6-13-16 FDA-AACR Oncology Dose Finding Workshop – Session I Panel Transcript

Geoff Kim: We'll begin the panel discussion and I'd like to introduce five new panelists. I'm going to actually just go in order. I'm going to ask for you initial response, can you introduce yourself and then what interests you here, and what discipline you represent, then also your initial response to this very dense series of presentation that we just received, and your reactions to it. Kelvin, can I start with you?

Kelvin Dickenson: I would start by saying I'm a patient representative. Sorry, I would first start with I'm a patient representative. My understanding of modular subject matter is naïve to say the least. However, I think my overall observation is while much of this is scientifically very evolved, it is difficult to articulate probably, to a clinician, and certainly to a patient. There definitely is a need for a transitional layer, to be able to really explain why is this the dosing methodology, what is a dose that's effective, X instead of Y, and how should a patient really react to that. I think my observation beyond that would be two fold. One, and only one of the methodologies to really see a big emphasis on negative side effects, which was the one that we talked about in HP a lot. The other would be, I think most of these dosing methodologies don't really go beyond the initial phase of therapeutic effect and discuss how we can discuss disease resilience at different dose levels, and how that changes over time.

Geoff Kim: Great. Dr. Jin?

Jin Jin: I'm Jin Jin from Genentech, I had the mono-simulation drug in clinical pharmacology there. Oncology dose, up to my addition, is really a topic very close to my heart. This is also a very fast evolving area with a lot of methodologies and also new mechanisms of action such as immuno-oncology. I think this morning's session is a great one highlighting multiple new examples such as Pembro and Nivo that give a very comprehensive overview of the entire clinical develop paradigm, and the quantitative analysis together with also a small molecule example, a couple of them in the early clinical development phase. Also, in addition to that, there are also two very good methodology talks covering the statistical aspect of the dose response and also have [inaudible 01:16:00] base QSP approach. This is a very comprehensive overview and I'm very much looking forward to the coming panel discussion. Especially, this is such ideal forum integrating multiple disciplines with clinical pharmacology, the quantitative scientists, modelers and bio-statisticians, and also clinicians, regulators, and also the pre-clinical research scientists. I'm very much looking forward to the discussion.

Geoff Kim: Thank you. Dr. Mandrekar?

Dr. Mandrekar: Hi, I'm Sumithra Mandrekar. I'm a bio-statistician at the Mayo Clinic. When I was first invited to this panel, I think my primary interest is because I'm very much interested in dose-finding study designs, and adaptive dose-finding study designs. I think this workshop and all the talks this morning sort of spark my interest a lot in terms of how these studies, especially the immuno-therapy studies were designed

earlier on, I know all of the phase three trials. I didn't know much about the phase one data. To me, even now, after listening to the talks, I'm still sort of [bazercked 01:17:15] at the numbers of patients that were treated in the phase one, supposedly the phase one, portion of these studies. This is very intriguing to me. I know there's a lot of data that was presented today on exposure and bio-markers and PKPD modeling that probably necessitated the numbers of patients that were put on the trial. Often the patients, or 300 patients, in these cohorts ... I'm still interested in maybe having an offline discussions on what prompted those. Whether we're all hypothesis driven, whether more driven observational and then trained to test a hypothesis later on, or the pre-specified. It's just a lot of those unanswered questions that are still in my mind.

In the second aspect of the day that I hope we can get to in the panel discussion is how do we, going forward now that we've learned a lot from these initial development of these agents, how we do a more informed dose-finding study designs with these agents. Especially because you probably are not ... When you first did these studies, you probably weren't aware, what is the right endpoint, and how did these agents really impact efficacy. What are the bio-markers to target, what do you mean by exposure. I guess now we have a better understanding, so I guess my interest is in learning or trying to understand how we will go forward when we are trying to design studies with these agents.

Geoff Kim: Agree. Remember the session three is dedicated exactly to that question. Dr. Siu?

Lillian Siu: Hi. Lillian Siu, medical oncologist at Princess Margaret Cancer Center in Toronto. I'm a phase one trialist. I would currently our phase one program are by and large immuno-oncology, I would say at least 80%. This is a very helpful, and useful workshop. Two point I think I learned from this morning, or at least got my thinking. One is I think you still hear quite a bit about three plus three and I think it's time for us to really think about, in early phase clinical trials, how to learn to go beyond where we are and think about new, and even beyond adaptive design, innovative methods that are really helping us to use what we have learned so far in the long term to help us understand the surrogate short term endpoints that we need in phase one trials. By and large, these patients do not stay on trial long enough for us to see those long term endpoints. We have now learned enough currently to inform us that kind of knowledge.

