

A Potential New Approach to Assessing Progression-free Survival in Solid Tumor Studies Using Independent Review Committees

Kristin Borradaile, Director, Clinical Operations
Kevin D. Duffy, Senior Vice President, Global Clinical Operations
CoreLab Partners Inc.



January 2013

EXECUTIVE SUMMARY

On July 24, 2012, the Oncologic Drugs Advisory Committee (ODAC) convened to discuss the merits of a potential independent review committee (IRC) approach to evaluating the efficacy of solid tumor clinical trials with a primary endpoint of progression-free survival (PFS). The original intent of the IRC was to eliminate the perceived bias associated with investigator assessments, specifically for trials where the investigator may be unblinded to treatment arm. Based on analyses conducted on past trials that included independent review, the PhRMA PFS Working Group and the FDA demonstrated there was no systematic introduction of bias in investigator assessments. As a result, the recommendation was made to evaluate the primary efficacy endpoint using investigator assessments while utilizing IRC in a sample audit capacity. Based on this analysis, a full IRC review would only be required in the event bias was identified in the investigator assessments. The ODAC agreed the IRC audit approach is a viable alternative, however many unanswered questions remain and the approach is still under review by the FDA. At the time of this article, no final decisions had been made regarding the size, timing, or methodology of the IRC audit and there are concerns regarding the timing of the decision to require a full IRC review. One of the primary concerns is the dramatic shift in regulatory burden from the IRC to investigator sites. This includes the cost and training of site radiologists, difficulty and cost of mandating standardization, the management and monitoring of the reader process, and the ability to maintain an audit trail of image annotations at all trial sites. Without formal guidance from the FDA regarding the recommended audit methodology, Sponsors should use caution and consider the challenges this proposed process presents, including the potential to increase trial costs and the potential delay in time to market when a full IRC audit is retrospectively required. If the IRC audit approach is employed for a given trial, a “Collect and Hold – Read Ready” approach with the option for an upfront audit of the first subjects enrolled is recommended.

The ODAC agreed the IRC audit approach is a viable alternative, however many unanswered questions remain.

INTRODUCTION

Approximately four years ago, an industry committee, the PhRMA PFS Working Group, convened to begin discussing the use of the PFS endpoint in oncology trials and the value added by independent review facilities. This group was specifically interested in making recommendations that would reduce the costs and burden associated with conducting oncology trials with imaging endpoints. This resulted in the recommendation for a modified independent review approach for large registrational studies which was published in numerous journal articles in early 2011. In 2012, the FDA referred this discussion to ODAC for their opinion. On July 24, 2012, ODAC held a public meeting to discuss the *Evaluation of Radiologic Review of Progression-free Survival in Non-hematologic Malignancies*¹.

ODAC MEETING SUMMARY

Background

The original intent of the IRC was to eliminate the perceived bias associated with investigator assessments. This was thought to be particularly relevant for unblinded studies where the investigator knew which patients were enrolled in the treatment and experimental arms of the trials or for blinded studies where the toxicity profile associated with one or both treatments had the potential to un-blind the treatment assignment.

The implementation and management of the IRC process is costly and burdensome for sites and clinical trial Sponsors. The average cost to include IRC was estimated at \$4,500-\$7,500 per patient (\$1-\$3M per study) depending on the size and complexity of the trial. In addition to this, statistical issues associated with the use of an IRC have been widely discussed within the industry over the past few years. This includes an average discordance rate of 30% between investigator & IRC assessments and within the IRC (when utilizing the two readers with adjudication reading paradigm). Discordance between the investigator and IRC assessment of progression results in informative censoring when the investigator identifies progression and removes the patient from the study without an IRC-identified progression event. Disproportionate informative censoring between treatment arms leads to the introduction of bias into the trial.

