

6-13-16 FDA-AACR Oncology Dose Finding Workshop – Session 1a Transcript

Pasi Janne: Okay, good morning everybody and welcome to our workshop on oncology dose finding. I am Pasi Janne from Dana Farber Cancer Institute from Boston, Massachusetts. It's my pleasure to welcome all of you here today, along with my co-chair persons for the meeting. This is the second year we're doing this meeting, and in fact the first year was quite a success as you can see from the AACR Focus publication that describe the findings of the meeting that was just recently published. Also from the fact that there are many more participants here today than last year, in fact we have 700 participants today. Now, it doesn't quite look like we have 700 in the room and that's because many people are viewing this through the webcast as well. We hope today will be an informative and interactive meeting, we have the scheduled the meeting with talks as well as panel discussions and hope all of you will find it fruitful and informative.

With that I'll turn over to my other co-chair persons.

Amy McKee: Good morning, I'm Amy McKee from the FDA. I want to say a quick word about what the difference is between last year's workshop and this year's. Last year we focused on the small molecules in oncology and really focused on how we move through those findings from our safety side. This year we broadened it to include monoclonal antibodies and also to take a look from the efficacy side. That's the primary difference between the focus of this one day workshop, this year and last year's two day workshop, so we hope you enjoy it. I'll turn it over to Eric Rubin.

Eric Rubin: Hello. I'm Eric Rubin, I'm a Medical Oncologist from Merck. I'll say that, first of all I'd like to thank AC and FDA for putting this dose workshops together. I think we all agree that it's very important to try to get the dose right. There are recent examples of where I don't think we've done a great job collectively in getting the dose right. In part I think this is because of the unmet need and the pace of development on oncology oftentimes we take doses into later stage studies that might not be, then turn out perhaps not to be the best dose and we end up doing post marketing types of things to try to sort that out. I do think you'll hear today about ways where we can try to keep that speed and yet do a better job. I think there are both advances in pre-clinical models, the mathematics behind dose finding and others that enable a better approach. I'd like to thank the speakers who have come together today to work on this. They've worked hard and I hope you enjoy the discussion. Thank you.

Geoffrey Kim: I'm Geoff Kim from the FDA. One of the main goals of this type of workshop is that we want to really spearhead a multidisciplinary approach. One thing we've learned is that not everybody talks to each and there may be gaps of information and inefficiencies introduced when groups don't talk to each other and we stay solely in our respective disciplines. We really want to promote a multidisciplinary discussion here and hope and expect that basically the talks will be very robust and involve multiple areas. To that point we've integrated the panelists so that all the disciplines, including our patient representatives who will have a voice and we

expect the audience to participate and ask multiple questions.

Just a couple of housekeeping items, as we introduce, for the sake of time, we're just going to ask all the speakers to follow the agenda and when they get up we're not going to breach the gap by introducing the different speakers. We ask the speakers to introduce themselves and just give us a brief introduction to who they are, what discipline they're representing and what they want to basically accomplish with their talk. In addition, after the talk, speakers are welcome to return to their seat to the audience and we'll move up the panel discussion to the four, two-row discussion panel up here when it comes time for the panel.

Without further ado we really thank you, we thank AACR, our co-host for this workshop for providing just invaluable support and a sponsorship for this workshop and to be able to make it happen. This workshop only happened because there was a lot of buzz from the first workshop we held last year. If people are interested please let us know and then we can start keeping it in mind for our part three next year. Without further ado I'd like to start off and we'll just call speaker one to the podium.

Dinesh DeAlwis: Good morning everybody, so speaker one is in fact Dinesh DeAlwis. I'm from Merck Research Laboratories and I head up the Late Stage Development in Quantitative Pharmacology and Pharmacometrics. I was very actively involved in the Chemical Development of Pembrolizumab. I'd like to actually thank the organizers for inviting me to speak here, actually for the second time. I spoke last year, so I guess I must have done something right, so thank again for your vote of confidence and I hopefully won't let you down this time round.

The title of my talk is the optimal dosing for targeted therapies on oncology. I'll focus on the development of Pembrolizumab. I'm actually going to split this into two sections. The first section is actually going to be brief but I'm actually going to try and set some guiding principles in terms of where we should be focusing our dose finding efforts. Obviously, the session is appropriate because it's going to be exposure responsive, there's no secret there. Really as to what we should be doing and how we should be doing it in terms of high level points. Then I'm going to talk about the case study of Pembrolizumab and take you through what's in all, and as close possible to the chronological events of how we got to that dosing and dose finding there.

Obviously I have some poetic license, things obviously look a bit drowsier when you look back at things but hopefully I'll be as close to the events as possible. On that note, this slide actually I think sums up a number of things in fact of what we've been doing for probably the last 30 years. Especially this slide, if you look at the left hand side of the slide, the figure here. We spent over 30 years concentrating on getting to a maximum tolerated dose, and finding numerous methods of trying to get there. We did numerous algorithms from starting with the 3 plus 3 to continued reassessment methods, modified CRM, tight CRM, accelerated titration and the list goes on. It's exhaustive.

Really this was appropriate for cytotoxics but really toxicity and efficacy don't really relate. In many cases, with the target agents don't really correlate that much. I think we should be moving away from really trying to identify the MTD to really getting towards what is in fact the biologically dose and what is the minimally efficacious dose, and really understanding the dose exposure response relationship of our agents. I think really this is where we should be maximizing our efforts. It isn't easy, it isn't a simple algorithm that can be implemented and it depends on the programs, there's a level of individualization. There are some overall guiding principles that I think we can adopt in terms of trying to identify this.

On that note, I will start to ... If I can get this to work. Okay. It starts with pre-clinical. I think we really need to use all available PK biomarker efficacy and safety data to really develop translational encyclical models that can predict the biologically effective dose and also the MTD to really increase the probability of technical success in the clinic. That's if MTD cannot obviously be quantified. Doing the right PD and efficacy experiment is really required. What do I mean by that? In many cases, most pre-clinical experiments are done with the end point of trying to get efficacy. What's actually very very useful for us is where the minimal, where we're just seeing the cusp of efficacy occurring and really understanding why dose exposure range. Because this is when we can understand what's actually driving the efficacy. Is it C max, is it AUC, is it time on target, for example. Really identifying target media to drug disposition.

This is very very key, so we need to move away from just trying to get experiments and just give us efficacy and really understanding what's driving efficacy. In phase one, we need to be really informed by pre-clinical information, we need to focus on collecting information for dose finding. Phototherapy is really collect appropriate PD bar markers and this can be direct target engagement in the case of biologics or something more indirect, maybe further downstream from that. The other thing we need to do is actually, we have obviously an excellent measure for solid tumor and that is using a continuous measure of change in tumor size. We tend to result to things like response rates and resist which effectively actually we lose information, so the power you get from continuous tumor size measurement is about 50% more for the same effect size. We really need to use that information, it's actually very very powerful and we can look for exposure response relationship with that.

The other thing actually we need to do in phase one is actually look for implementing noble designs and actually implement things like inpatient escalation and carry out PK-PD assessment where the patient access is on control. You can get a lot more valuable information that way because of variability with respect between subject variability is controlled in that circumstance. The other part I think we need to ... I guess the figure here is a little bit deceptive because the arrows are only going one way. In fact, once we're in the clinic and we have an understanding of clinical efficacy is occurring, we may actually have to go back and do some pre-clinical experiments and further refine our understanding and doses,

and schedules. This is actually very important, that it's not a one way street here.

In fact, even if you go back to crizotinib, it was developed as a C max but I got to the clinic and we understood it was an alkaline inhibitor. Then [Faizer 00:15:12] went back and did pre-clinical experiments really understanding the PK-PD relationship in terms of alkaline inhibition. We may have got it wrong and we understand this better once we get to the clinic and then figure it out, in that sequence. I'm not sure why this is ... Okay, in terms of phase 1b, 2 studies, this really needs to be randomized. This is not just randomized against the control ammo, or standard of care, but really randomized by exploring two doses, at least two doses of your oncology agent. I think this is something we really need to move towards and move away from this paradigm of just taking one dose forward into later clinical studies.

