed degree of precision to rule out harm, consider the following: disease setting, known safety profile of the product or drug class, input from patients and physicians, expected number of events, length of expected survival, duration and severity of expected adverse events, use of subsequent therapies or crossover, potential for non-proportional hazards, evidence of benefit from other endpoints, expected benefit of the control, acceptable level of uncertainty, feasibility of follow-up, and other available data.

- There are multiple approaches[11] to assess if study data can rule out harm, and the approach will depend on the degree of precision that is feasible (or degree of uncertainty that is acceptable) and the overall survival summary measure used. If there is high uncertainty and/or marginal overall survival results, additional follow-up may be required and may be the only reliable way to adequately assess harm. In cases where a study is not designed to test for overall survival superiority, uncertainty in the estimate of the overall survival HR is expected, and there may still be uncertainty even with long-term data collection in diseases with long survival times.

  - When the sponsor expects that sufficiently mature overall survival data will be collected or the study is designed to demonstrate superiority of overall survival, one option that can be used is to evaluate if a 95% confidence interval for the overall survival HR excludes a clinically relevant threshold. The threshold should be justified, and the study should be designed to estimate the overall survival HR with a sufficient degree of precision.

[11] Rodriguez LR, Gormley NJ, Lu R, Amatya AK, Demetri GD, Flaherty KT, Mesa RA, Pazdur R, Sekeres MA, Shan M, Snapinn S, Theoret MR, Umoja R, Vallejo J, Warren NJH, Xu Q, and Anderson KC, 2024, Improving Collection and Analysis of Overall Survival Data, Clin Cancer Res, 30(18):3974-3982, doi: 10.1158/1078-0432.CCR-24-0919.

- In cases where a study is not designed to test for overall survival superiority, there may be high uncertainty in the estimate of the overall survival HR. In this case, one or more pre-specified thresholds greater than or equal to 1 can be used. Assessment of harm can be based on the point estimate of the overall survival HR or the confidence interval for the HR at a pre-specified level (e.g., 90%) excluding the threshold. Another approach may be via a Bayesian decision rule, if an effective sample size has been pre-specified for safety evaluation purposes.

- If the sponsor anticipates that overall survival data will be immature or there will be high uncertainty in the overall survival summary measures at the time of the proposed final analysis, sponsors should conduct additional calculations or simulations to assess if it is likely that harm may be ruled out with additional follow-up time (which may or may not be feasible). Various methodologies can be used to calculate the probability of ruling out harm based on observing hypothetical future data. Sponsors should specify in the SAP approaches to rule out harm under a variety of scenarios, justification of methods, and justification of assumptions (including the assumption of observing additional events over a specified time period). Sponsors should include a variety of assumed scenarios for future data. However, the uncertainty in these methods generally increases for longer looks into the future, and evaluation of harm cannot solely be based on hypothetical future data.

3. *Post Hoc Analyses*

- Sponsors can conduct post hoc analyses to supplement their pre-specified analyses. Post hoc analyses of overall survival can provide additional insight into the potential for harm.

- For studies lacking pre-specification of overall survival analyses (as recommended in this guidance), analysis of overall survival should still be performed. In this case, sponsor should perform additional supplementary analyses and simulations to investigate the robustness of results under diverse assumptions.

- Post hoc thresholds for harm can be based on similar considerations for a pre-specified threshold; however, in the post hoc setting, multiple measures that reflect the full survival distribution should be considered, including the overall survival HR, and multiple credible thresholds, or grid of thresholds should be explored under a variety of scenarios. However, it may not be possible to detect harm based on studies not adequately designed to assess overall survival.

- There are several caveats regarding assessment of overall survival in the post hoc setting, including:

  - FDA considers post hoc analyses exploratory and suitable for hypothesis generation but not suitable for statistical testing with alpha allocation.

- It is not possible to determine if bias is introduced when retrospectively selecting statistical approaches to analyze overall survival or when selecting approaches to evaluate harm.