Data Analysis

The PhRMA PFS Working Group analyzed data from 27 past randomized clinical trials where an IRC was utilized. Despite patient-level discordances, they concluded there was a high degree of correlation between investigator and IRC-determined PFS treatment effect at the population level. Because FDA drug approval decisions are based on population-level statistics, the use of IRC did not change the outcome of the trials. It was concluded that no systematic bias was introduced by the investigators in the 27 trials. Additional analyses based on trials submitted to FDA produced similar results.

Based on their analyses, a recommendation was made to utilize IRC in more of a limited role. Specifically, IRC would be used as an audit function in a sample of subjects to evaluate the consistency in treatment effect measured by the investigator sites and IRC. The intent of this approach would be to reduce costs and logistical burden, avoid missing data issues at the IRC, and mitigate statistical issues encountered, such as informative censoring.

Proposed Audit Methodology

With the proposed approach, the investigator assessments would be used as the basis for the primary endpoint analysis instead of the current standard which is to utilize the IRC data. For unblinded trials, an IRC audit of a sample of subjects would be performed to confirm the investigator assessments were free from bias. It was estimated that 100-200 subjects would be required to be read by the IRC as part of the audit, however the sample size would be variable and dependent on the treatment effect determined by investigator assessments. Based on the required sample size, small studies (phase I and phase II) would require 100% IRC review. For truly blinded trials, it was discussed whether IRC would be needed. The consensus was that IRC review would be desired in order for the investigators to be aware that a level of oversight was included in the trial.

Two potential audit statistical analyses were proposed:

1. Analysis of population-level hazard ratios between the investigator and IRC assessments, or
2. Analysis of “differential discordance” among treatment arms between investigator and IRC assessments.

Based on a pre-defined threshold, if bias in investigator assessments was identified, 100% IRC review would be required. This approach was recommended for solid tumor registrational studies utilizing PFS as the primary endpoint. Further discussion would be needed for different tumor types and complex studies that require multiple imaging modalities.

CONSIDERATIONS FOR CLINICAL TRIAL SPONSORS

Research Sample

The data presented during the ODAC Meeting was comprised of a 27-trial analysis by the PhRMA PFS Working Group of published data and a similar analysis conducted by FDA of submitted trial data. Both analyses concluded that no systematic bias had been introduced by the investigators. A high degree of correlation was observed between the investigator- and IRC-determined PFS treatment effects as measured by hazard ratios and objective response rate as measured by odds ratios.

These analyses may have been conducted using a biased sample of studies for the following reasons:

1. All included trials had independent review performed by an IRC. Therefore the investigators participating in the trials knew their work would be 100% reviewed by a third party. It is questionable whether their assessments would have differed without this knowledge.

2. By including only trials with published data or data submitted to FDA, it is reasonable to believe that all or most of the trials had favorable results by the IRC. With a larger treatment effect, a high degree of correlation between the investigator and IRC assessments would be expected. It is questionable how the results would have differed if trials with favorable investigator results but unfavorable IRC results, or vice versa, had been included.

Preliminary Discussions; No Final Decision

While ODAC agreed that the data presented suggests the IRC audit methodology is a reasonable alternative in certain settings, it was noted that additional research and discussion was needed on many points:

- The size of the sample audit: may be based on treatment effect observed during the trial.
- The method of random subject selection: a truly random sample would likely exclude some investigator sites from the IRC audit.
- The timing of random subject selection: if based on treatment effect, could be selected towards the end of the trial.
- The statistical method used to compare the investigator and IRC results.
- The audit threshold of differential discordance or hazard ratio differences that would result in the need for 100% IRC review.
- The timing of the above decision.
- The scope of studies eligible for the IRC audit approach: The FDA indicated that the proposed approach applies to solid tumor studies and not to hematologic tumors since other factors (such as blood counts and physical exam findings) are incorporated into the assessment. Regarding solid tumor studies, ODAC indicated further discussion and research was needed in order to make recommendations for tumor types that are more difficult to measure, those which incorporate both radiographic assessments and biomarkers (such as prostate cancer), and those which require additional types of imaging assessments other than CT scans.

