Pharma Imaging Network for Therapeutics and Diagnostics (PINTAD)

MEETING SUMMARY

MEETING SUBJECT:	**PINTAD 2013 Telecon**
DATE / TIME:	**25OCT13 / 11:00 AM EST**
PREPARED BY:	**Barbara Chandler**
LOCATION:	**Teleconference**

DISCUSSION POINTS:

1	**Q & A with Dr. Lori Dodd (LD) and Dr. Ohad Amit (OA) on their thoughts and suggestions on the implementation of the Audit Methodology. Discussion led by Annette Schmid (AS)** **1. Available data:** • **Have Drs. Dodd or Amit worked on <u>any audit designs that have been submitted</u> to the agency? Details that they can share.** • **Some big Sponsors are so far reluctant to use the industry (Ohad's) method due to comments made by the FDA favoring the NCI method. Is there <u>any progress to validate both methods</u> for use in clinical trials to allow for some flexibility by sponsors based on their own needs?** • **Have they been able to evaluate any <u>prospective data to compare</u>?** • **Critical to the acceptance of either of the methods will be data supporting the successful implementation. Are either of the authors already involved or interested, <u>willing and able to collaborate on some road testing</u>?** OA: Personally hasn't worked on any audit designs that have been submitted but aware of other sponsors who have been working with FDA, don't have many details. • Refined our method from initial publication and do need to get that out there. • Key provision is how to choose the threshold which would move a study to a full audit review • In process of finalizing that, some validation work needed LD: Hasn't specifically worked with any sponsors on implementing the audit method. • Has provided code to sponsors from paper that allows them to evaluate the method. • Hasn't heard back about specifics or how they are planning or not planning to implement the audit methodology. • Not aware of any prospective data that would allow us to compare the methods • Not as involved as she was a few years ago now that she's at the Infectious Disease Institute working on tuberculosis drug testing. • Certainly interested in participating and collaborating depending on how much involvement needed. Thinks it's needed. • Could turn to colleagues at NCI to see if they're interested and have time. OA: Also working in infectious disease now • More work needed in road testing, just a matter of who's going to do it. AS: Are you close to publishing a revised paper: OA: Don't know date of publication yet. Working through a few issues then need to write it up. AS: Are you planning to present it at a conference? OA: Yes. Procedure was revised to make it conditional on full review (gold standard). • Revised threshold value to make it conditional on what is the gold standard which is a comparison of hazard ratios in full independent review vs local evaluation and also controlling the sensitivity of procedures, choosing the threshold value to control the sensitivity of the procedure at 90% to allow specificity to vary on the basis of the sample size. • Main missing piece was how to implement that in a single sample. Can do this empirically and choose the threshold value in a simulation, need to work through how to do that in a given sample to make it practical.

- Didn't define threshold value very rigorously. This is an attempt to be much more rigorous in the definition.

2. Methodology:

- **With implementing audit, the informative censoring issue is not mitigated. Any suggestion for how to handle informative censoring during audit.**

OA: Both procedures aren't designed to address informed censoring but to detect imbalances in informative censoring that arise as a result of bias and tell you to evaluate robustness of local evaluation.

- Issue with informative censoring is that it leads to bad estimates of treatment effect. Neither procedure gets around that. They're really designed to detect its impact.

LD: Agree with that. Try to make that point every chance I get because it's a source of confusion. Would never want to imply that it does mitigate informative censoring.

- Would be great if real time central review is the chance to mitigate it by allowing you to get additional scans if there is a discrepancy at the time an event was determined locally. Doesn't work in all trials. Has a lot of practical implications.
- Alternative is to get additional scan after the time of locally determined progression. That would reduce informative censoring considerably. Not practical in all situations but would be helpful.
- Approaches to impute the progression time in case of censoring are model dependent and from a regulatory perspective presume only to be used as sensitivity analysis

AS: Question about central reviewer impacting patient management since fast turnaround of progression confirmation is possible today.

OA: Some experience doing that. Doesn't think it would impact patient care. Just another piece of information for investigator to have in making decision about patient care.

- Investigator's judgment about whether to continue treatment in spite of PD assessment if a patient is achieving some clinical benefit. Depends on how the protocol is designed.