- Summary measures may have high uncertainty if based on immature or limited data, making the assessment of harm unreliable.

- Summary measures may also have high uncertainty if there was extensive crossover, non-proportional hazards, or other potential confounding factors not considered and accounted for at the time of study design, making the assessment of harm unreliable.

- The level of data maturity based on the observed number of events may not be well-defined if the trial was not prospectively designed to assess overall survival.

- Post hoc evaluation of intercurrent events is likely to be unreliable. Sponsors should assess the impact of multiple alternative intercurrent event strategies if they were not pre-specified.

C. Subgroup Considerations

This section includes considerations for evaluation of overall survival in subgroups. Based on prior information, there may be known subgroups of patients where the overall survival effect may be expected to vary across subgroups, such as biomarker-defined subgroups or populations vulnerable to adverse events. When those subgroups are not identified a priori in the study protocol or SAP, FDA considers subsequent analyses of these subgroups to be post hoc.

- Known or biologically plausible subgroups should be prospectively identified in a study, and a plan for adequate overall survival data collection within the subgroups should be pre-specified. This includes provisions for extended follow-up periods to enable a robust evaluation.

- Depending on the intended use, analyses of overall survival in subgroups can be exploratory and need not be alpha-controlled; however, these analyses should be pre-specified.

    - Pre-specified analyses should include subgroup analyses by demographic characteristics (age, sex, race, and ethnicity), region, and anticipated prognostic or predictive characteristics or biomarkers.

    - Sponsors should account in the study design for subgroups known to potentially experience differential treatment effects on survival, and analyses of these subgroups should be pre-specified. The importance of enrolling subgroups of sufficient size and conducting formal hypothesis tests for overall survival within these subgroups depends on the expected difference in survival or potential for

harm in each subgroup, and the expected contribution of each subgroup to the entire overall survival treatment effect.

  - Additionally, if there are known subgroups and potential for using the overall survival subgroup results for labeling purposes, pre-specified study-wise control of Type I error is strongly recommended for hypothesis testing within these subgroups (and for all primary and key secondary analyses).

- The following should be included in the study protocol when designing a trial with where the overall survival effect is expected to vary across pre-specified subgroups:

  - Biological or mechanistic rationale for expected differential treatment effects on survival based on subgroups;

  - Anticipated survival benefit or detriment, within each of the subgroups;

  - Expected contribution of the subgroups to the overall treatment effect;

  - Prevalence and plans for adequate enrollment of patients in subgroups;

  - Feasibility of and justification for any planned stratified randomization for prognostic or predictive subgroups;

  - Comprehensive analysis plan for the subgroups, detailing objectives of each planned overall survival analysis (whether for safety or efficacy), and pre-specified thresholds for harm;

  - If a formal hypothesis test for overall survival is planned within subgroups, the number of events that would be required to detect a clinically meaningful difference in survival in each of the subgroups.

- Subgroup results should be interpreted with caution, especially if overall survival data are immature or analyses were conducted post hoc. Findings based on subgroups defined post hoc or post hoc analyses of pre-specified subgroups may result from random patterns and are considered hypothesis generating, not evidence of benefit. However, such findings may be sufficient to raise safety concerns, warranting further investigation.

  - There should be consistent results across subgroups, any observed differences should be further evaluated (i.e., additional post hoc analyses should be conducted).

- Subgroup results cannot be used to salvage a trial that was not successful based on evaluation of the overall ITT population.[12]

- Subgroup analyses indicating potential detrimental overall survival could result in halting the trial, indication restrictions, or post-marketing requirements.

- If there is a known subgroup with potentially poor survival, this subgroup may be excluded from enrollment if there is sufficient prior knowledge that these patients would not benefit from the therapy. Alternatively, if these patients are included in the trial, early interim subgroup analyses should be pre-specified to mitigate concern for potential harm, particularly in cases where patients require longer term therapy and where there is expected to be a sufficient number of events in the subgroup.

- In addition to pre-specified subgroup analyses, additional post hoc analyses to explore relevant subgroups can be used to investigate any safety concerns in a successful trial.

