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

Geoff Kim: Especially if you were the speaker of the next session. Still awaiting our next speaker. Is she here? I can't ad-lib ... As soon as anyone sees our speaker number four, please let me know. I could sincerely like to thank all the speakers, and especially for the first three speakers of the day. It kicked off to such a great start. Highlighting some very important topics that we see as an agency, that we are going to be faced with.

Right now, five years ago, I presented at a oncology drug advisor committee. At that presentation we presented exposure response analysis for PFS with a very poorly soluble drug with a very tricky pharmacokinetics characteristics. We presented this graph, and it clearly separated into four quartiles of progression for survival. Where the highest exposure was at the lowest quartile. We were certain that there was a negative correlation between exposure and PFS. Now, to see how the field has shifted in this methodology and interpretation, is really amazing. I think some of the caveats regarding the clearance about the dose exposure versus dose exposure response are really highlighting some of the necessary discussion we need to have to refine and build this field. The utility of this methodology for, not only drug discover, and drug selection, dose selection, but really for the characterization and optimization for those with cancer.

I think, really to ... Multidisciplinary effort. When I say multi-disciplines, not only are we talking about all the scientists involved, and you'll hear this in our discussion too, but really making these discoveries. How do we communicate it better. How do we make these discoveries and inferences, be able to translate to those people who don't readily ... aren't in this field on a day-to-day basis. When we talk loosely about PKPD exposure response, how does that translate to something that's going to really be impactful. How do we know that everybody knows we are talking about the same things and on the same level. Keep that in mind as we move into discussion, and speaker of course.

Diane Wang: Sorry, for being late. Thank you committee to invite me for this presentation. I'm Diane Wall. I'm the [inaudible 00:04:33] of oncology lead advisor, oncology. I heard several wonderful presentations, and also now, but I'm going to focus a little bit on the small molecule side. How to base the selection and the study design has been. Whether to use other therapeutic areas. However, the oncology area, this is a still area that we are exploring right now. For small molecule oncology compounds, MTD is still the most often used approach. The MTD determine in the phase once study often use as the recommended phase two dose. It is rare for phase two dose for two studies, to have more than one dose levels. I know, we just heard from Mark and and [Yemes 00:05:42], they have multiple dose levels in their programs, however, not all of us has that kind of luxury in our development program. As a result, these border response are really well-defined at the MTD selection stage. The inter-patient viability, the advocacy, and the safety is not adequately evaluated most of the time.

The first reason that is, and also often conducted in listed cancer patients with multiple tumor types, therefor the long-term cumulative toxicity is not adequately addressed due to the short life expectancy of these patients compared to phase three studies. Also, maybe the intended patient population. These factors, among others, contribute to the higher [inaudible 00:06:43] rate of the phase three oncology clinical trials. Six in all reported that, many recently approved oncology drugs label that the dose levels that may not be optimal. What are the challenges? Why we're in this situation within oncology therapeutic area.

The first, and the most fundamental reason is that there's a lack of predicted PK markers to guide the selection, especially for solid tumors. It is not always feasible to collect tumor [palpisies 00:07:29] to measure modulations. Sometimes there's no suitable essay for the PK marker for certain targets. As you can see, the procedure of the ISA may not be suitable for PKPD analysis. Sometimes P markers is not stable enough in the tissue so fresh biopsy is required to do the [mass 00:07:57] at the site, which may not be feasible.

The last, but more importantly, the correlation between PK marker and [effix 00:08:06] endpoints are not well-established. Therefor, the prediction is always in the question at that stage. This is especially the case for the first compound class, in class compound. From business and the patient's perspective, considering a medical need and competitive landscape, the need for developing the new drug quickly often precede the need of finding the right dose. Trials with multiple does levels would require a lot more time to patient enrollment, and also longer time to the advocacy endpoint read out. Of course, factor trials will cost more. With the limited budget, and multiple promising agents available in the development pipeline, sometimes is really hard to justify for the investment. However, as we now know, MTD approach was designed for the cytotoxic therapies. For [Cyto 00:09:20]therapy, the optimal logic dose might be more relevant. A comprehensive strategy, that includes those optimization, but not delayed market entry, is needed.

How are we going to do that? There's no one-size-fits-all approach for those actions. [inaudible 00:09:43] approach recommended has been to find biologically, the effective dose using the effect marker. If the effect marker could be a PT marker that measure target modulation, or a disease progression marker that, for example, measures tumor size change. There is examples of this approach, as reported in the literature. These are some examples listed here.

In my presentation today, I am going to discuss two case studies. Case one is the dose selection for the small molecule in an incubator that constantly marker modulation, exposed response relationship, and potential drug-drug interaction, and safety and efficacy data. The phase two is modified MTD approach for [remission 00:10:44] therapy by balloting the over lapping toxicity of the combination. Using the simulation approach and introducing growth factor during the DT termination period. We provide a modified MTD criteria to communicate the over lapping toxicity.