We need to look at surrogate end points to make early no go decisions based on safety and efficacy, we need to look for factors contributing to individualized response, exposure, demographics or tumor genetics. Now, as I mentioned, if you use something like a continuous tumor size measurement you get much more bang for your buck, you get much more power. You can actually start to look at co-variants and sub-groups that actually might be affecting the response. You can start to look at things like demographics and tumor genetics in this circumstance. The last point is really establish POC. POC means many things to many people but really what I'm talking about here is has the drug reached the site of action. What does a POC look like? Has the drug reached it's site of action? Have you seen some kind of target engagement? Have you seen some kind of down regulation of what we proposed initially from the pre-clinical stages? Has proof of mechanism been established? Has efficacy been established in the targeted patient population? This is extremely important. Is it better than standard of care?

Now we're in an age where we have some amazing therapies, I mean modulators which are just showing dramatic efficacy and impressive efficacy. I think we owe it to the patients, so when we bring therapies forward, they're actually differentiating from standard of care. Or at least going into patient populations that aren't responding and we're seeing some kind of benefit. Last, but certainly not least in this particular slide, has an adequate therapeutic range been established? Now, to do this you can't really do this with one dose. You need to do at least a decent dose range of something like 5 to 10 fold. Obviously that would depend on the particular compound. We need to take more than one dose forward and trying to establish therapeutic ranges. It seems to be standard practice in most other therapeutic areas but why is oncology different? I don't think it should be different, I think we need to take this forward.

it also enables us to do many things during the course of clinical development. We can change dosing, we can change schedule, we may even be able to change route of administration. We could do this by, once you've established a therapeutic range, we can do this by PK bridging for example, as opposed to actually carrying out a formal efficacy study. There's a number of things that it actually enables us to

do. The fact that speed tends to be the paramount optimization in most circumstances isn't really a good answer. I think I'm going to show you at least an example with pembro where we could do it with speed. It was registered, clinical development took around three and a half years and we were able dose finding in phase two.

On that note, just go back. I am going to ... There's someone else who is controlling it. I think I'm going to talk about Pembrolizumab and the trade name is Keytruda and during the course of its development it was called MK3475. As I said, I'm going to try and take you through as close as possible to chronological events and also show you some of the significant discussions I think we had once the data came forward. Okay, on that slide I'm going to ... It doesn't work. Okay, thank you. It's important, Keytruda is important and highly selective humanized monoclonal antibody for the IgG4 Isotype. It really blocks interaction between program PD1 and it's ligands PDL1 and L2 and enhances the functional activity of the targeted lymphocytes to facilitate tumor regression and ultimately immune rejection.

It really blocks interaction of PD1 on Peripheral T-cells enhancing T-cell response against tumors. Next slide. Okay. The pre-clinical data actually suggested that Keytruda would have an anti-tumor activity in multiple cancers and hence the reason for taking it forward. The Usind was opened on the 7th of January, 2011 and it was a phase one study of simulation in MK3475 in patients with progressive locally advanced or metastatic carcinoma and melanoma. Now, this is ... Next. Yeah. The part A1, dose escalation study, the objective was to really define DLT, MTD and maximum administered dose and really characterize POC. Very standard objectives. The design, I'm almost embarrassed to say, so it was a 3 plus 3 design. We started with the traditional paradigm, non-randomized studies followed by small expansion cohorts of 32 subjects.

We had 1, 3 and 10 every two weeks with the dose that we investigated. The reason for that selection was one was, in terms of the projected marginal safety, was around 400 and at 10 it was 40 and we had some pre-clinical data in terms of PK, PD data showing that the EC 50 at one ... I mean, we would get the relevant EC 50 with the every two week dosing. In terms of results, there was no DLTs at any of the tested doses. The objective response was observed in two out of the three first melanoma patients and the first response was seen at 3 mg per kg every two weeks in melanoma. Based on this strong activity signal, amendment was issued to expand the melanoma cohort and we went into 10 mg per kg every two weeks and we established the maximum administered dose and this was selected as the first dose to take forward.

Okay. No. Going backwards, yeah. Forward, I have no control over this. Next slide. Okay. Once we had enough data, we were able to actually characterize the peak. The peak is actually very standard stuff, it's exactly what you'd expect for IGG therapeutic oncology body, low clearance, limited volume distribution, low viability. We were able to estimate a 10 to 6 day half life. This is where we start to question why we're doing an every two week dosing schedule, we had a 10 to 6 day

half life. We had a lot of discussions around this and we started to deviate from plan, and we started to drive the clinical development towards an every three week dosing paradigm. We switched and we explored on every three week dosing paradigm.

Next slide. Meanwhile, obviously that was part A1, we had rushed through to part B1, which is looking at 10 mg per kg at the time. This is a non-randomized part and as you can imagine, we had very impressive efficacy and this is the efficacy that was seen in melanoma patients, both ipil pluralism at naïve and ipil pluralism at treated. There was 40% response rates in all melanoma patients, 43% in ipil-naïve and 33% in ipil-treated. That was a lot of excitement, there was a lot of energy, excitement and lots of things were going on in parallel. I know it's like everything needed to go and move forward with this. We actually submitted for breakthrough designation in November of 2012 and we received the breakthrough designation on January 17th, if my memory serves me well, in 2013. This was in fact the first oncology product to receive that. Very exciting stuff. Now, what are we going to do going forward?

Next slide please. As you probably know about this phase, one study became a very large phase one study and it had over 1200 patients, and we had something like 655 melanoma patients and around 515 non smoker lung cancer patients. That was something like 5 amendments just between December 2011 to September 2013. To answer emerging questions as you can imagine, there was efficacy here and then suddenly we had expanded it, we looked at certain things. There was a number of amendments and there was something like four-phased to study like parts including three randomized dose comparison sub-studies. All this was happening, meanwhile, I don't think many of you are aware but there was a small amendment called part A2 which was effectively in 12 patients but it was the PK-PD. It's actually hidden in this huge study, but this is probably one of the valuable amendments that we made and I'm going to talk about that.

We managed to convince the team that, "Can we just have a small amendment. Just give us 12 patients and we can explore PK-PD with this." Next slide. It works, great. Thank you. The part to dose expansion part was to evaluate PK-PD on every three week dosing schedule. We did this with intra-patient dose escalation to explore the PK-PD with Keytruda from doses from point 0.005 to 10 mg per kg every 3 weeks. A 2000 full dose range, so you can imagine when we first suggested this, the kind of reaction we got from some of the clinicians. We calmed them down and said we're going to do this in a manner that no patient will actually get sub-therapeutic doses. The way we did this was, patients were escalated in 3 steps at days 1, 8 and 22 from low, so 0.005 to 0.06 mg per kg to 2 and 10 mg per kg. These were the high doses.

Every patient got, if you like, 2 and 10 mg per kg eventually. The PD acid that was used was an XV4 IL-2 acid and in this basically, there's no IL-2 released from lymphocytes when there is an activated PD1 pathway. The superantigen endotoxin B actually causes IL-2 release. When you give Pembrolizumab this is significantly

enhanced by the effects of pembro and you can have a nice dose exposure response relationship with the ranges of doses. Effectively, when you give pembro at the highest dose, you should have maximum release of IL-2 which basically indicates that the IL-2 release mark really affects the function of blockade of the PD1 pathway by pembro, and it basically utilizes the measure of target engagement.

This is what the data looks like, and we have a lovely PK-PD relationship, so they have IL-2 stimulation ratio on the Y axis. This is the range of exposures, obviously as I said there's a 2000 full dose range here and we have a lovely inhibitory Emax model that you can estimate here. You start with an IL-2 stimulation ratio of two and even here with maximum release, you have an IL-2 ratio of 1. This is basically, you can estimate using an Emax model. Based on this, we could estimate that with 1 mg per kg every 3 weeks, we could get 95% target engagement or such a ratio level. Simulation showed that at Ctrough of 1 mg per kg, we got 95% engagement.

Therefore, if you like, 1 mg per kg every 3 weeks, if you like, is a lower boundary of full clinical efficacy. Now, we didn't stop there, we had a number of modelling simulation exercises. We looked at, at 1 mg per kg every 3 weeks, what is the probability of achieving full target engagement? The probability is 64%. If you went to a larger patient population, 64% of the time we'd get that kind of engagement. If you moved this to 2 mg per kg we would get 90% probability of getting this engagement. If you went to 10, actually we didn't see that much improvement. Basically, 2 mg per kg then is, if you like, your minimum dose that is maximally efficacious. Okay. Our proposed efficacious at least based on this particular piece with this party to expansion was 2 mg per kg every 3 weeks. Okay.