Second, I think is having heard at least five companies, there's a lot of collective wisdom here. Not just in the clinical space but also in the pre-clinical and even metho-medical space, that I think we have not really taken advantage of. Certainly, if we talk about modeling, this is a great place to model such that we can actually learn in a collective fashion to help inform the next design for these kind of biological agents.

Geoff Kim: And a man that needs no introduction with this group, Dr. Wang from FDA.

Yaning Wang: My name is Yaning Wang. I am the acting director of the department of

pharmacometrics with the office of clinical oncology at FDA. You see, I think everyone agrees that this morning's session is a great one because it covers a wide spectrum of application from all the simulation in oncology, from small molecule, to large molecule, from pre-clinical to late phase, and to a more [inaudible 01:20:55] modeling. I've been working in this field for the last 13 years, but seeing oncology as a medical field moving forward at this moment on those issues when everyone is rushing through approval is very impressive. I still remember how difficult it was, maybe ten years ago, to [inaudible 01:21:21] those issue within oncology. Therefor, at this time, we dedicated this workshop on dosing from oncology, has, I think, a very important message for the whole industry. If I have to pick two key points from the morning, I would say one is this ... While you only study one dose level, you can imagine how difficult it is to figure the exposure response, as demonstrated this morning.

In fact, we realized this many years ago when we first published paper for exploring response for reception. Back then we already realized that if you only had one dose level, you try to rely on the individual exposure range to tease out patient response. There was a strong confounding effect between the drug exposure and the patient disease severity that's always given a very steep exposure-response relationship. In fact, we never believed that it was only due to exposure. Even that paper, we point out most of the difference in efficacy between the low exposure and high exposure was due to the unbalanced disease severity, not due to the exposure. In fact, in the last one year or so we learned even more. That is, not only the baseline disease severity is confounded with the drug exposure, but due to the treatment, the improvement of certain disease characteristics also confound the drug exposure making the typical steady state exposure not a very ideal PK merit to link exposure to response. Which highlights how important it is to study, at least two doses. To have a randomized dose. In addition to the dosage though, we still can do exposure response, but it gives us a much better sort of balanced patient population to tease out this exposure response.

In other ways, the importance of having a more mechanistic model in today's time to justify or optimize what regimen we should combine and what kind of dosing frequency, whether it's concurrent, or in sequence, because we have so many options these days. Chemotherapy, target therapy, immuno-therapy, everyone is moving to this combination therapy. With so many combinations, there's no way you can study them all. Therefor, I still remember almost ten years ago, within the FDA, we applied the first mechanistic model to justify the combination of two drugs with different mechanistic action for each regimen. Which made it a very highlight for the end of this 2A meeting. These days I think oncology, I see a similar trend moving to this multi-regimen, with different mechanistic action, with different dosing frequency. I guess the only option is to rely on this more mechanistic model. Of course, the mechanistic nature of this kind of model gave the strong predictive power of this model, which doesn't come for free. Every piece of this system pharmacology model requires data. I understand there's still gaps between the pre-clinical animal model to translate to human. I guess, only with continued learning can we further improve this model, eventually make it more predictive. That's my

two cents. Thank you.

Geoff Kim: I think that's a topic, definitely, we need to discuss about ho are we going to integrate and utilize the data we have to move forward. I want to be greedy and ask a question that got brought up with me. Basically, the concept of using tumor growth as a continuous variable rather than bucketing into the Reese criteria is a very important one. However, I'm interested to know whether in the phase one trials that were presented, were Reese's criteria used, which for those not familiar, Reese's 1.1 basically allows for investigators to call only limited number of lesions, and basically the rest lesions there follow. Did you use all the tumor measurements available, or were you following the criteria where you only had a limited ... Or did you simply just use the targeted lesions only? To kind of couple that with something that Dr. Shoe brought up, which was the variability assigned when we started looking at what we put into the model. We have a lot of variability when it comes to tumor assessments, there's a 15 minute coffee break study where an independent reviewer, or radiologist, reviewing the actual scan will have some variability from read to read. There's also machine variability and very technical variability when it comes to the actual scan itself. How is that built in and addressed when using tumor growth models as a continuous variable? I'll open it up to anybody.