ODAC agreed the IRC audit approach is a viable alternative and more cost-effective, however more research is needed and the specific audit strategy per study needs to be determined on a case-by-case basis with the FDA.

Upcoming Guidance for Industry

In August 2011, the FDA released their draft *Guidance for Industry: Standards for Clinical Trial Imaging Endpoints*². Within the draft guidance, recommendations are made regarding when centralized image interpretation is important:

“The need for centralized (core) image interpretation process is contingent upon the role of imaging within the trial. In situations where image interpretation results in measurements representing important components of trial eligibility determination or safety or efficacy endpoints, and these measurements are vulnerable to considerable variability among clinical sites, a centralized image interpretation process is needed. A centralized image interpretation process also is critical to controlling bias in open label trials. In general, compared to a site-based image interpretation, the centralized process can better provide verifiable and uniform reader training as well as ongoing management of reader performance, ensuring that the process is accurate and that bias and variability are minimized.”

The exception to the above includes, “a randomized, double-blinded clinical trial of an investigational therapeutic drug where the imaging technology is widely available, the image is easily assessed by a clinical radiologist, and the investigational drug has shown little or no evidence of unblinding effects”.

The draft guidance also stresses the importance of outlining detailed information about the blinded review in the Charter, blinded reader qualification, reader training, re-training, and reader performance monitoring. These requirements are logistically difficult to achieve in the site read setting.

The final version of the guidance was expected to be released by the FDA in October 2012. Due to the open items remaining after the 24Jul2012 ODAC Meeting, it is unlikely that the recent discussion or opinions from the meeting will be incorporated into this guidance. This represents a potential discrepancy in the current thinking at the agency. As a result, it will be necessary for Sponsors to seek study-specific guidance and recommendations on the type of independent review approach required.

Shift of Regulatory Burden to Clinical Trial Sites

Radiology reads performed as part of routine clinical care and those performed by the IRC for the purpose of regulatory approval differ greatly. Radiologists in clinical practice are typically unaware that a patient is enrolled in a clinical trial for which specific protocol requirements apply. Outside of major research institutions, clinical radiologists are largely unfamiliar with the response criteria required in clinical trials and rarely select and measure target lesions at each time point. In addition to this, it is not uncommon for a different radiologist to read each time point for an enrolled patient, increasing the variability of

assessments. In order to meet clinical trial requirements, a high percentage of trial sites utilize study coordinators or oncologists to measure the lesions that are qualitatively referenced in radiology reports in order to fulfill the CRF requirements associated with clinical trial protocols.

As discussed earlier, the draft *Guidance for Industry: Standards for Clinical Trial Imaging Endpoints* goes into great detail regarding the burden of reducing variability in assessments during imaging trials. This includes specifications regarding image quality assessments, qualification and training of readers, standardization of image display and interpretation, the measurement tools and reading system, options and requirements for image manipulation, and monitoring of reader variability and assessments. The need to make site reads auditable, as they will comprise the primary endpoint data, simply shifts the cost and burden from the IRC to the sites and is logistically infeasible. The following considerations apply:

Making site reads auditable shifts the cost and burden from the IRC to the sites and is logistically infeasible.

- Cost and training of radiologists (the number of clinical radiologists can be high at each individual site)
- Training of 100+ sites when some may only enroll a couple subjects
- Ensuring site compliance as it relates to image quality and adherence to protocol requirements
- Difficulty and cost of mandating standardization (image display, image analysis tools, etc.)
- Management and monitoring of the reader process
- Process for replacing readers whose performance falls outside of acceptable thresholds

The IRC workflow is specifically designed to produce greater consistency and reduced variability in image interpretation. The main components of the IRC workflow are listed below:

- Consistent small group of radiologists
- Standardized reader training, qualification, and testing
- Controlled image viewing
- Standardized measurement tools
- Derivation procedures and edit checks to ensure response criteria are applied appropriately and consistently
- Variability monitoring
- Fully auditable source data

Timing of IRC Audit Decisions and Regulatory Filings

There are concerns regarding the timing of IRC audit analysis decisions as it relates to the timing of regulatory filings. Depending on when the decision is made regarding whether a 100% IRC review is needed, there could be a delay in the Sponsor's filing resulting in increased costs for the Sponsor and a delay in access to treatment for patients that need it. The potential that the FDA could retrospectively request a 100% IRC review during the time of NDA review further increases this concern.