That's one part. Then we decided we need another piece of evidence because when we looked at our pre-clinical data it wasn't that optimal for what we wanted in terms of translational. We decided to actually out, so we're in the clinic, we already have some idea where the clinical doses might lie clearly. We went back and did a pre-clinical experiment in over 300 mice. Chandni Valiathan is actually going to talk in great detail about this in the second session so I don't want to seem like I am ... I'm going to give a few key-points in terms of how we did this. The details will follow shortly.

In step 1, we developed this small model. By the way, we split this relatively large study into both training and a qualification data set to really understand how predictive this was within mice. We started with the mice where we had a mouse antibody and the PK there which is linked to the target engagement in occupancy in blood and occupancy in tumor which in turn was linked to tumor growth inhibition. This model was a Syngeneic Essen 38 tumor model. That's great, and as you can see we had a lovely PK-PD data here showing wonderful fits to the various doses. What does this mean to man? We needed to do some translational, we need to humanize this model. Clearly, the PK we had already had in man so that was easy, we just substituted with the human PK. Remember this is the target binding part was a mouse antibody so we needed to substitute the CAY on and CAY off-rates

from the human antibody, so we did that.

Then we had to link it to tumor growth inhibition. Essen 38 tumor growth inhibition in what we're looking at here in melanoma patients, so we looked at the literature. Now, when we fitted these models there were two key parameters, there was the kill rate and the growth rate. The growth rate, we actually looked at melanoma tumor growth rates in the literature and we substituted that, because that's a drug independent parameter. For the kill rate, we actually used different scaling techniques and Chandni is going to talk about that in terms of how we did it. One of the methods was arithmetic scaling. Based on all these, we were then able to simulate out what human tumor growth would look like on the right. That's this lovely very colorful bar chart here.

I'll cut to the chase, we had obviously resist with a few more categories including partial and major response. Major response here is greater than 50%. What we can see here is that the 2 mg per kg dose would be very similar to the other doses. Based on these two pieces of independent pieces of information, we decided we will take 2 mg per kg every 3 weeks forward, okay? Now, this is all well and good but meanwhile if you like, back at the ranch we had actual efficacy death coming from patients. Remember this is coming now from non-randomized studies. Please remember this point, non-randomized. If there's something you go away with today, please don't do non-randomized studies when you're trying to do these kind of analysis. Because it does cause potential confounding and bias.

Here we have observed response rates in both ipil-treated and ipil-naïve and what we found here was actually very interesting. We had 10 Q2 showing 56% response rates versus 27% for 10 Q3 This was actually very confusing and people started asking if we needed to go back to 10 Q2 or Q2 in fact. Then we looked at ORR in ipil-naïve and then we had 45% response rates with 2 mg per kg and 46% in 10, and 37% in 10 Q3. Very confusing results. As you can imagine, this caused more discussion and more confusion than actually shedding light. What we did was, actually instead of just looking at response rate, we actually started looking at the tumor size, the change in tumor size, a continuous tumor size measurements.

Here we have, the sum of longest time, we just changed from baseline at week 24. This two graphs are effectively the same, one is on a normal scale here, one is on a log scale. As you can see, this is actual exposure from the doses of 2 all the way to 10 Q2 and as you can see we have relatively flat exposure response relationship. It's a very weak relationship, in fact the alscARD is about 0.006. What this actually told you was that actually the 2 mg per kg was at the plateau of your exposure response relationship. When you looked at something like categorical response measure like response rate, you ended up putting it into one category or another. You have this artifact of relationship appearing when you look at categorical data. This is another reason why we shouldn't be doing when we look for exposed response relationships. Look for it with a continuous measurement such as tumor size, okay?

In addition to just exploring it in that manner, we did a full model based analysis as well. Here you have the observed sum of longest time which has changed from baseline and this is the observed pattern of tumor size change and this is actually in days. You have a classic pattern that we've now observed with immune modulators. You have these, obviously people who progress, and if you like, escape immune surveillance. Those who are, if you like, in stable disease or kind of an equilibrium and they are in the state for a significant length of time, and those who effectively show elimination or significant regression of tumor. You have these 3 types. We were able to model this and this is the predictions from our model, this is using only mixed effects. Actually we used the mixture model to categorize these categories. As you can see, the model does a pretty job in characterizing this. The basis of this was really not to just show this and say we can do a good describing the data, but it enables us to actually look for co-variants.

We looked at initial tumor size, we looked at ipil status, naïve versus refractory, we looked at different demographics, et cetera and try to assess how does that impact tumor growth change, et cetera. This is very very important going forward. It has the ability to simulate what the next trial might look like, et cetera, and make extrapolations. Based on this, now we can reduce this to resist response, this is the simulator, it predicts response as you can see. Basically there is really no difference between 2 and 10 in terms of progressive disease, stable disease, partial response or complete response. 2 is very similar to 10.

In addition to the efficacy, we also did analysis on AEs, and this is AEs of special interest with respect to immune modulation. This is a logistic regression model, these are AUC bins and as you can see we have a nice flat exposure AE relationship similar to the exposure response relationship. Again, 2 were similar to 10. Now, later, this is data from the randomized portion of our phase one study. That was part D and part B2, and this is ipil-naïve population and ipil-refractory, and as you can see confusion goes away when you do randomized studies and you have now 33% and 35% between 2 and 10 and the 10 month OS rates was 72, it was 64. Similarly if you look at the ipil-refractive group you have 25% response rates and the 6 months OS rate was 79, was at 77. Here are the ongoing response rates based on the follow up here. As you can see, there really isn't a difference between all the randomized parts, part B2 and D of the study.

I was very very encouraging, we had validation of the fact the two dose was really the right dose going forward. I think exposure response analysis were really key to identifying dose. A wide therapeutic range was established based on exposure response and exposure and I think this is really critical because once you do that this really enables a significant flexibility. When we want to go off into different indications and explore different things, and I think this is very very key. The approval of Keytruda was based upon positive response benefits and this is really efficacy based on the cohort B2 with 80 patients at the 2 mg per kg recommended dose. It received accelerated approval on September 4th, 2014 and as you know, you need to do two confirmatory trials based on this registration. We did that, this was with P002 and P006 and were conducted to confirm the safety and efficacy of

Keytruda and I'm happy to say they confirmed the safety and efficacy of Keytruda.

On that note, I'd like to end and really acknowledge a number of individuals. I've also had a couple of individuals here who are involved with a paper that we wrote around the guiding principles of what we should be doing. The significant list here really comes from the number of individuals involved with the effort that went to dose finding, the number of amendments that were made and being able to also back and forth between pre-clinical and clinical. Really, I think we've got to a very good place, despite the significant pressure with respect to speed. I just said we got approval in three and a half years of clinical development. On that note, I'll stop. Thank you. [crosstalk 00:38:29].

Geoffrey Kim: Kevin?

Kevin: You said that not having a randomized portion of the study was a really bad thing, so do you have some ... Did any explanation for the confoundness of the non-randomized portion emerge?

Dinesh DeAlwis: I mean, when you have non-randomized as much as you try you're never going to have ... You're not comparing apples and apples because there's obvious going to be with differences with respect to the patient populations. You could try accounting for it in terms of making sure what you think you know as being different. As you know, it's what is it, the known knowns, the known unknowns, and the unknown and unknowns. I think it's one of those things, and it's going to be the stuff you don't know that gets you. If you want to have a clinic experiment, do a randomized study, and I think that's there basis are written. Avoid confounding and other bias, et cetera.

Kevin: [inaudible 00:39:42] good explanation as, okay these patients are sicker than the others and they got randomized.

Dinesh DeAlwis: Right. We looked at that very very carefully as you can imagine Kevin. We looked at that very carefully. Yes, we couldn't see anything obvious, right. Things emerged later with respect to PDL1 status, et cetera. Fundamentally we couldn't simply say this was really the basis of such a difference because it was almost double at least. Always when you start phase one you go into initial patient population then you go to a next one, that can change things. You go to different sites, there are a number of things that can happen. If you do it randomized, I think that's the best way. Sorry.

Kevin: My question was, in the segmentation of [inaudible 00:40:31]. In the slide where you showed the segmentation of the subjects you had approximately half of them in non solid tumors, so my question is how did you measure response and tumor size in that population?