Dinesh De Alwis: Is this on? I'm looking at Eric, who can actually comment on the very first part of your question. It was around the, I think it was [inaudible 01:26:27], in terms of resist 1.1.

Eric Rubin: Eric [inaudible 01:26:32], I'm the medical oncologist. We looked at just about every thing Jeff, to be honest. I think one of the issues, as you know, is that resists 1.1 is sort of the standard and I think for interpreting our data and looking at historical data, we had to do that. We did recognize that since this is an immuno-therapy that there could be this issue around pseudo-progression due to the mechanism of action so we used IRC, which we didn't publish at that time, later we developed our own one-dimensional version of that known as IR Resist. We looked at that and then of course, we also looked at the individual tumor growths. I think Dinesh, you presented some of that data here right?

Dinesh DeAlwis: Yeah. With those individual tumor growths all of the data itself ... You were not relying on the radiologist read, but had it reviewed for all lesions or the target lesions only. What I'm trying to get at is that for future use, because the whole point of this workshop is really to establish best practices moving forward. Is it the best practice to rely on just the conventional Reese's one four, the efficacy endpoints that we have, or do we need to do more comprehensive look at the tumor itself to look at all of the lesions even though we're not required to do so. Is there a better methodology besides the Reese's two-dimensional way. Has anybody started to look at volumetric imaging to see if we can improve the models even further.

Eric Rubin: Dinesh, I think, but I could be wrong, because most of our data was from central

independent review, where they're using assist algorithms. They are looking at target algorithms, so you get in the data base from measurements of the target lesions over time. I don't think they looked at every measurable lesion, but I could be wrong. Dinesh, do you know?

Dinesh De Alwis : No, I think you're right in terms of they don't look at every lesion. In terms of your point around volumetrics, we also did have volumetric measurements, which I didn't present here. Actually, results weren't, in fact, different. At least for melanoma. It might be different for different indications but the melanoma that the outcome, with respect to exposure response analysis, was the same.

Eric Rubin: What did it apply, because as you probably know, FNIH is sponsoring something called VOL-PACT, which I think is looking at your question where they're trying to access large data bases to see whether volumetric or other approaches can be more predictive for things like OS. There's a component looking to try to get immuno-oncology trials as part of that as well, which I think is fairly important.

Speaker 16: [inaudible 01:29:07] from BMS. We also looked into all the comprehensively of tumor tumor measurements. Resist, IR resist, pseudo-progression, and the tumor shrinkage over a period of time as a continuous measure to reality market of what is actually happening for cancer immuno-therapy early on in the program. For the tumor growth regression modeling, we actually used all the target lesion time on it, in fact, ten further comment on to the additional efforts that are going at Vienna's to utilize this as a measure for the activity.

Dinesh De Alwis : I just wanted to also confirm it was the target lesions and not all the lesions that would have a new measurement. To your other question, what is the best approach ... I don't think that this is necessarily the nest approach. Often times the approach that we have taken has been a function of what tools are available. Certainly, even with taking the ... Arguably, looking at the longitudinal continuous measure has a higher resolution than looking at categorical research responses. We didn't like looking at longitudinal measures, whether you do it one-dimensional or two-dimensional, they're actually quite comparable as is volumetric. We're still sort of lumping together multiple lesions. It's quite possible that if you only have a few lesion that the dynamics of the sort of overall tumor growth may be quite different compared to a subject who has many small lesions that have the same target tumor [burn 01:30:46]. There are methods that are now becoming available to actually start looking at individual lesions and modeling that sort of one-level additional hierarchy. There's been a paper published recently by Lina Freeburg from the university of Oxford that's looking at this. I think the methods to looking at this are evolving, the best practices are evolving, as new methods and sort of sessions like this and ideas come forward.

Sumithra Mandrekar: I can make general comments based on my experience and awareness. I think on the drug develop application side, as mentioned, is still mainly a summary of target lesions that's made a basic scenario for our tumor dynamics modeling. However, the field is very, very fast evolving. There are several efforts being done in multiple

places looking into for example, incorporation of new lesions, looking at individual lesions, even inter-lesion variabilities from that perspective. Also, not only the target lesion but also looking at the lesion at medistat inside. There are also effort, not only looking at summary diameters, but also looking at the tumor from 3D pattern, and also using the new imaging techniques, measuring the tumor size. Especially for oncology, the tumor size is composite of not only tumor cells, but also [inaudible 01:32:08] cells. That's actually, by looking at the tumor size by itself, may be confounded, especially for the early times with heavy immuno-cell infiltration. I think there are a lot of methodology development, both on the experiment design side, and also on the analytics side on using these methodologies. With the evolving science, I guess, many of these are more ... Haven't fallen into the best practice stage. These are more of methodologies developing using some implications into specific case examples. I think that really is the exciting part of these, is the science is evolving.

