

# Pharma Imaging Network for Therapeutics and Diagnostics (PINTAD)

## MEETING SUMMARY

<b>MEETING SUBJECT:</b>	<b>PINTAD 2014 Telecon</b>
<b>DATE / TIME:</b>	<b>31JAN14 / 11:00 AM EST</b>
<b>PREPARED BY:</b>	<b>Barbara Chandler, Annette Schmid</b>
<b>LOCATION:</b>	<b>Teleconference</b>

### DISCUSSION POINTS:

1	<p><b>Q &amp; A with Dr. Ohad Amit (OA) on his thoughts and suggestions on the implementation of the Audit Methodology. Dr. Lori Dodd was unable to join the call. Discussion led by Annette Schmid (AS).</b></p> <p>1. The NCI method is right now implemented on Local Evaluator (LE) success with a sponsor-determined Clinical Irrelevant Factor (CIF). Should the CIF be predetermined prior to the study or are there any recommendations on the CIF relative to the LE declared Hazard Ratio (HR)?</p> <p>OA: You would want to pre-specify some clinically important threshold and how that ties in with the local evaluation. Don't leave it open ended to say we want to determine the target after the fact.</p> <p>2. What would be the scope of the proposed audit process?</p> <p>OA: Aim to try to keep it relatively limited - 100-200 patients. What you can do is control sensitivity of the procedure. There's a trade-off between sensitivity and specificity. I could do an audit of 150 patients and think I could get to 90% sensitivity. By that I mean I'm going to detect some systematic bias relatively well. My trade-off would be specificity. Maybe only 50% of the time I may be proceeding to full central review in absence of bias. Lori's procedure is more based on choosing a fraction of the overall population and that's what drives the operating characteristics and to some extent that's true of the Pharma procedure as well. If you start having to do this with 400-500 patients, then that defeats the purpose. We published some information of operating characteristics. Within the public literature there needs to be a good understanding of operating characteristics of different audit sizes. You can't just pick at random because it feels good. At a minimum you need to convince FDA that you have the right size to detect what you are looking to detect if you had an issue. It speaks to more work needing to be done to understand it better.</p> <p>3. How would the quality/interpretation be measured? (aiming at hazard ratio vs events)</p> <p>OA: Our procedure is looking at differential discordance which is a good surrogate for systematic evaluation bias. The gold standard is hazard ratio. You can look at either. Both are valuable tools. What you're looking to do is evaluate the robustness of what you've seen with the local evaluation.</p> <p>4. When would the audit be conducted (before or during or after the study) and by whom,</p>
---	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## **Pharma Imaging Network for Therapeutics and Diagnostics (PINTAD)**

would it be remote, or on-site at the sponsor, or at the study investigator sites?

OA: This is where there may be some difference between the two approaches. In an ideal world what you do is when you have your clinical cut off, that's when you perform central review of scans. The results from the independent review are available alongside the results of the local evaluation. When you unblind the study, that would be part of the analysis at that time, presumably part of the analysis by the sponsor. You'd look at the local evaluation and simultaneously look based on the audit for evidence of bias. If there was, then potentially you would have to proceed to a full review of the remainder of the scans. The one difference between the Pharma procedure and the NCI procedure is the NCI procedure needs the local evaluation results and the hazard ratio from the local evaluation which would only be available after unblinding in order to dictate the size of the sample for the audit.

Question: Why does the study need to be unblinded for audit to take place?

OA: Two aspects to the audit. One is the actual reading of the scans. That you don't need to wait in the Pharma procedure to be unblinded. In the NCI procedure you would need to get a sample size based on the local evaluation hazard ratio. There you would need to be unblinded to get that hazard ratio.

In terms of the analysis of the results of the independent review, the analysis of the audit, clearly you need to be unblinded there, because what you're evaluating is differential bias between the treatment arms.

AS: People are looking into options to at least start central read portions as the study is ongoing, not wait until there is evidence of superiority to start an audit. That would save time. That's in Dr. Dodd's latest paper. It's an evolution that's occurred since the original paper. In the NCI method you'd need to have the unblinded data to determine the sample size. Dr. Dodd suggested coming up with values to allow to sample at the front end.

OA: You could start early in either method. Once you're enrolling you can identify the sample, samples of different sizes.

5. What would the model be if a pharma wanted to use this on an ongoing basis instead of end of study; could it be linked to an Interim Analysis- if so, what would be thresholds to establish the site reads are not adequate?

OA: Do you need to wait until end? During the study can you get a read on quality of local investigator and be done after that? There's some challenges to both methods. The Issue is you need progression events and a representative sample of the overall population. You can say you're going to take the first 10 progressors. But at the end of the day you still don't have a lot of follow up on a lot of patients. Methodologically we need to think about how to implement something like this. The sample you have early on looks different than the sample you have after enrollment and follow up are done.

AS: The question is do you have the right sample? This is another area where more discussion would be helpful.

6. What should be the focus of the thresholds for determining acceptable site reads? Would it be based on the clinical outcome at the end of the study to be most accurate?

OA: Yes to second question. Look at HR and differential discordance.