Dinesh DeAlwis: Sorry, which-

Kevin: Where you segmented out the subjects. I think you showed that about a half of them were in non solid tumors, was that correct?

Dinesh DeAlwis: No, they were all ... Yeah, non responsive.

Kevin: Okay, thank you.

Dinesh DeAlwis: Yeah, it was non responsive lung cancer, sorry.

Kevin: For clarification.

Dinesh DeAlwis: Okay.

Mark Ratain: I am Mark Ratain from Chicago. How much of your very building clearance attributable to body weight?

Dinesh DeAlwis: Very little.

Mark Ratain: Why do weight based dosing? It just adds time and effort and I can't tell you how many oncologists get calls from pharmacy or nurses saying you miscalculated your dose or patients lost 5 kilos since you started this treatment or maybe gained 5 kilos and you got to recalculate. I don't know if you guys considered this or you could consider a label change to get rid of weight based dosing.

Dinesh DeAlwis: Okay. The short answer is we have and we're going with the fixed dose. I think Mark you make a very excellent point. There has been excellent review about a year back looking at, I think over 20, 25 monoclonal antibodies and assessing the impact of weight and I think the exponential clearances, roughly in the range of about 0.4 and 0.6 or something. That tells you that really weight based dosing is not really important for monoclonal antibodies. It's one of those remnants that we've left with ... That happens with monoclonal dosing. Yeah.

Geoffrey Kim: I think this will be your last question.

Dinesh DeAlwis: Okay.

Sandra Girish: I'm Sandra Girish from Genentech. That was a really nice talk, thank you so much for sharing. My question was a little bit around the intra-patient dose escalation that you have in your 12 patient cohort for PK-PD. Personally I'm a little intrigued by the fact that you're doing that in an immuno modulatory setting. Do you worry about any immuno-modulatory responses that might have happened with the lower dose and escalating it to a higher patient. A higher dose in the same patient and then trying to tease that out from the lower dose effects. If you could please comment on that, that would be greatly helpful.

Dinesh DeAlwis: Yeah. With respect to Pembrolizumab, we obviously didn't have an issue with this. We could escalate intra-patient without having these, you know, your issue of

immuno-modulatory effects having an impact on a higher dose. From a general perspective, yeah that's a fair point but I think if you're starting at these very low doses and you're escalating, remember you're almost characterizing TMDD, right? You're getting an understand of TMDD. That can be an issue if you have a significant ADA that's confounding this, then can be difficult to interpret, right? That's important and in some case where immuno-modulation may actually cause ADA, then that can confuse the picture. In general, if you're starting with very low doses, quantifying the PK-PD relationship in this way, I think is the way to go.

It could be a different issue if you're starting at a much higher dose and then escalating up. There can be then issues around, perhaps what you may get. This was an antagonist of the system as well, we probably didn't fear that was going to be an issue here, okay? Thank you.

Shruti Agrawal:

Hi, good morning everyone. Am I audible? Maybe I should ask am I visible here. First of all I would like to thank organizing committee to provide an opportunity to discuss the case study of Nivolumab. My name is Shruti Agrawal from Bristol-Myers Squibb. I have been involved in Nivolumab right from it's early clinical development and I hope to cover ... After a very presentation of Dinesh, I hope to cover a very different perspective for the dose selection of very similar against the same target here. I will cover Nivolumab as a case study for cancer immunotherapy dose selection. What were the challenges, in retrospective I will again try to cover real time as much as possible. Lessons learnt, there was a very interesting phenomenal that we observed and opportunities for cancer immunotherapy and the cancer dose selection and the development in general.

The outline of my talk will cover what is cancer immunotherapy, case study of Nivolumab immunotherapy dose selection across tumor types. Along the way, we did the integrated exposure response and analysis for safety, efficacy and pharmacodynamic markers. I will try to touch upon what we are understanding in the dose selection for solid tumors. Then in the process we made an interesting observation that disease state in fact is having effect in peaking of monoclonal antibodies. This will have implication, not just on cancer immunotherapy, not just on anti PD1 target. It could have implication across monoclonal antibodies for cancer agents. Also then I will try to cover briefly what are the translational approaches, futuristic translational approaches to accelerate cancer immunotherapy development and some of the future directions that still need to be addressed also.

Let me start with what is cancer immunotherapy? Of course it is now a well established fact that cancer immunotherapy is a type of treatment that we design to boost the body's natural defense to fight the cancer. On the left hand side I have shown one of the survival plot from phase 3 Nivolumab trial in first lane melanoma that this one is actually a standard of care and the one here is Nivolumab and we can see there is significant increase in survival for a maximum number of patients. On the left hand side, those are typical splatter plots that we see in the clinical development. The small for the, a handful of patients from only clinical

development in melanoma. I would like to point out some of the distinguishing factors for cancer immunotherapy in spite of plots that we also compared to targeted agents of chemotherapy.

Once we actually achieve some tumor shrinkage, even if it is in the range of stable disease, there is a potential to have a durability in the tumor reduction there. That is actually a distant compared to targeted agents. In targeted agents over a period of time we might expect the growth of tumors. That may also lead to ultimately the decreased survival here. One more point that is unique to cancer immunotherapy is even if there is new Ligand, there could be reduction in overall target burden despite of appearance of new Ligand as shown here in this particular patient. Also there could be phenomena of the disease progression, that patient initially might have some increase and then decrease. Let's see how this also affects when we select efficacy measures for overall dose selection.

Let me spend a little bit of time on the mechanism of action for Nivolumab. It's similar to Pembrolizumab, it's anti PD1 inhibitor. The one in the center is T-cell, the exhausted T-cell. Exhausted T-cell does express check points such as PD1 which are inhibitory in mechanism. The tumor cells are the antigen presenting cells expressed ligands such as PDL1 or PDL2. The interaction of PDL1 and PD1 on T-cells lead to the exhaustion of this T-cell and the patient becomes immunosuppressant. Agents such as Nivolumab bind to PD1 on T-cell thereby inhibiting the interaction of PDL1 and PD1 and also that leads to the release of immunostimulatory cytokines. Over a period of time T-cells becomes activated and the proliferation of activated T-cells boost the immune system and leads to their prolonged survival.

What is Nivolumab? It's a fully human monoclonal anti IgG antibody targeted against PD1 on T-cells. It binds on T-cells at a very high affinity in low nanomolar ranges both to human as well as monkey PD1 target. It does not elicit ADCC and it has shown to increase cytokine production and also proliferation of T-cells. In animal models, it had shown to have anti-tumor activity both in PDL1 positive and PDL1 negative pre-clinical models. Little bit about the overview of Nivolumab development, and I am giving the example of melanoma development here because that was the first approval of Nivolumab. The IND started back in 2006 with first in human single ascending dose study. The multiple ascending dose study started in 2008 and the phase 3 started in 2012. Randomized phase 3 studies.

However, what I would like to focus here more is the duration for a multiple ascending dose study that was taken compared to phase 3 trials. This actually duration focuses on a dose optimization, dose ranging trials across tumor types as well as coming up with the safety management algorithms for cancer immunotherapy. First in human studies started as a single ascending dose study with alliance of subsequent doses if patients are benefiting every 12 weeks, which are really sparse dosing here. The dose selection for the first in human was based on allometric scaling. It was from allometric scaling we could see that doses anywhere between 1 to 10 mg per kg, could be efficacious in human. The starting dose of 0.3mg/kg was selected based on the toxicology data. The first human trial

was actually 0.3 to 10ml/kg. The tumors were both high immunogenic as well as low immunogenic tumor types.

High immunogenic tumor types were represented by melanoma RCC and low immunogenic tumor type was represented by non screamers lung. Of course screamers lung was also there. It was highly refractory patient population. However, even in that setting we could see some of the encouraging activity very oddly on. There was one PR and 15 carcinoma, one CR in colorectal cancer and SD in melanoma. That led to the multiple ascending dose trial. Again, I would say we started also with a traditional approach of 3 plus 3 design here. The multiple ascending dose study was 1, 3 and 10ml/kg. 3 plus 3, very traditional oncology trial. The MTD was supposed to be wither highest tested dose or the highest tolerated dose. The expansion cohorts were in melanoma and RCC at the highest study dose. The patients were scanned every 28 days for the tumor size.