Speaker 17: I was wondering, from the clinical tester point of view, when we're calling it as this criteria, if there is a total percent tumor reduction rate, we are calling this upon everybody, there is 29% tumor reduction exposure responder. With the advancements in the tumor measurements, I was wondering how it would transform in the clinical practice when it comes to the treatment choice for individual patients. Whether we should continue treatment for 29% tumor reduction thinking that, with the cancer immune therapy and multiple modalities, that would lead to long-term survival. That is something I always wonder when I look into when I look into resist versus non-resist studies of tumor measurements.

Gabriel Helmlinger: I just wanted to add, I tend to think, as an engineer, that's my background, to me it's obvious that if there are measurements around but you don't use them in your analysis, you're just going to miss something. In terms of lesions, we have done some exploratory work in that domain when we're [inaudible 01:33:48] with people like [inaudible 01:33:51]. Dr. Wang, you might be aware of some of that work where looking at all lesions just improves the predictability situation looking at longitudinal data as well. We also looked at, in the few occasion here we had that luxury, looking at more than one baseline measurement of lesions like a pre-baseline and a baseline. Again, you can predict much more later on with what would be happening. Also this point from you [inaudible 01:34:20] on arbitrary thresholds. We have seen from the systems pharmacology work that some things work via thresholds and others work via continuous behavior. Just coming in early on and make a cut somewhere, you might be missing many things.

Eric Rubin: For those that don't know the story of how Reese's developed, that's a fascinating story of how to go back and understand how it was measured and why 30% was chosen and all that stuff. Again, we were faced with massive technological developments, so our standards and everything should be open to development as well. It's going to be hard to roll out certain things, especially ... From [inaudible 01:35:03]'s perspective, sometimes it just works. 30%, there's variability there, but that doesn't mean that that has to be applied directly from the development stage to get the knowledge. That knowledge should be incorporated when possible.

Male: Everyone on this panel agrees that tumor size is a useful endpoint, but in phase one setting when you have multiple tumor types, how do you integrate a tumor size changes across different tumors doing pharm. Does anyone have a comment on that?

Sumithra Mandrekar: The only way you can probably do that is if you have expansion goals where you write a whole bunch analysis of patient population, then try to understand the efficacy there. Not when you're dose-finding with two patients with multiple tumor types. I also have a comment going back to the tumor size versus using Reese's grade criteria. I think, Jeff, you brought up a very good point. Reese's itself was doubled up for a different reason, then now we are trying to expand it to use it to new agents which probably don't follow the same mechanism of action. I do have to say, there's lot of work out there that says continuous metrics have better relationship or correlation with survival compared to Reese's based metrics. Our group, we published three papers which mined the entire Reese's state of warehouse that was used to come up with the 1.1, so over 8,000 patients, multiple tumor types. We actually showed that none of the continuous metrics have any better predictive ability over categorical research. It wasn't just correlations, we tried to look at the predictive power, we tried to look at the C index, the concurrence index, and those were not targeted agents.

That database was not set up for targeted agents or immuno-therapy, those were more chemotherapy and cytotoxic. Completely different world. I just think we should be ... We just have to be careful when we start saying we won't use volumetric data, yes. But where's the benchmark. I don't know what signifies a meaningful improvement. I think we just have to be careful. Then also, when you talk about pseudo-progressions, we don't normally, I don't know maybe these trials are different, how long do you continue treating your patients beyond progression. How much data do you have beyond progression? It's very hard to just change something and then say, okay I don't want to use Reese's, I'm going to use the pseudo version of Reese's. I'm going to change this from trial to trial. It's very hard to benchmark them across trials.

Male: I just have a follow-up question. In terms of using different tumor types, does it matter what the primary tumor location was, or the organ of the tumor. Could we control for that in some of the new looking at tumor size changes.

Sumithra Mandrekar: I'm not a physician, I'm a statistician. I can tell you that yeah. What other variable do you think is important. Primary versus medistatic. If those are important than you want to make sure that your expansion goal is able to address those variations, then try to homogenize your population as much as you can.