Requirements during Regulatory Submissions

There are concerns regarding the auditability of site reads, especially if they will be used as the primary endpoint in regulatory submissions. The source data in imaging trials is the image with the lesion measurements and annotations included. The data transcribed onto trial CRFs do not represent source documentation. CRFs do not typically contain screen shots of annotated images and annotations are not typically saved on clinical review systems at the sites. Therefore, monitoring typically occurs between radiology reports and CRFs, not the annotated images. Conversely, the IRC is able to provide the FDA with annotated image archives demonstrating all tumor assessments contributing to trial outcomes.

Statistical Concerns

With the proposed audit approach, additional statistical issues can arise when 100% IRC review is retrospectively required. Currently, Sponsors employ a reading paradigm where the IRC conducts real-time progression confirmation reads to prevent statistical issues associated with informative censoring. If an IRC audit approach is implemented, and 100% IRC review is later determined to be necessary, informative censoring could become an issue since real-time confirmation reads would not have been prospectively performed.

Recommended Image Collection & Read Approach

With the need for an IRC audit of a random sample of subjects and the potential need to conduct 100% IRC review, it will be necessary for the IRC to collect 100% of subject images throughout the trial in a collect & hold scenario. Random subject selection will occur towards the latter part of the study (after the completion of subject enrollment, completion of an interim analysis, or completion of the entire study). To mitigate unnecessary delays and the potential for missing image data, it is recommended that the image collection occur up front.

A "Collect & Hold – Read Ready" approach is recommended in order to ensure a timely IRC audit or 100% IRC review upon notification from the Sponsor. This includes the creation and finalization of the Charter and database design towards the beginning of the study.

A primary concern will be the potential for the FDA to require 100% IRC review following the completion of the IRC audit. Depending on the timing of this decision, significant delays could affect the timing of the regulatory filing and review. Another risk mitigation strategy is an upfront IRC audit based on the first subjects enrolled into the trial. While these may not be the same subjects randomly selected for the FDA-required IRC audit, an early analysis of the investigator/IRC assessments of the first enrolled subjects could provide more information as to whether or not it is likely that a 100% IRC review would be required. If the initial results suggest investigator bias is present, the Sponsor may opt to have the IRC read 100% of scans prospectively to mitigate the potential delay associated with a late decision to require 100% review.

CONCLUSION

The IRC audit is a potential approach Sponsors can choose during the conduct of registrational solid tumor studies. Without formal guidance from the FDA regarding the recommended audit methodology, Sponsors should use caution and consider the challenges this proposed process presents, including the potential to increase trial costs through substantial investigator site burden and the potential delay in time to market when a full IRC audit is retrospectively required. If an IRC audit is employed for a given trial, a “Collect and Hold – Read Ready” approach with the option for an upfront audit of the first subjects enrolled is recommended.

REFERENCES

1. FDA Briefing Document – Oncologic Drugs Advisory Committee Meeting: Evaluation of Radiologic Review of Progression-free Survival in Non-hematologic Malignancies. July 24, 2012.
2. Draft Guidance for Industry: Standards for Clinical Trial Imaging Endpoints. USFDA. August 2011.

CoreLab Partners is an industry leading, independent core lab providing centralized cardiac safety and efficacy services, and independent medical image assessment solutions. Our clients are global pharmaceutical, biotechnology, and medical device sponsors of Phase I-IV studies.

For more information, visit www.corelabpartners.com or contact us at +1 609 936 2600



Driven by Science