However, after the encouraging activity, we really didn't want to continue with the traditional approach of oncology dose selection. With the unique potential of cancer immunotherapy that we're not targeting directly at the tumor cell, we are in fact boosting the immune system by targeting the peripheral PD1. There is a unique potential that we could select one dose across tumor types. With that thinking, that mechanism of action led to multiple amendments and the expansion of cohorts that we wanted to have dose ranging and at least 2 to 3 tumor types ranging from low immunogenicity and high immunogenicity for the tumors. Melanoma, we already had 1, 3 and 10 ml/kg and because already there was activity, we also wanted to go below these already clinical studied dose so we actually started with 0.1 and 0.3 also. However, there was a possibility for patients to get 1ml/kg if they are not responding, so I may not focus much on 0.1 and 0.3 because of the confounding effect on the tumor shrinkage there.

For screamers and non-screamers we expanded trial to 1, 3 and 10ml/kg and for ours we already had 10 so we actually expanded to 1ml/kg. Overall we have 3 tumor types. Four, including 2 histologies of lung and those ranging trial across tumor types here. The approach for Phase 2 dose selection was in fact a very integrated approach. We had total 306 patients for safety as well as similar population for efficacy across tumor types. We focused onto observed safety and efficacy along with integrated quantitative analysis for dose, and exposure-response relationship for pharmacodynamic markers, for safety and for efficacy. We actually focused on a long term relation as well because we wanted to make sure that the selected dose actually leads to long-term survival. We had minimum fall-off for efficacy population kept at least at 8 months.

For safety we focused on AE profile, grade 3-4 AEs, SAEs and AEs of special interest, and of course dose-response relationship for multiple efficacy endpoints. The early endpoint was objective response rate measured by resist criteria. We also looked for higher resist but focused more onto confounded resist responses. PEER Phase at 24 weeks, PEER Phase at 6 months was selected as a durability of the response endpoint along with the objective response rate. The exposure response

relationship was also done for multiple endpoints. Receptor occupancy was selected as a major for target engagement. We also had activated T-cells and the interferon gamma release available. Of course, population peaking, which was standard for this monoclonal antibody with a half-life of 26 days.

Exposure response was done for multiple endpoints, specifically, again I was saying earlier, endpoint the durability of the response and tumor growth dynamics. Which is not impacted by either follow-up or the bias of at what time the measurement is done for the response rate, and the dose safety for the probability of the high grade advancements. This is the dose information from 306 patients. The duration of therapy across those levels were similar, approximately 16 weeks. Patients received around 8 to 10 infusions on an average, but I would like to focus onto dose intensity per patient per 2 weeks. It was fairly very similar to the intended dose that patients were supposed to receive, patient at one mg per kg received one mg per kg every 2 weeks. Patient at 10mg per kg also received 9.8 on average per 2 weeks so there were not many dose reductions. The dose intensity was more than 90% across all dose levels.

Along with the DLT period we knew that DLT period only was 28 days and is not a reflector of the long-term safety so we in fact focused on the long-term safety with the minimum 8 months follow up. The MTD was not reached for Nivolumab and hence MTD was either 10mg per kg 2 week or above that. If you see any great events, grade 3-4 events, serious adverse events and AEs leading to discontinuation, they're very similar. It's like 71-81% here for 1-3 mg per kg. Grade 3-4 events were also very similar. All the AEs were manageable and reversible with the use of immunosuppressants. Along with all causality, drug related AEs and drug related SAEs were also very similar across those levels. Of course, there were few events, death early on in the program, that were associated with pneumonitis but over a period of time, a very robust safety management algorithm was created for cancer immunotherapy.

This is the exposure response relationship. There is time to event analysis for the multiple dose levels that we have studied and of course probability of the adverse events were a little lower for 0.1 and 0.3mg per kg. However, very similar if you could see it's overlapping over a period of 15 months for 1, 3 and 10mg per kg dose levels. That's true for treatment related AEs as well as AEs leading to discontinuation. That led to ... Even if we give 10mg per kg to 2 weeks for a prolonged period of time, Nivolumab was considered safe and tolerable from the odds of safety, as well as the probability of experiencing adverse events.

Let's focus on what was the activity and the durability across tumor types. This was a very demarcating thing for cancer immunotherapy that we have observed. For melanoma and RCC, which are supposed to be highly immunogenic tumor types, Nivolumab was very active. Both in terms of the objective response rate, as well as PFS and we could see that there is similar objective response rate and PFS. Of course numbers are limited, and there was high confidence interval and PDL1 expression status also might affect here. Overall, there was activity across those

levels, starting from 1mg per kg for melanoma as well as for RCC. However, for lung there was indication that 1mg per kg is not very active compared to 3mg per kg and 10mg per kg. We could see that the response rate as well as 6 months PFS is lower compared to 3 and 10. We also wanted to figure out why and how and how does it impact dose selection across tumor types. This is just an example.

Apart from the objective response rate we could also see the durability and the depth of response. The maximum tumor shrinkage was actually achieved anywhere between 16-24 weeks. Once there was maximum tumor shrinkage there was a very good durability of response. This is the example for melanoma. However, for lung and RCC also we have seen similarly in these plots. Let's see, what did we see in terms of the exposure response relationships? Peripheral receptor occupancy was selected as a measure for target engagement. Receptor occupancy at 8 weeks after 4 doses of Nivolumab in melanoma patients, and we could see that even at 0.3mg per kg the peripheral receptor occupancy was very saturated. These are exposure response relationships done by the Emax modeling concentration receptor occupancy. For the exposures which are leading to close to 0.3mg per kg, again we have seen a saturation in the peripheral receptor occupancy with the Emax of 71%. This was because of the frozen PBMCs.

Later we did collect data with fresh PBMCs and we did achieve 100% of the receptor occupancy. The EC50 was very low with the EC90, again was in the range of 0.3 milligram per kilogram dose. However, when we have more data from the later biomedical trials from RCC as well as for lung, this did not demarcate between what we were observing for the observed activity across tumor types. That may be because at that time limited understanding about how a peripheral receptor occupancy downstream is making an impact in tumor microenvironment. There was little bit of gap in understanding and maybe peripheral receptor occupancy or activated T-cells or the interferon gamma release for one tumor type did not fully explain the differences in activity across multiple tumor types. Hence was limited in terms of utilizing it for the dose selection.

These are the exposure response for efficacy for the probability of achieving any response. On X axis we have steady state Cmin. We selected steady state Cmin as this is a most conservative pK measure. We had similar analysis done for all measures of Cmax as well as the AUC and similar transfers. One on left hand side is melanoma, then in the middle, lung and RCC. This line is actually the probability of achieving the response rate with the 95th percent prediction interval. The horizontal lines are actually the mean Cmin concentration along with the observed activity that we were observing. Based on the trends we could see that melanoma exposure response relationship is very steep and is actually at plateau at the lower concentration. However, there is chance that for lung, the higher exposure is required for achieving the higher response rate.

For RCC we did not actually pay a lot of attention because of the gap in the dose response rate. We only had 2 dose levels and then of course it is expected to be linear. However, one point I want to point out here is, we could see the observed

response rate on the higher side here for 3, 10. Lower even for melanoma. However, we knew that there is something going on here, it's not just the exposure that is impacting the exposure response relationship. Of course there are other prognostic co-variants like ECOG, LDH, that need to be taken into consideration when we do exposure response relationships. Apart from that there was something that was going on that observation was made, which I will touch upon later.

In brief, for this slide, melanoma exposure response was steep, however higher concentrations were required for lung. Same patterns were observed for PFS at 24 weeks, which is actually a measure of durability of the response. This is the exposure tumor growth dynamic modeling done for all three tumor types, melanoma, lung and RCC, which was in fact extended to cancer immunotherapy from a model proposed by FDA. It was well established that for cancer, for the targeted agents, tumor growth dynamic modeling works very well. We have extended that to cancer immunotherapy. This is actually drilling down to the cellular level. That's what actually is happening when it comes to concentration effects. For melanoma, lung and RCC, the exposure response for the shrinkage rate was pretty flat. As long as there is some Nivolumab available in the system, it is able to shrink the tumor.

However, the demarcation was seen in progression rate. I would focus melanoma and lung more compared to RCC because we had gaps in 3mg per kg data there. For melanoma you could see that the shrinkage rate is decreasing very fast at the very low exposure here. However, for lung higher exposures are required, which are actually falling approximately at 3mg per kg to plateau the progression rate for lung. This was actually demarcating for us, to select 3mg per kg 2 week as immunotherapy dose selection across tumor types. Of course we also considered Pidalvan expression. However, data was limited and the Pidalvan expression continuum was still to be defined. Our goal was to select a dose which would be useful across the patient population, across multiple solid tumors as well as different flavors of Pidalvan expression.