Geoff Kim: Before we go, I wanted to get Dr. Siu's impression from this because ... It's really important because I think ... Only in oncology can we talk about principles and dose-finding phase one studies and automatically put caution around the interpretation to clinical benefit. Because we like to see things move that quickly if

possible, when impressed. Normally, this would be a step-wise kind of fashion. In the phase one population that you see, are you seeing, especially with the homogeneity of tumor types. Sometimes you get, because of the way we write our protocol eligibility criteria, we get very indolent tumors that perhaps they've been through several therapies. They definitely need therapies. Certainly, their course may be kind of truncated. That may reflect and kind of give you an impression of the nature of, especially about growth dynamics, and how lung cancer behaves. If this a lung cancer patient that has been through seven different therapies and is still able to hang on, is that different type of growth pattern that is not representative of the general population as a whole. Do you see that type of heterogeneity when you see a phase one trial population?

Lillian Siu: I guess stepping back a step, phase one ... It is a very heterogeneous group of patients. It is really, the primary endpoint is still safety. I know we are all looking at efficacy as our secondary endpoint. Of course, every patient we want to see that there is some evidence of preliminary anti-tumor activity. I think we do have to be careful not to draw conclusions about efficacy right off in a small number of patients of 30, 40 patients. When you have 1,00 that is a different story obviously. I think I want to echo Sumithra's point that it's hard to jump and make decisions based on what we've seen so far without the body of evidence to continue as measured is better than resist. I think if the body evidence tells us that volumetric measures or continuous measure is better, then it would be a new criteria.

Until then, we have to really speak the same language. Otherwise, you have people in one country or one center doing a different kind of cutoff or threshold for response, and another center doing completely different. How do you even compare data? Last thing, I think it is very important to take patient context into question. I think the question previously about should all lung cancer be considered the same primary versus medistatic. I think each case is different, and for me, certainly, making a decision with a patient in front of me, I need to know the dynamics of the tumor before and on treatment, and make a decision based on the clinical symptoms as well.

Geoff Kim: I think you bring up a critical point, and this illustrates the need for these multi-disciplinary conversations because I think, and the panel can correct me, the speakers can correct me if I'm wrong, the decisions we're talking about are not treatment related decisions. They're not, basically, decisions relying on whether or not a patient should be treated with a drug. These are developmental questions that need to be answered. The modeling and simulations that provide information for this type of thing is not ... The life or death critical situation where you have to make that decision for the patient. I think that's a really important concept to understand because we hear so many disciplines ... How do I put this the politically correct way. Basically, there's a lot of differences of opinions, and different related disciplines in development to answer these questions. I think some of the nomenclature that we use from the speakers that have spoke in the morning, and then the speakers who will be speaking in the day, that nomenclature differs, and those words carry a lot of loaded weight to it. I think that hinders a lot of the

disintegration conversation that we need to have take place.

Kelvin Dickenson: I just wanted to draw one point from a patient perspective. I think there is a material difference between how you would want to treat efficacy in metastasize tumors than non-metastasize simply because it's much more progressive in the earlier stage. The tolerance for side effects and the tolerance for aggressive treatment is probably higher in those cases, and also the need for higher efficacy.

Don Stanski: Don Stanski from AstraZeneca. I work closely with Gabriel Helmlinger who's presentation. A quick observation as to why this meeting is important and a question to the panel. It was intriguing to see the journey that Mark and BMS took to both dose-finding and getting approval of these molecules. I saw that at AstraZeneca when I had arrived there. It reminded me of a dart board. In other words, you're throwing darts and you're kind of hoping that some will get close to the center. Out of all of the darts you throw, and boy you threw a lot of them, in term of sub groups and changes, some of them landed and you made a decision. There was efforts to do PKPD modeling, but if you look at the actual immediate contribution, it generally was low.

Retrospectively, both speakers were trying to extract out PKPD modeling processes to try to rationalize the dose, but the clinical teams had gone forward just looking at efficacy and safety. Seeing these stories is important for all of us to see. The question to the panel is this. We're clearly moving to combination therapy to try to get response rates from 25 to 30% of the 50 or 60. We can't do, as Yaning has said, we can't do dartboard with combinations because there's too many ... Now you have two drugs, you have dose responses for each of them, how do we combine them. What has been the learning we've had to date? To try to get a more efficient drug development, more efficient regulation so that the next generation of combination treatments, both large-large and large-small molecules, can get the patient sooner and cheaper compared to what we've gone through. What smart person will take this on?