This is case one study. For both of these examples the compound name indication, and the target are masked for confidentiality reason. PFX is a small molecule inhibitor currently under development in phase two. It is evaluated in the hematological and the solid tumor patients in the phase one study. The first human study was conducted in the best hematological agency over a wide range of the dose. From five milligram to 600 milligram dose QD. The PD of the PFX is linear with the dose range study, and is mainly metabolized by Z3A. The early signs of advocacy in the wide dose range studies from ten milligram dose to 600 milligram dose. The response is observed from 22 patients out of 47. The protocol defined the criteria was not met. A 400 milligram dose was considered to be the MTD, since a safety signal was noted at the 600 milligram dose. A comparative analysis, [inaudible 00:12:33] was collected from 26 patients in three clinical trials. This is also in the wide range of 25 to 640 milligram dose. The [bell marker 00:12:44] biopsy was collected at both baseline and at a steady state to measure [bell mark 00:12:53] expression.

A quantitative real-time PCR was used to measure the expression of a transcription factor, which is a marker for the pass was activation and signaling. In the inhibition of sickling resulting in activation of the bio-marker and suppressing the expressant genes that mediate the tumor growth. The incubator data from PK, PD, CD and and advocacy was used for the selection to move forward to the late-stage development. This plot shows the [bell marker 00:13:31] trend from baseline versus steady state exposure as measured by AOC 24 hours at steady state. The different colors here represents the different dose levels. If you look at the red triangles here, these are the bell marker change for the 100 milligram dose. Basically, similar levels of [inaudible 00:13:59] modulation was observed at 100 milligram dose and above. There's no advantage from bell marker perspective to go higher in dose. This bell marker is also used for dose selection of other compound in the same class, however the correlation between the [inaudible 00:14:18] modulation and advocacy endpoint has not been established. Therefor, other factors are also evaluated to determine the dose moving forward for the output.

The time to those levels were considered within the team. 200 milligram dose, and 100 milligram dose. As I mentioned earlier, the safety signal was identified at 600 milligram dose. The simulation for the PK profile was conducted with the goal of having minimum overlapping between the dose selected to move forward, and the 600 milligram dose. This are the result of the simulation, which include the fifth to the negative fifth percentile interval of the predictions. As you can see, there are significant overlapping between these two dose levels in terms of the exposure. At the time, a clinical oncology also read out. It showed that [inaudible 00:15:22], which is a strong inhibitor of Z3A, increased the PFS mean by .4 fold, and the C max by 1.4 fold. Intended indication for this compound is a hematological cancer.

The use of a fungal drug, [inaudible 00:15:44], can come delay with the PFS is not avoidable to be able to measure the patient with hematological cancer. Therefor, with the potential drug-drug interaction between PFX and [Azzos 00:16:04], the overlapping between those two dose levels will be more significant. This results

shows the comparison of 100 milligram dose to 600 milligram dose. As you can see, there is no overlapping in C max, which is actually the driver for the safety finding between the 100 milligram dose and the 600 milligram dose. Even, if you consider the potential interaction between Azzos and PFX, and that they overlapping will be minimal. This is block plot, so the distance showed here actually is more than it appears in this plot. Still, in summary, 100 milligram dose was selected for further evaluation in phase two clinical trials, based on similar down regulation of the path of activation at dose levels within the 100. Also, the observed clinical advocacy across the wide dose range study, starting from the 600 milligram dose. The minimum overlapping between the 100 milligram dose and the 600 milligram dose was prime to AUC, which provided more confidence from the safety perspective. Of course, the lower dose level of all safety profile and therapeutic profile are also better for the compound.

The second case is an automated study design for the MDT determination in the combination therapy. Drug A was approved for the treatment of solid tumor X. Solid tumor X is a solid tumor that was very limited treatment options for patient. Therefor, there is a medical need. The approved dose for drug A is 300 milligram dose. Drug A actually is a chemotherapy. The drug B is a targeted therapy that is in phase one clinical trial as a mono-therapy. 15 milligram dose is determined to be the MTD. This is also the dose that we see some clinical advocacy with the mono-therapy. Pre-model showed some [inaudible 00:18:48] in solid tumor X when the B is in combination with A. Phase one B clinical trial is designed to evaluate the safety, Pk, and the early sign of the advocacy in patients with solid tumor.