In summary, for the dose selection part, peripheral receptor occupancy, because of the limited understanding that time did not provide meaningful differentiation of the observed activity, 10mg per kg 2 week was safe and tolerable based on the observed activity. Integrated exposure response relationship for various efficacy endpoints for early measure, durability of response and mechanistic tumor growth dynamic modeling indicated that higher Nivolumab doses or exposure are required for low immunogenic tumor types. Based on this totality of the evidence, based on the observed data, based on the exposure response relationships, Nivolumab 3mg per kg 2 week was selected for immunotherapy dose for further evaluation in our clinical program. This dose was shown to provide long-term survival benefit across multiple tumor types. In fact, 5 year survival from this study was recently presented at ASCO, and there was similar survival for 1, 3 and 10mg per kg for melanoma. However, survival of 1mg per kg, although very active was less compared to 3 and 10mg per kg.

I did allude to the fact that there was confronting effect going on in the exposure response relationship. On the one side, for the probability of the response, it's the same slide, but shown differently here. Instead of the exposure, these are actually what the patients received. Individual patients received response or not. Responders are on the upper side and non-responders are on the lower side. This is for melanoma, lung and RCC. What I want to show here is, there was an interesting observation if you see. Maybe I would focus on RCC first, that it's very easy to observe. Responders by default, they're on the higher side of the exposure range compared to non-responders. That was across any dose level. Say for example, the exposure of non-responder for 10mg per kg is way higher to the responder here. Still it's non-responder and within this dose level, responders are always on the higher side compared to non-responders. This was actually at all those levels and in all tumor types. We were actually wondering what is going on here.

This visual observation then led to very stimulating discussions and I would definitely like to thank Ahmed to have multiple rounds of discussions and coming up with such illustrative figures here. The data presented here is probability of response rate for Cmin, however plotted now within the dose level. You see this is 1mg per kg, 3mg per kg and 10mg per kg for melanoma. We are not showing 0.1 and 0.3 because patient didn't receive 1mg per kg there. You see there is distinct exposure response relationship within those levels. If you have very limited data, there is a possibility of misinterpreting that higher exposures are required to achieve the response. However, in fact, it's if you plot it differently, the probability by clearance now, you would see that those are parallel also. Of course there is this funny trend because there are other confounding effects which are not taken into consideration like ECOG and LDH.

There is very well a phenomenon that it's actually clearance also along with the exposure is playing a role in exposure response relationship. This needs to be monitored very carefully. The leading hypothesis here is, at the tumor side, our patients who are really progressed, there may be higher proteolytic activity by default. That's the reason we see patients are losing their body mass very rapidly. It's inherent proteolytic enzymes are increased so much and that also affects the endogenous protein as well as proteins such as anti-pidialvan or the monoclonal antibodies. When it comes to monoclonal antibody exposure response relationship, the confounding effect of disease status needs to be evaluated carefully along with the exposure response. We need to have data at least from 2-3 dose levels to demarcate these. This was actually lessons learnt that can be applicable to monoclonal antibodies for the cancer development.

Now that we have a lot of data from multiple check point inhibitors across even Phase 3 data, we also wanted to see how can we further utilize the comprehensive database to accelerate the cancer immunotherapy or in general cancer drug development. This is actually leveraging the totality of the immuno oncology data. This graph shows data from 1,500 patients, again further expanding the work that FDA did on the tumor growth dynamic modeling. This slide shows tumor growth dynamic modeling. However, to show whether the early tumor shrinkage can be

predicted for the cancer immunotherapy as well. The one on the left hand side is the prognostic variable only, not including the early measure of exposure. The smooth lines are the predicted lines for the various stress groups. The dotted lines are the observed survival here.

We could see that there is not a very good agreement between various stress groups for the observed versus predicted if we only use the stress group. However, if we are including early measures such as maximum tumor shrinkage or week 8 tumor shrinkage, there is a very good agreement between the various stress groups when it comes to the survival prediction. With this, of course, this is a validation that even for cancer immunotherapy, week 8 tumor shrinkage can be utilized for the early dose selection or the early decision making in the drug development.

Another approach also, is very interesting, is utilizing IO systems pharmacology models to predict the immunotherapy dose or selection for the combinations for the efficacy. The one shown here is for melanoma, that it is under development at BMS. Earlier I did not mention that pharmacodynamic data was not very meaningful. However, if we consider the entire system's biology, entire comprehensive mechanistic modeling, definitely it might lead to utilization of the entire knowledge for accelerating cancer immunotherapy development. This particular model has various compartments as blood, or plasma transport in tumors and lungs along with the multiple cells and the various mechanisms. The one example shown here is only one virtual patient as of now and it's ligand and the response to the combination immunotherapy. If we don't give anything, untreated, you could see that there would be an increase in tumor. However, with therapy A, although there would be increase in tumor, still there is a reduction there.

With therapy B, which may be more potent or maybe this particular checkpoint may be expressed on the tumor cells of this particular patient you would see a stable disease type of response, however, with combination or the higher dose of the therapy that may be for the reduction. This may also lead to the individualization of the cancer immunotherapy as well. This work is in progress at BMS. However, our work is not done here yet. Although we have encouraging Phase 3 data from multiple compounds now, Nivolumab, Pembrolizumab and Natalizumab now, there is a lot needs to be done for cancer immunotherapy, especially in terms of establishing the optimal regimen. Either sequenced or concurring depending on the mechanism of action, either as a monotherapy or combination for the patient selection.

Dosing frequency also needs to be optimized, duration of the treatment also could be one of the important factors that clinicians may be interested in. Of course, work is further ongoing for triple combinations or combinations with multiple treatment modalities to aggressively attack the cancer early on right after the diagnosis. Those are some of the future directions that we need to actually think through. With all that, I would certainly like to acknowledge all the patients enrolled in clinical trials. Without their consent it's very hard to further advance the

science. Clinical pharmacology and pharmacometrics group for stimulating discussions especially for demarcating the confounding effect of clearance on the exposure response relationship and of course Nivolumab and ipilimumab clinical development. Thank you.

Male: Thanks. A very nice talk, very stimulating. I'm very intrigued by the probability of objective response versus C_{min}, steady state, you have stratification by dose, but on the other hand this clearance show they are aligned across there. Clearance, essentially is the AOC, right? The probability of that response over AOC you see they are aligned very well. That indicates that the matrix of the exposure is very important, is that right?

Shruti Agrawal: Yes. I would say that both matrix are very important. Exposure of course to tease out the exposure response relationship, but we need to tease out the confounding effect of clearance. If you have only one dose level, then yes, clearance. Or if you have narrow dose level, you have clearance and the exposure highly correlated. If you have wide exposure or the wide dose level, the correlation between clearance and exposure goes away, and that's what you need actually to tease out the difference.

Male: What I'm trying to see, disease status, you say, affects the exposure response? Actually, You may explain by AOC change.

Shruti Agrawal: Not fully.

Male: Clearance as you say, AOCs indicate the failures, right?

Shruti Agrawal: AOC overdose right? If you have narrow dose, then it's highly correlated. If you have wide dose, then there is no correlation there. The correlation coefficient would be very low there.

Male: Okay. Thanks.

Female: Very comprehensive insight for presentation, thank you.

Shruti Agrawal: Thank you.

Female: It is very interesting to see you guys go back with multiple amendments to test more dose levels, especially testing the low doses in melanoma for 0.1, 0.3 and 1. Regarding that, I have 2 questions. One, is there any ethical concern potentially to go back to patients to test very low doses, where you know there may be sub-optimal efficacy? That's a common question we get all the time with clinical teams obviously. Another question I have, given these different dose levels were tested at different times, different cohorts, is there any combinability concern, or confoundings because these data from different doses coming from different cohorts and different patient groups, different parts of the study?