Speaker 16: I'll just try to share some of the perspectives at BMS for the ion-ion combination. Of course, the very easy step for us to take the approved dose of the one ion, and then add on to the dose escalation components for the other new and novel ion-therapy. However, I must congratulate and commend BMS clinicians that they took a different approach of optimizing or adding early on various different permeation combinations to really learn the safety and efficacy balance for particular ion combinations across tumor types. Not just for meta-normal, or not just for lung. In fact, one of the trial CO2O9O12 trial, which is presented at ASCO, is an example that how a PKPD as well as the various combination regiments led to benefit-risk ratio which is favorable for ion-ion combinations. I would highly encourage teams to look into the different modalities, including the various different dose levels for ion-ion or the different modalities. Ion-chemical combination as a ... But at the same time, not compromise the split to patients. I know that's really hard, and a really general answer.

Don Stanski: The dilemma is, it's still the old model of multiple combinations, multiple patients. What is the efficiency we can get? Yaning, you raised the point. What do you see as a more efficient way to study these combinations besides just doing every possible combination?

Yaning Wang: I work with, across all the disease areas. I you ask this question to different medical divisions, probably you get different answers. Again, oncology is evolving. At this very exciting moment, we're talking about dosing, and then the next step is combination, what regiment to use. If you look, again, try to learn the lessons or experience from other therapeutic areas, there are many other issues, in addition to what regiments to combine and what doses to use. One thing I can immediately think of is this ... We try to, for example, in the example presented. When you combine two approved regiments, you try to lower the dose for both components, believing there may be synergistic effect on efficacy, at the same time maybe lower some toxicity. One design challenge is it would still require factorial design to demonstrate the added value of each component. At least I, personally, have experienced at least two cases which basically, the single component at a lower dose is unethical to study. At the same time, we still, based on the law or the guidance, we still need to demonstrate the added value.

That's another area we need to address. Basically, do we need to demonstrate the added value relative to each single component, or as a combination regimen as a whole. As long as that regimen is better, or not inferior, to some control that's enough. There are many challenging issues, maybe different disease areas, they have different strategies to handle this. I guess, at least in oncology, given this momentum, I would say in addition to the empirical data collection, to look for optimal combination of whatever regiment or whatever frequencies. I think the mechanistic modeling, at the early stage, at the limit the ranges of testing possibility, at this moment is probably the most efficient way. Given, of course, the caveat of this gap between pre-clinical to clinical translation. I think over time we will figure out where those gaps are, and those pre-clinical mechanistic model will become more and more, I guess, predictive.

Eric Rubin: Can I just comment onto, sort of, Don's statement and also perhaps [inaudible 01:49:17]'s question on combination. Don, with respect to throwing darts on a board, I think the bar may be very low in oncology so we ... Any improvement is an improvement. I think that there would only be one dart thrown on that board, which would have been maximum administered dose, in all intents and purposes. The fact that we were able to actually take a significantly lower dose forward into phase two, I think, is significant progress, and also being able to explore 2004 dose range does enable you to really characterize the PKPD relationship which in fact, even if you look at other therapeutic areas, it's rarely do you actually have such a wide dose range, that means you can actually investigate the exposure-response relationship. In many ways, whilst I admit that was a number of arms and aspects of a type of test, within that there was actually some good PKPD driven stuff which actually drove the decision in the end.

With respect to the combination question, my approach, and that's why I was actually intrigued by what Gabriel was presenting, is I think, kind of the way forward, there's in that ... We figure out, and we do need to kind of restrict the number of combinations, the factorial design for like, in terms of things that we haven't tested in the clinic. The only way to do that is using some kind of informative knowledge-based pattern. That need to be from data coming from the clinic, as well as the pre-clinical experiments that are planned out to really inform what you're going to test out. The key issue though, is how do you best test this out in the clinic? You really need to define your patient population, that you need to test this out, based on your hypothesis. Based on either non-responders or you clearly define that. Then you can actually go in and try the combination out. That way you may be able to get a quick answer to if this works, rather than try every possible ... By the way, we haven't even got to sequence of administrations, at the moment we're just talking about giving them together. When you get into sequence of administrations that's a whole different guideline.