  - Sponsors should identify any post hoc subgroups that appear to benefit the most, explore any subgroups that appear to lack benefit, evaluate potential overall survival detriment in any subgroups, and explore subgroups identified from data external to the trial (if applicable). Sponsors should include a justification for identification of any post hoc subgroups, including clinical relevance.

  - Post hoc analyses of subgroups can include investigation of whether there is repeated evidence of overall survival detriment in the same subgroup from other trials using the same product or class of product.

  - Also see Post Hoc Analyses recommendations in Section III.B.3.

D. Regulatory Considerations

- Benefit-Risk Determination

  - If the trial is intended to support a regulatory submission, the SAP should specify that at the time of the primary analysis of the primary endpoint (if the primary endpoint is not overall survival), an analysis of overall survival will be conducted to assess the potential for harm.

  - Overall survival data may not be mature at the time of an early interim analysis, but an assessment of the available overall survival data is important in the determination of the benefit-risk profile. Sponsors should contextualize the

[12] Amatya AK, Fiero MH, Bloomquist EW, Sinha AK, Lemery SJ, Singh H, Ibrahim A, Donoghue M, Fashoyin-Aje LA, de Claro RA, Gormley NJ, Amiri-Kordestani L, Sridhara R, Theoret MR, Kluetz PG, Pazdur R, Beaver JA, and Tang S, 2021, Subgroup Analyses in Oncology Trials: Regulatory Considerations and Case Examples, Clin Cancer Res, 27(21):5753-5756.

available overall survival information and feasibility of obtaining additional overall survival data to inform regulatory decision-making.

- Accelerated Approval

  - The accelerated approval pathway is intended for drugs that address an unmet medical need in the treatment of a serious or life-threatening disease or condition. Specifically, FDA can approve a marketing application on the basis of, among other things, adequate and well-controlled clinical trials establishing that the drug product has an effect on a surrogate endpoint that is reasonably likely to predict clinical benefit, or on an intermediate clinical endpoint that can be measured earlier than irreversible morbidity or mortality that is reasonably likely to predict an effect on irreversible morbidity or mortality or other clinical benefit, taking into account the severity, rarity, or prevalence of the condition and the availability or lack of alternative treatments.[13,14,15]

  - Often in oncology, accelerated approval is based on a clinical trial that evaluated response rate or another intermediate clinical endpoint (e.g., progression-free survival, etc.).

  - If there is significant uncertainty in the overall survival results, but efficacy is supported by an intermediate clinical endpoint, accelerated approval may be the most appropriate approval pathway. Traditional approval can then be considered at a later time when there are sufficiently robust and interpretable overall survival results available.

- Postmarketing Requirements or Postmarketing Commitments

  - FDA may require a sponsor to conduct postapproval studies or clinical trials as a postmarketing requirement (PMR) where the statutory criteria are met[16], or FDA may enter into a written agreement with the sponsor to collect postapproval data as a postmarketing commitment (PMC).

[13] See Section 506(c) of the FD&C Act.

[14] See the guidance for industry *Expedited Programs for Serious Conditions – Drugs and Biologics* (May 2014).

[15] See the draft guidance for industry *Expedited Program for Serious Conditions – Accelerated Approval of Drugs and Biologics* (December 2024). When final, this guidance will represent FDA's current thinking on this topic.

[16] See, e.g., section 505(o)(3) and 506(c) of the FD&C Act. See the guidance for industry *Postmarketing Studies and Clinical Trials—Implementation of Section 505(o)(3) of the Federal Food, Drug, and Cosmetic Act* (April 2011) and the draft guidance for industry *Postmarketing Studies and Clinical Trials-Implementation of 505(o)(3) of the Federal Food, Drug, and Cosmetic Act* (October 2019). When final, this guidance will represent FDA's current thinking on this topic.

576 – If the overall survival data at the time of approval is immature, FDA may require
577 additional postmarketing overall survival data through a PMR or request such
578 information through a PMC depending on the situation.