Methodology (cont.)

- **In an event-driven trial where the number of events will determine the end of the study, there is some very serious concerns from study sponsors on patients lost to follow-up since, that a future central radiological read would declare that no progression had occurred would mean the loss of an event with the possibility of delaying the filing. In the case of a first in class trial where multiple studies are competing for that honor, what is the <u>advantage of an audit over a confirmation of progression central read</u> that reduces the chances for lost to follow-up in near real time rather than waiting until it is too late?**

LD: Back in 2005/2006 we were thinking about real time confirmation. Still thinks that's ideal. Losing events – lose power. Votes for near real time reads if at all possible. Second best is to have the second scan even if it can't be read in real time. From limited data majority of the discrepancies occurred around the window of +/- 1 cycle of chemo. If you can at least capture one time after, less likely to have loss of events.

OA: Seen similar pattern in our trials.

Methodology (cont.)

- **If we are willing to accept that independent reviewer may also be highly qualified and experienced in reading the trial data, and we know that we never know who and how many people will be evaluating the patient status at the site- <u>should we be focusing more on what data needs to be presented to the central reviewer to be able to make an adequate assessment</u>?**

OA: More data is better. More you can have a robust read the better. Never going to replace the assessment of progression at the site.

LD: Nothing can replace the site determinations in terms of having the patient there and patient information in front of you. Risk in giving all that information to the central reviewer.

- May inadvertently unblind the central reviewer. Need to keep in mind the types of bias we're trying to eliminate.
- Depends on what data you present and how and is disease dependent to an extent. And potentially drug dependent.

AS: Key question is why we think there is a difference between the site and central? Is it because relevant data is missing or someone isn't doing a good job. It's easier to train central reviewers.
LD: Clear tradeoff between the variability at sites with informative censoring and loss of events that occur that's harder to control when it comes to central review. Both require increases in sample size to maintain the power that the study was designed to have.
AS: Good point that it will need to be the balance of informative censoring and variability.

Methodology (cont.)

- **With Cheson, MacDonald/RANO and modified RECIST already incorporating clinical data into the central evaluation of radiological progression, what is the advantage of relying on local evaluators since it will very likely cost the sponsors more to run an audit by several model estimates not accounting for the implicit cost of filing delays? Again, why not use confirmation of progression instead.**

OA: Agree with the last comment. Confirmation of progression in real time is more desirable that running an audit.
LD: Agree, absolutely
OA: Could almost do that in audit type fashion, just monitor first X number of progressions that come into a trial and if you see that you can exclude bias based on what is happening in the first X number of events that come into the trial and then stop it there rather than having to do confirmation of progression on every single patient. Subset wouldn't be a random subset but an operational audit where you look at the first 30-40 progressions then if things are going pretty well, would you need to keep reviewing the subjects?
AS: It wouldn't help avoid informative censoring if you stopped?
OA: Wouldn't be using that information to generate an estimate of treatment effect from central review.
AS: Just do confirmation of progression to see if there's a level of discordance but go with the site assessment.
OA: Right. In absence of bias informative censoring is a non-issue. Site evaluation is reliable and not subject to informative censoring.
AS: Variability will be greater and if we don't deal with the variability you would need to increase sample size.
OA: Statistically that's a valid comment. Haven't been increasing sample size to account for variability to date and seem to be doing a pretty good job getting drugs approved. Risk is when looking at drugs with more marginal treatment effects.

Methodology (cont.)

- **Have either of them modified their methodology since the original paper and if so, can they share details?**

LD: Joint publication with Dr. Zhang et al. No other work since that April 2013 paper.