Jin Jin: I just want to make original comments on the combination therapy. I think at this stage pk model is still a very good way to start with. For example, now we for for [inaudible 01:51:45] therapies, we also have multiple drugs already as the combination for certain tumor types. However, if you add another agent to the combination therapy in the clinic, it sometimes is even harder, from efficacy perspective, and to have a feasible combination. It's important to evaluate the combination in the pre-clinical model, whether we really need the both compounds for the cure to be in the combination, and whether we can skip some of them and still have the synergy factor, or at least as good as the synergy factor by including all the agents. If that is tested in the pre-clinical model, and showed that the combination with one of them may provide better efficacy, then the original combination, then that can be a mechanism moving forward in the clinical setting. However, of course the dose question is still a need to be answered in the clinical setting, because it's really hard to translate the dose that you used in the pre-clinical to clinical setting.

The other comment I want to add is using the continuous measurement of the tumor size versus the Reese's criteria. I believe in the development of [inaudible 01:53:11] actually used continuous measures for to do the model, for the prediction. I think is also important for some of the drugs, they might be low action, actually is to lead to a [inaudible 01:53:28] disease, not necessarily tumor regression. By only based on the response rate, sometimes you will miss opportunity for those kind of drugs with certain mechanism of action. If you can prolong this new disease for a significant period of time, say six months, or even longer than that, that will translate to meaningful improvement to PFS or OS. That's another reason that reason that I think continuous use of variable should be sued in the development process for the evaluation of that efficacy.

Lillian Siu: Just one addition al comment regarding the combo. I think combo development definitely double, even triple, the complications. However, as a modeler, and given my opportunist nature, I also feel combo actually provide very unique opportunity

for modelings to provide value added, to increase the quality and efficiency for drug development. Because generally, when you develop combo, you will have some information regarding the single agent. How to leverage this historical knowledge of the single agent would be very critical and actually have the opportunity to improve the efficiency of the combo development. The single agent data may have teach you, already, somewhat aspect of the pre-clinical to clinical translation, or early to late clinical endpoint translation perspective. That's actually for both advocacy and safety. Even on the experimental side, anything you learn from the single agent, like how to monitor or manage adverse effects, many of these are ... You don't want to make the same mistake in combo. I feel, actually, these are also unique opportunities for our area to provide more value added to increase efficiency in drug development.

Mark Ratain: Mark Ratain, Chicago. I think we need to make some major changes in how we thin about all this. Unfortunately, it probably requires a reboot of the oncologist mind, particularly those involved in clinical trials. A complete, wipe the computer and put in a new operating system. Again, one doesn't need to reinvent the wheel. The wheel's out there. It's been rolling around for cardiovascular disease, pulmonary diseases, rheumatology, it's just like take that wheel and put it on our car. I'm mixing a lot of metaphors here. I think we would do well to get rid of the concepts of phase one, phase two, phase three. I think we have registration trials, and then we have, if we want to call them phase three, that's fine, then we have pre-phase three, or Learn And Confirm, as Lew Sheiner put it. That gets rid of this concept, as Lillian said, she does phase one trials, but she doesn't do phase two trials, because she said I'm a phase one trialist. That means Lillian's going to find your dose at the end of phase one, but she can't. Maybe she's going to learn that, that she just can't, so stop saying that that's what she does.

Lillian Siu: Tell that to 1,000 patients.

Mark Ratain: If we would sort of get rid of this mantra, phase one is you find the dose, phase two is you prove whether the drug works, and we started actually putting in what's the primary objective of the study. When that primary objective changes, fine, put in an amendment. We changed our primary objective, that's why we're enrolling 1,200 patients, because we have a different primary objective now. I think that would help us greatly to just sort of change this. There's only so much that individuals can do. I think it would be great if FDA took the lead on this and maybe out out some suggested protocol format, and subjects to nomenclature for studies. If we got rid of this ... I know it's in the sea afar, phase one, phase two, phase three, but it would be nice if we could change that.

Geoff Kim: Well, that's why we're here.

Speaker 19: I'm a [inaudible 01:57:37] from BMS. I just wanted to come back to a comment that actually Mark had made earlier on about the appropriate use of the pharmacogenetic models, and PKPD models, and perhaps extending to QSP models as well. I think of all these models as a sort of a working model of our

understanding of the system, particularly the QSP model. It's certainly not, I think we would all agree, that there's certain gaps in our knowledge. Any prediction that you might make from the model have, not only parameters, certainly there's model uncertainty, just because the whole model structure may not be quite correct. What's the appropriate use of the predictions from these kinds of models in coming up with a dose.