Shruti Agrawal: Right, to address the first part of the question, yes there were a lot of discussion with the investigators when we wanted to explore 0.1 and 0.3 dose levels. There were some concerns whether those would be efficacious doses or not, especially when the elementary scaling was showing the efficacious dose of 1-10mg per kg. That's the reason patients were allowed 1mg per kg escalation within the patient if they are not responding earlier on. That was first part. The second part, of course, there were limited sample sizes earlier on, and there were effect of confounding factors as well. We went back exploring various dose levels across tumor types. However, the understanding of the exposure response relationship matured with the availability of additional dose ranging trials. We had dose ranging trials in RCC as well as Phase 3 data contributed and further solidifying the understanding for Nivolumab.

Female: Thank you.

Dinesh DeAlwis: Going back to the issue around disease affecting the pK, can I just clarify something? The analysis was done with 1, 3 and 10 but they weren't actually included in the study, right? They were actually different studies-

Shruti Agrawal: No. This is only from one study. What I showed here is only the 003 study data. However, we did a similar analysis from RCC dose ranging trials, which was in fact a randomized trial with 60 patients. A very similar finding was seen. Again, lung also, but we only had 3mg per kg as a Phase 3 data and there is some effect there as well.

Dinesh DeAlwis: Okay. The factories that there are obviously, for any drug is on board, there are patients who are going to have a higher catabolism versus a lower catabolism. That's going to be reflected in clearance. That's got nothing to do with drug. Then you've got to actually assess how it actually may have disease modifying over the course of time. There are 2 fundamentally separate things that you've got to de-convolute when you do this analysis.

Shruti Agrawal: Of course. The full analysis is not shown here because we actually wanted to focus on the dose selection part. However, a very comprehensive analysis that actually incorporates, not just the clearance across the dose range but also including the other prognostic covariants as well. That still shows that clearance by default may have some implication in the dose selection then.

Dinesh DeAlwis: That's fine. Clearance, not exposure though. Basically what I'm trying to say is clearance just happens to be a good surrogate for what's going on.

Shruti Agrawal: Yes. Good surrogate for something that we don't know.

Dinesh DeAlwis: That can be down to albumin changes, weight changes, any other factor. I just want to kind of delineate, it's not an exposure driven phenomenon. Thank you.

Shruti Agrawal: Right, it's not exposure. You're very right. One point to your question Dinesh, that,

we have in our model, we also have weight, albumin, the ECOG for RCC. I would say that the clearance is not fully accounted by these already known covariants. That's something which is reflective of the protein catabolism that is clearance, we might consider as surrogate here, but not actually fully taken into consideration by other covariants.

Male: I have a question about your receptor occupancy at a very high dose of 10mpk, receptor occupancy was around 60 or 70%. Can you explain that because, would you expect PD1 to be that highly expressed or that highly turning over, that even 10mpk dose cannot saturate at 200%?

Shruti Agrawal: Yes. One of the limitations of the data that was shown here was, data was with frozen PBMCs earlier on in the program. The maximum occupancy that we have noticed was approximately 70%. Within the variability range 10mg per kg also was in the range of 60-70%. However, later clinical trials were done with fresh PBMCs, again, it was highly saturated up to even at 10mg per kg it is 100% receptor occupancy. It was a data collection limitation I would say, earlier on.

Male: Thank you.

Shruti Agrawal: Right. Thank you all.

Chyi-Hung Hsu: Good morning everyone. My name is Chyi-Hung Hsu, I'm a modeler. I'm also a statistician. Today I'm going to provide some information regarding the potential benefits of exposure response, overdose response, this is the subject we are going to discuss in this session. Before I start I'd like to thank the organizers for this opportunity. Thank you everyone for coming as well. The talk consists of 4 components starting with motivation. Then the modeling framework which ties the exposure response and dose response together. Then we'll have some simulation and finally will be some comparison among these exposure response models and there's some conclusion regarding that.

The motivation, I think, drug development ... I think we all agree that it is very important to have a good understanding of those responses and also how to pick the right doses for development. We do have some difficulties. Particularly, for example, the number of doses that can be investigated may be very limited. Also the sample size, for example may be sufficient for the hypothesis testing but might not be sufficient for the estimations. Therefore, the alternative exposure response which takes into account the pK inter-subject variabilities should be a very viable alternative for dose response. The question we are going to answer is, how can we quantify the potential benefit of the exposure response modeling? This is the whole subject area we are going to discuss.

I'm going to start with the modeling framework I'm going to use a very over-simplified model compared to the previous speakers. I think it's over simplified. The first one, exposure response model I'm using here is the true model which tied exposure and the response together. This is how the data was related with

exposure. We are going to use the sigmoid Emax model in AUC. Before you all is the average PD response E_0 placebo response maximum effect. EC_{50} is 50% of E_{max} . H is here coefficient. AUC will be equal to dose over current. A couple things I'd like to address here is, in this true underlying exposure response model, the occurrence we use here is the true occurrence, which might or might not be observed during the time of the experiment.

The second one is, another thing I want to emphasize is, EC_{50} . EC_{50} stands for all subjects. As long as the subject reaches the EC_{50} , then we know that 50% of the E_{max} will be reached regardless. There's no subject specific EC_{50} in this model. In this particular model, there are several sources of the variabilities, which we probably need to consider. The first one is PD variability. If we go to this graph, what the PD variability is trying to do is giving same exposure. This graph, the y-axis is a response. X-axis is AUC. The thing PD variability is going to describe, why given the same exposure we're going to have the different PD response. The PD variability, the notation used is called the σ_y . In the numerical term this is dollar sign sigma.

The next thing is pK intrinsic variability. In this particular example we're going to use, we're going to focus on the occurrence, which means different subjects have different occurrence. That's one demonstration. The distribution of occurrence for this population. For example, the variability associated with a pK intrinsic variability are going to use the notation σ_{CL} . CL represents occurrence. This 2 pieces of variability I think that we're probably very familiar with. The third part of the variability, sometimes we might or might not pay attention to is, the pK measurement variability. Where is this thing coming from? In practice, the occurrence for the model available to us is probably measured with error. There is also an estimate. It's not the true occurrence.

However, they are related to the true occurrence. What is the main thing is the observed occurrence basically centered on the true occurrence, with some variability called the σ_u in this case. In this model, we have 3 different sources of variability, PD variability, pK intrinsic variability and the PD measurement error or pK measurement variability. The importance of the pK measurement variability, they give us a different perspective from the modeling perspective. Then we look at data will be very different with or without error. Let's see. Right now there are 3 panels in this graph. The y-axis is the response. The x-axis is observed AUC, which is observed exposure. Let's say σ_u is a measurement error. Suppose there's no measurement error, from the model perspective while we see the observed exposure over the response. This, while we anticipate, what will they look like? At this moment, observed exposure is equal to the actual or the true exposure.

However, when we have the measurement error, for example ... By the way, the grey band, 90% of the population will be falling into this grey band. However, if we look at the measurement error, we have 20% measurement error, then we're going to introduce more variability. The reason with the more variability is, even we observe for example with AUC equal to 5, there's mixture of responses from

different true AUC's. Maybe it's from true AUC of 6 or true AUC of 4, because of that, we increase the variability. This is the grey band, the darker band, basically it's measurement error. This will increase the variability. What happens if we have 40%, they're even bigger. From the perspective of a modeler, we see the data. We can maybe look at this, maybe the middle one or maybe the last one. This totally depends on how precise our pK measurements are. They have different shape, if you look at this. Probably they won't have some impact on the parameter estimate during our modeling exercise.

The next thing we are going to look is, how can we relate exposure response to the dose response in order to compare these two approaches? Starting with original exposure response model, then we just replace AUC by the dose over occurrence. This will be this part. Then the modified occurrence after simple mathematical calculations, we are going to go to the right hand side. This will be the dose response model. In this model, ED50I, there is ED50 we know. There is the dose, which 50% of the maximum effect is attained, but because of the mathematics we apply, ET50I will be equal to occurrence I. Occurrence I is the individual level occurrence tied to EC50, which means that the ED50 right now is subject specific, which means that they will be able to accommodate the inter-subject pK variability. At this moment, these 2 models are equivalent.

Dose response-exposure response, at this moment is equivalent, provided that we are able to estimate ED50I, EEO Emax H as before. Another thing is, if we look at the distribution of ED50, the variability of the ED50I is basically controlled by the intrinsic pK variability, which means that the ED50I basically absorbs all the intrinsic pK variability. At this moment, these two models are the same. They are equivalent. There's no ambiguity, we all agree. The problem is, in practice, the ED50 cannot be estimate. In order to estimate that, each subject has multiple doses in the experiment, unless in the crossover. Typically, they are not estimable. For necessity, we have to make the ED50 same for our subject. Because of that, those responses, exposure responses will be different. The dose response is no longer able to accommodate the inter-subject pK variability at this moment, so these 2 models become different.