- **When would the audit be conducted (before or during or after the study) and by whom, would it be remote, or on-site at the sponsor, or at the study investigator sites?**

LD: Method was initially designed to be fully retrospective, didn't' think it was worth investing in audit unless there was an effect based on local evaluation. Nothing to prevent sponsor from generating list of cases to be reviewed as study goes on. If it turns out there's no effect based on local evaluation, that money wouldn't have been well spent. In terms of where and who should be doing the central review, don't feel qualified to answer that. I haven't designed a central review.
OA: Sample size is pre-specified and subjects chosen at random ahead of time. Results for central review of subset are complete at time of study unblinding. Sponsor would perform analysis of audit results. Risk is that you detect some measure of bias and then want to proceed to full review.
Peter Eggleton (PE): Most of our studies are done using SPAs. If we supply an audit plan, would we also need a SPA to provide a complete independent read plan and charter in case the audit was not passed? Would FDA ask to see the full audit in any case?
OA: Why would charter be different for subset than for full set of reads?
PE: Number of radiologists would be different. The structure would be slightly different. The methodology would be fairly similar.
OA: Would do the central review in the same way whether I was doing a subset or the full trial

population.
AS: Typically you design in your charter how many cases or which cases are going to be read. And if you add additional cases, you would need to have at least at a high level the plan on what that is driven by.
PE: We had in mind something like an adaptive design where you go down path A or B depending on the outcome of the audit and it would be pre-specified how you make that decision.
OA: Haven't had experience discussing this with FDA. Can't envision they'll be completely dogmatic about this. Know they are encouraging sponsors in certain situations to do audits. They would want the procedure specified in terms of how many subjects you're going to audit and what are the potential outcomes of that audit and under what circumstances you'll do a full review. Pre-specify if full review is necessary, will augment the number of readers.

Methodology (cont.)

- **We see typical hazard ratios that are fairly close to the 100% audit mark. That could mean a delay of several months for filing if the audit occurred at the end of the trial with even more delays if a full central review was indicated. What is the approximate cost to the sponsor for these methods? Is it likely to be less than the maximum possible $200k savings for local assessments? Did your estimates of cost savings account for premium pricing for an expedited/emergency central audit and read?**

LD and OA: Cannot speak to cost savings

- **Both methods currently require unblinding to identify potential bias which is associated with logistical challenges and potential delays to market. What are the major roadblocks that prevent us from using reviewer performance metrics to determine whether the site assessments are adequate that may be implemented without unblinding?**

LD: Little confused about what is meant by unblinding. Thought we had come to an agreement in the context of a blinded trial central review wasn't needed. I know there's an ongoing debate about whether a trial can really be blinded. That is one paper on my to-do list because I think there's an advantage to having a partially blinded trial depending on how much unblinding occurs.
OA: The way I read it you have to break the randomization code. Even in an open label trial we don't break the code until we're ready to do the analysis. There's certainly nothing to prevent from using metrics to evaluate how well the site read is doing in real time. I would challenge that first statement. Don't think it adds a lot of logistical challenges and delays to market by having to have randomization code available to do these procedures.
What potentiates a delay is having to do a full review. Short of that I don't see anything else in the methodology.
AS: That's the real risk. You will need to wait until after the unblinding to see if you need to do a full read.
OA: Either you're trading off a cost savings for that risk. If you have a reasonable amount of confidence and in 90% of trials there's really no bias, then you 'll take that risk for the 10% of trials that you may end up with a chance finding or that you really do have bias.
Don't think having to unblind the studies in order to run these procedures really incurs a delay.
AS: Heard that delay of bringing a drug to market has cost associated in the millions. So 200k is irrelevant if every day counts. Fair amount of head to head trials are competing for the same patient population using similar drug activity. Immense pressure to be the first.
OA: Don't disagree. Sponsors don't want to delay a file. Experiences sorting out differences between site and central data even in the full review can derail you. Not getting that data until after the trial. Don't really know the issues until after unblinding. High level of discordance in informative censoring between two analyses can delay a filing.
AS: What is the gold standard?
OA: Don't think there is a gold standard. FDA has been very clear on that. And I agree. Don't think the central read is going to represent the truth or the gold standard.
AS: If our solution to bias is using a central read, essentially that is the gold standard. Could say undoubtedly there is no bias with a central read and it's a good tool for evaluating bias.
OA: Could say there is no bias associated with a central read therefore it is a good tool for evaluating

	bias. LD: Need to be clear what kind of bias we're talking about. Informative censoring does introduce bias in the central review. OA talking about reader bias. There's a trade off in biases. There's no true gold standard. Central review has the potential to have informative censoring. If no bias and site reader calls progression, they have an event. They're not censored. Then there is no informative censoring.
2	**Next meeting TBD.** Annette will send out Google poll to find a date that works for most attendees.