The approach that BMS and Mark can take was to use the PKPD data, use some of the model predictions, to come up with a minimum dose predictions. A dose that you expect would be efficacious in patients, that is not unethical to give to patients, and then go beyond that to some level that, not necessarily to the NTD, but going back to the idea of having ... You need to have multiple dose levels because that's how you would tease out your uncertainty in your model predictions, and then what other factors might be important. It ultimately might be possible to actually do a randomized multiple dose study because that's really the ... will give you the unbiased clinical data on which to make a solid dose recommendation.

Geoff Kim: Unenviable, last question before lunch.

Speaker 20: [inaudible 01:59:27] from Boston. I want to followup with [inaudible 01:59:30] raised the question about random trial. In an ideal world, you have two or three dose level randomized trials, but before we get into that there are a lot of situations, still there's one dose level, and then you have imbalance in the baseline characteristics. Is there anyway that we cannot throw away that data of [inaudible 01:59:53], so is there any way we can use an innovative analysis to address those imbalance using modeling or simulation. I know you have [inaudible 02:00:04] patient use case controlled analysis. That's more comparison. Do you have any more comments on that.

Speaker 21: Through the years we learned that the closure we move to the final clinical endpoint, for example, [inaudible 02:00:22] or even PFS. There are more confounding factors. Meaning, in addition to drug exposure, there are many other patient characteristics that can easily influence the clinic endpoints in a magnitude that's probably even larger than the drug exposure. That's why, in recent years, when we only have one dose level, we still tried to answer the exposure-response question. We tried to focus on this early PD marker, which is less influenced by those patient characteristics, and more influenced by drug exposure. We focus on, let's say, tumor size, especially target tumor size. Even there, we see, still, some confounding. Most patients with this less tumor reduction. Again, if we have to extract the most information from this single-dose study, I'd say move towards earlier response. In receptor occupancy, or some upstream bio-marker, which is less influenced by this confounding. I'll just give you a more real exposure response on those endpoints. Stay away from this very late phase endpoint.

Actually, another point is about this phase one trial. We have seen example where phase one data where you have multiple doses, even though every dose has a smaller subject, smaller sample size. When combined with phase three, or late

phase trial, the patient population is pretty much the target population, but clearly less severe than phase one population. When we pool the data together, and look at PFS for survival, you see a U-shaped dose response. That is because the final study dose in phase three is somewhere in the middle. Those patients were the less severe patients. Again, that's another point about why you try to pull this early phase one most severe patients with later phase, that's another issue. Again, that can be also classified as confounding because they have different severity. Therefor, pulling them in such a simple way will give you wrong, basically, even dose response, not even exposure response.

Speaker 22: I just have one follow-up question for [inaudible 02:02:47] I guess. As we know, most of the exposure-response analysis has been done using phase three data. One of the major reason is that, as you said, the confounding factors are very important in this analysis, and even with larger sample size, we cannot sometimes tease out this kind of confounding factors to the effect of the efficacy. I earlier stated with smaller sample size it will be even harder to tease the exposure-response relationship. Although, we have seen some examples of using multiple dose levels at early stage of development, but it is really not a ... I don't think it is a practice across all companies. I'm wondering from [inaudible 02:03:39]'s perspective, how many examples you have seen that people have done via thorough evaluation in the early stage of the trial that can actually provide meaningful guidance for the phase three dose selection, and how successful they are.

Speaker 21: I would say, we are seeing more and more examples, not just the large molecules like BMS and Burk, even small molecules. We're seeing more dose ranging. Not necessarily randomized dose ranging, but different dose level studying in phase one, not only just to identify MTD, but look for dose or exposure response for efficacy even for small molecules. I guess, it's just a historical sort of tradition. People felt somehow oncology is special, you don't do dose efficacy study, but I guess the landscape is changing. We have successful examples where the early, I guess, dose efficacy information was very helpful to help sponsor, to identify the ... Not based on the toxicity, but based on efficacy, sort of the optimal dose. Which turn out to be successful in the phase three. Again, these days, even with all these efforts, you still see a lot of phase three programs with dose that are too high with so much ... In fact, we have a case study in the afternoon to demonstrate you have so much dose reduction during phase three trial. Like you said, in probably practice after approval it will not be a very good thing for the patient, even for the company if you have such a high rate of dose reduction.

Geoff KIm: On that note we are out of time and we need to break for our lunch, so right now it is 12:08 on my watch. We'll reconvene at 1 o'clock with session two. I thank every speaker and panelist for their participation. It's been very fascinating, thank you.