However, for drug B, there is no available bell marker to measure the modulation of the target for those selection. Therefor, the traditional MDT approach is used to determine the recommended phase two dose. One of the biggest challenges of the combination therapy oncology is overlapping toxicity. Unfortunately, neutropenia was shown to be one of the major side effect for both drug A and drug B in their respective clinical trials. The combined therapy of drug A and drug B, believed to enhance neutropenia effect, limit the MTD dose for both compound. Therefor, potentially lead to stop therapeutic advocacy for the combination. A prediction of the potential neutropenia of the combination therapy may help better study design and neutropenia management.

Fortunately, the PKPD model for neutropenia were available for both compound. The neutropenia model for drug A is available in the literature since it's an already approved product. For drug B, neutropenia model were developed using the data obtained from the phase one clinical trials. For the predicted combination effect on neutropenia, the additive model was assumed for the drug A and drug B combination on the neutropenia reduction. Of course, if the effect is synergistic, the effect will be more profound. However, the additive effect was assumed as a base case scenario. This is the dose isolation scheme with combination therapy. The dose levels will start as the 200 milligram dose for A, which is one dose level down from the approved dose. For drug B, is 30 milligram dose, which is two levels down from the MTD and also the dose that we showed single agent activity.

Since we don't really know which drug can be a little bit more to the combination therapy in solid tumor X. There are two parallel isolation pathways. The one on the top part has the escalation of drug B by keeping drug A at the one dose level down at the starting dose level, while the escalation for drug B, I mean the isolation at the lower panel, prioritized the isolation for Drug A. The highest dose will be the center dose for drug A and also the MTD for drug B at 15.

For this dose evaluation, this is a traditional three by three design. The observation time is one cycle. During this cycle, [inaudible 00:22:33] use is not allowed. Also, if we observe 30% of the patient with the DLT, then that will see the MTD level. There are also ... If the drug cannot be tolerated at least those two drug levels, there will be further dose reduction from these two dose levels. Totally, there are eight possible dose levels that will be tested.

This plot shows the simulated neutropenia profile. On the right-hand side, the purple curve here is the neutropenia profile of drug B given alone. The blue curve here is the neutropenia profile when drug A is given alone. The black curve here is the profile for the combination of drug A and drug B. This is the highest dose level, which is level eight. On the left-hand level, it shows the neutropenia dose level across given dose levels. These are the mean neutropenia profiled for each dose level. This dotted lines here shows the neutropenia grade cut off. This .5 here is the cut off for grade four neutropenia. Between .5 and one is the neutropenia grade three. You may say, at the starting dose level, which is somewhere here, that we're not leaning too bad because we're hardly touching the grade three neutropenia. However, this is the in profile, and if you look at the distribution of the patient population, and then we'll see a different picture.

This is the simulated profile from 50th percentile to the 95th percentile of the neutropenia profile at dose level one. Dose level one is actually one level down from our setting dose. You can see there will be a significant portion of the patient that will get into grade three neutropenia. A few of them will be grade four. Grade four is considered a dose-limiting toxicity. This is a picture, it shows how it looks like for the dosage levels at different neutropenia grade. Vertical panels here are neutropenia grade one, two, three, and four. The horizontal panels are different dose levels. For example, at dose level one, it's most likely you see 15% of the patient has grade one neutropenia. 3% of the patients has grade two neutropenia. 3 to 5% of the patients has grade grade three neutropenia. About 10% of the patient may have grade four neutropenia. This is at the dose levels one level down before our setting dose.

If you look at the left panel here, at grade one neutropenia, with dose increase you will see a left shift of the distribution in that when dose increases, you will have fewer patients with grade one neutropenia. However, this is not really a good thing because the patient actually shift to the right. To the higher neutropenia grade. If you look a the grade four neutropenia, with dose increases the personal patient with four neutropenia will be increased. The shift is to the right-hand side. This a

numerical presentation of the mean percent of the patient in each neutropenia grade. Again, grade four neutropenia is considered as a DLT, and is also what triggered dose reduction. While grade three neutropenia is not a DLT, however you will need to have dose reduction for those patients. A starting dose level, as you can see, the neutropenia grade four will occur probably in 20% of the patient. Based on neutropenia alone, this will not be exceeding MTD. However, if you consider both grade three and grade four you will have 55% of the patient that will require dose reduction. Even this dose is not exceeding the MDT, this will not be a dose that we will like to move forward, because you don't want to have a dose that you require the majority of the patient have dose reduction in the phase three clinical trial, and also in the clinical practice trial after approval, if we can get approval of course.

As you know, remember, the center dose is 300 milligram dose for A and the clinical advocacy is shown for drug B at 50 milligram dose. If we have those reductions clinically a the lower dose levels, the concern is whether we can achieve meaningful advocacy for this combination. To address that to all concern, especially this is the tumor type that is a medical need, we decided to introduce these FFU's during the DLT evaluation time period. Instead of dose reduction to address the neutropenia issue, we introduced GSfF before the dose reduction to mitigate the neutropenia for the combination therapy. Hopefully, we'll boost the neutral field column at each dose level to maximize the potential for advocacy of this combination.