At the same time, in practice, we also know the exposure available to us is not the true exposure. It's something we observe or estimate from the model. Both of them probably will derail from the model we are sure earlier. Our exercise, trying to see, what happens about this two approaches? One is better than the other? How much benefit from one model over the other? The simulation's started. We are going to have a design of five parameters, the placebo plus 25, 50, 70, 100 milligrams. We are going to assume 30 subject parameters, then we'll consider 120 scenarios. There are 4 sigmoid Emax exposure response models. This is true underlying model, so at least here. The PD variability where 2, 10, 20%. Intrinsic pK variability sigma u 30%, 50% and 70%. pK measurement variability 0. That means there was no measurement error. Then with 20, 40, 60 and 80%. This is the simulation study we have.

The exposure response model considered looks like this. Some of them very steep, some of them very short EC50 that reached the steady state PD very quickly. There's the 4 exposure response models we have. Let me give a little bit of detail about simulation studies. How do we do the simulation? There will be 1,000 trials for each scenario. How do we do that? Starting with simulated PD response under a scenario, the first step, we're going to simulate the true individual occurrence which is the true driver for our PD response. The second one is, we're going to simulate the observed occurrence by just adding a little bit the measurement error to the true individual occurrence simulated in the step 1. The third one, we are going to simulate the true PD response according to the true occurrence.

In terms of model fitting we're going to use the basing approach. For the exposure response estimate, we're going to use the model we described earlier. However, only observed occurrence is available to us so there will be the information which will be the input we're going to use for the model fitting. For dose response, we're going to use population ED50. That means we're now able to account for the inter-subject pK variability in this model, so everybody will have the same ED50. Another thing important for this one is, because of focus on dose selection, I'm going to define how the dose will be selected. We call the MED, Minimum Effective Dose that consists of 2 components. The first component is DOTA. The clinical relevance in proven DOTA over placebo, I'm going to assume 60. In addition to that, I'm going to say, they're going to cover 80% of the population we are studying. We have two components, one is DOTA another is coverage, this is the MED we're going to have. I'm going to show some results.

The first graph we're going to see here, I'm going to see the impact of the measurement error on the estimates we have. Right now, for this particular example, sigma y is the PD variability 20%. Sigma u, this is intrinsic, pK variability is about 30%. MED on the y-axis, which is going to give us the summary of a thousand simulations. The band we have is 90% confidence interval. The blue band represents the median. The red line here is the true MED. On the x-axis will be different levels of the measurement error from 0, which means there's no measurement error up to 80%. As we know that dose responses have no measurement error because the input is just dose. From here we can see a few couple things in terms of the impound measurement error on the exposure response.

First of all, the measurement error increases the width of the confidence interval. The second one is, in terms of estimate, they become biased gradually. When the measurement error increases, it will create some bias. The third thing we can observe from here is, if we want to compare the dose response at the last one against the exposure response. If you compare the least one, from here we can see, if there was no measurement error, there's no brainer. I think we can see the exposure response is much better than the dose response. However, when we have 20% at this moment measurement error, these 2 are closer. I think the exposure response is probably still better. When we go somewhere between 20 and 40, I think the dose response can be preferred. This is the one graph to

demonstrate the impact of the measurement error on both estimates, and how do they compare with each other.

The next one, there will be 6 of them instead. Previously, we looked at this one bottom left. Right now I'm going to put all 6 together. What's the difference between all these 6? On the top row there will be σ_y which is the PD variability, 10%. The bottom row is 20%. For each column represents different levels of the intrinsic pK variability. First one, let's compare the top with the bottom. We see the σ_y , the PD variability from 10% to 20%, the variability increases as the interval increases. We know the impact on σ_y . The next thing is, let's look at what impact we will have when the σ_u intrinsic pK variability increases. That means we can compare row-wise. We can see, when the pK variability from 30% to 50%, the variability definitely is increasing. The same thing if we look at the dose response, the interval also increases as well.

Therefore, we know the σ_u intrinsic pK variability will also have the impact on the dose response as well. The larger the intrinsic pK variability, the larger the confidence interval we're going to have if we use the dose response. However, again, we see in general, the dose response under this scenario should be probably somewhere ... Probably it's not as good as when we have 20% measurement error in pK, but probably will be preferable if we have something between 20 or 40% pK measurement error from this exercise what we can see.

There's another one, model 2. This is much more profound, we can see in this case. If you can recall that previously, that the measurement errors depends where the MED is. That can be have a great impact because the confidence intervals become so wide. This maybe has some impact. In this example, I think the dose response is probably as good as those exposure response with 20% of the measurement error. Another thing I want to re-emphasize again is if you look at the bottom 3 panels when the σ_u from 30 is at 70%, you see how wide the confidence interval increases. This exactly reflects what we noticed earlier, that those response models are not capable of, at this moment, to address the inter-subject pK variability. They are now very robust. In order to address, the confidence intervals become so wide.

Then come to the conclusion. The conclusion based upon this simulation, I think we realize that 2 things are very important in terms of the performance of these 2 approaches exposure response and dose response. First of all, intrinsic pK variability, σ_u , as we said earlier, the current approach for those responses, this piece of variability was not able to address, typically. Therefore, the larger the pK intrinsic variability, the more likely to prefer exposure response. The other thing we're going to emphasize in this simulation, we call them measurement variabilities. This particular one, we see when the larger the measurement error we have, the less ... When the measurement error between 20% or even more than that probably we might prefer in that case in terms of precision bias, probably we are going to prefer a dose response model in that case.

In terms of improvement for these both approaches, as we said, the pK variability is

probably going to drive the performance of the dose response. Unfortunately, pK variability, for example, intrinsic currents, that is nature, nothing you can do. However, if you're going to improve the exposure response model, then we can maybe use a better methodology. For example, we can have a better PK-PD model or maybe better bioassay for example. Maybe make the measurement error probably smaller. In general, we do not know exactly how much error we have, that indicates perhaps we can have both models being performed like the previous speaker. Probably we can see the robustness of the approaches we have.

One last thing I'd like to mention there is, in this case, we know the model uncertainty. Sometimes this plays a very important role. There's some methodology that's been proposed in the dose response, but they also can be easily explained with the exposure response through the model average methodology. Finally, I want to make acknowledgement for those BOMEs here for their wonderful valuable contributions and the suggestions. There are two references, the first reference is for the model and the second reference probably is most of the slides coming from today. They give you more information about the 4 models, have been simulated and evaluated. Pretty much it, I'm going to conclude my presentation. Thank you so much.

Male: I'm not trained in Stats so my questions can be quite naive. Can you give some intuitive interpretation for when the variability is high, like as you talked about, in the PK. Why does it lead to biased estimate of MED?

Chyi-Hung Hsu: Okay, good point. I believe that depends on the definition of MED. If MED is defined just based upon the BOT, the minima, then probably there is no bias. Right now because we are going to cover 70% population, as you can recall, the variability becomes bigger. It's going to push either forward this way or the other way, so there's the second part. The variability creates the bias.

Male: In that case, the variability should always lead to higher dose because you need to cover wider variable, but that's not the case in some of the cases you had.

Chyi-Hung Hsu: Correct, it depends where you are. If you are at that part, see the very flat or the middle, then probably I totally agree with you. If you go the other part, then I'm not sure.

Male: Okay, thank you.

Chyi-Hung Hsu: Thank you. Yes?

Mark Ratain: Mark Ratain, Chicago. I think you've over simplified your model. You have assumed that the EC50 is constant for all patients when in fact it's not constant. In fact, you may have a situation specifically for monoclonals where the EC50 and the clearance are correlated. That means you may be drawing the wrong conclusions without those response, so you need to really consider a model that would allow you to look at patient covariants and have a patient specific EC50 to really have a robust

model. Have you considered that?

Chyi-Hung Hsu: I totally agree with you, but this exercise basically is very generic. Just wanted to get some idea. What's driving those responses? What's driving the exposure response? I totally agree with you, in practice you are absolutely correct. Okay, thanks again.

Geoffrey Kim: The time I have is 9:47. Let's take a 15 minute break and come back here at 10:02. [crosstalk 01:49:59].