## ***Pharma Imaging Network for Therapeutics and Diagnostics (PINTAD)***

7. How would the gold standard be established? Would it be based on the evaluation of site data by a single central reader? If the reader does not agree with the site, would there be need for adjudication?

What would be the standard to compare with?

OA: I don't know it's about establishing a gold standard. It's about the extent you can have interaction between site and central readers and get agreement in real time on the assessment progression, is ultimately a good solution.

Question: What about the loss of independence?

OA: Probably would be loss of independence, it depends upon the ultimate goals.

8. Since the FDA seems content to let single local readers determine the patient date of progression, does that mean that they would also recommend that IRCs use only 1 reader instead of 2 primary readers and an adjudicator? If so, then the cost savings of a local evaluator/audit schema would be only a maximum of \$100k for a 600 patient/5 year study and very would increase costs to the sponsor due to fixed costs not associated with readers. Will the constraints of any audit be relaxed to ensure pharma can save money by this method? (Yes, this is veiled sarcasm)

OA: A few pharma have already gone to the single read model. I don't know that FDA has a huge issue with it. In terms of cost savings, I don't think I'm the one who can comment on it.

Comment: Most studies have 2 plus 1. ODAC expressed concern about high adjudication rate. That implies there is concern about the high variability of reads.

9. The C2325 study is often cited as the primary driver for the development of audit methodologies. The central review of futility was consistent with the OS results (HR=1.22!!) but the LE review was completely contradictory. With the indictment of local evaluation as clearly biased, why wouldn't this case study be reason to never use local evaluations? Did anyone look at the TMF to figure out why the LE assessments changed from the 2/3 interim to the final? Why did the FDA change its endpoint to BICR instead of the central adjudication?

OA: I don't know the data that well. The one piece of data I saw - local evaluation and central review at interim were not that far off. The truth is probably somewhere in the middle. I don't think that hazard ratio of 1.22 by any means is not an unbiased result. When you have bias in local evaluation, you clearly get a biased estimate from central review. If you really have systematic bias in a study, no amount of central review is going rescue that. The central review can detect bias but it can't get you to a good estimate of the hazard ratio. I think we've shown it gets you to a biased estimate of hazard ratio and the truth lies somewhere in between the two estimates. I don't know enough to say why it should damn the use of local evaluation moving forward. I will say that meta-analysis we've done and that the FDA has done has shown relatively good concordance in treatment effect. Will there be studies with bias moving forward? Of course there will be. It's a judgment call by FDA as to the amount of time and expense that we want to spend, and pharma as well, to detect that one study that has bias compared to the hundred that don't.

10. FDA guidance (3rd attachment, page 9, 2nd paragraph) indicated ""a randomized, double-blinded clinical trial of an investigational therapeutic drug where the imaging

## **Pharma Imaging Network for Therapeutics and Diagnostics (PINTAD)**

	<p>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". For a double blinded study we are planning, we propose not to utilize IRC audit due to the concerns listed above. INV evaluation is reliable with enhanced monitoring guideline, investigator training, and site monitor visit, study quality control etc... Any suggestion for our proposal to FDA about non-audit plan.</p> <p>OA: If it's a double blind trial, they've been on record to say go with local evaluation. The key thing to illuminate with them is - are you truly blinded? Their pushback has been, given the toxicity of cancer agents, investigators get functionally unblinded. Personally I don't buy that argument because to be functionally unblinded you'd almost have to have a zero percent chance on a control arm for a specific toxicity and a fairly high percentage of occurrence in the treated arm. Of course there are treatments where there will be some functional unblinding, but I think that's their key pressure point with that whole concept where you have to convince them that the trial really is blinded given the toxicity profile.</p> <p>11. Questions are based on a white paper – 4th attachment The white paper indicates that the primary concerns of audit are the dramatic shift in regulatory burden from the IRC to investigator sites. It includes</p> <ul style="list-style-type: none"><li>• cost and training of site radiologists,</li><li>• difficulty and cost of mandating standardization,</li><li>• management and monitoring of the reader process,</li><li>• ability to maintain an audit trail of image annotations at all trial sites.</li><li>• potential to increase trial costs</li><li>• potential delay in time to market when a full IRC audit is retrospectively required</li></ul> <p>Without formal guidance from the FDA regarding the recommended audit methodology, what is the author's recommendation to mitigate the above concerns?</p> <p>OA: We're powering trials based on site assessment and not increasing sample size. You power trials based on number of events and hypothesized treatment effect. Yes there's some variability around that treatment effect. We've designed several trials based on a local evaluation and haven't adjusted sample size. It's not like we're getting a lot of failed Phase 3 trials in that paradigm. It's not fair to argue now we need to increase the sample size of Phase 3 trials and that increases cost.</p> <p>AS: Other costs, delay to market, etc. will money be saved in audit method? Comment: In both papers no evidence that money will be saved with audit method.</p> <p>OA: Intuitively people are saying this will lead to cost savings. Putting some numbers around that will be a useful exercise.</p> <p>Annette asked if attendees are interested in workshops dedicated to particular aspects and to cost? Positive response and support for workshops</p>
<b>2</b>	<b>Next meeting</b> Friday, 28FEB14 11:00 am ET