This is how it works. For example, now we are at dose level 4A. The collective data have shown that, based on the data from 3A and 4A, now we have six patients for example. If we see there are more than 67% of the patient that has grade three neutropenia. Although, we may not achieve MDT because the total rate of the DLT may not exceed the 33%, we will not proceed to the next level of 6A here, but instead we will repeat the dose level of 4A, but including GSFF during the DLT evaluation period. This is shown as the red boxes here. Once we switch to this right box, and then this isolation will follow the red box path. In order to observe the effect of GSFF, and will extend the DLT observation from one cycle to two cycles. By doing this, hopefully, we will be able to get those to the level to have the maximum chance to see the combined therapy with advocacy in this tumor type. In summary, there is a high probability of the dose reduction that will occur in majority of the patient at lower dose levels before we reach the MDT. Therefor, the GCFF use is introduced during the DLT evaluation period to maximize the potential of this combination therapy in tumor X.

We'll also looking at, of course, other alternative dose schedules to minimize neutropenia without GSFF use. These examples that we are presenting here, we show that ultimately the simulation approach is a useful tool in dose selection, and a more informed study design. However, it's much less elaborated and comprehensive in those that we have been normally doing in other therapeutic areas. However, I believe that with the involvement of the science and technology, and the collective effort between the agency and the industry, with the little steps

that we take right now, we're going to enter a new dose-finding era for oncology drugs that will provide patients with better beneficial ratio to the cancer patients. With that, I would like to acknowledge my colleagues who contributed to the presentation, Navid, who actually is in the audience right now, is the [inaudible 00:32:34] for the first case study. If you have tough questions, I will defer to Navid to answer them. [inaudible 00:32:47], who did the simulation studies for the second case study. Of course, [inaudible 00:32:53], who is our head of the clinical oncology. He is a big advocate of the dose-finding oncology. Unfortunately he could not be here due to there conflicts. With that I would like to thank you for your attention.

Mark Ratain: Mark Ratain, Chicago. I believe your first case study used skin biopsies as your bio-marker. I believe that's your [smoothant 00:33:30] inhibitor, you didn't say that. That doesn't make sense to me, that you would define that as your optimal dose. There's a drug called Iressa, that's something you may have heard about. That was how they chose the dose for that drug. I think this would be a good opportunity to learn from their mistake, that yes, it tells you you have a pharmacologically active dose, but you cannot conclude that's an optimal dose. In fact, the whole point of a targeted drug is to take advantage of differences between the tumor and normal tissue. You shouldn't base your dose on a bio-marker in normal tissue.

Diane Wang: I don't believe bio-marker is the only factor considered in this dose selection. Navid, do you want to comment on that?

Navid: Mark, you do have a good point. Maybe the term optimal is not the right term to use. Let's just say that. It's certainly ... The bio-marker that was used is a marker for the mechanism of action validation, not for efficacy itself. Which is typically the situation that you face for most drugs, and most composites you double up. As Diane pointed out, that was one of the factors allowed us to settle on a dose to move forward in addition to looking at a lot of different things different things, as she mentioned. Efficacy, safety- or I should say signs of efficacy, not true efficacy because it is a phase one study- as well as drug-to-drug interaction potential, which again, leads into safety. I agree with you that optimal is probably not the right term. It's the dose that we chose to further investigate in phase two.

Thomas: Thomas [inaudible 00:35:35], the university. Got a question about the second case study. Essentially you're basing your simulations on the fundamental assumption of additive relationship. Can you comment on what basis have you excluded that there could be a synergistic or an antagonistic effect.

Diane Wang: There really no sense of a base for that. However, if you think about it, if you're the synergistic effect, the synergistic effect can be any kind of magnitude and synergy. That's why I said the additive fact, it was used as a base case. Actually, while I was showing that example ... Dose level 4A, as example, that we have the observations actually, that clinical trial is current, ongoing actually. The grade three and four neutropenia rate I've seen from the first two dose levels, are actually 67%. Which is pretty much in line with what we predicted. I agree with your point, but since this is

already a very profound additive effect, and that's the base we want to use to give us a chance to evaluate this combination.

Speaker 6: I just wanted to comment on the neutropenia model, and the use of that. First of all, commend you for using that, but also I ask you to comment on a photo used beyond this sort of GSF or alternative schedules that you might consider. Notably, NTD and dose selection often times based on MTD is, of course, an implicit MDT as a DLT period. Which often times is over 20 days or around there. What models, like the neutropenia models, allow you do to is look at the longitudinal time profile of lab value, in this case. You might have other models of adverse events over a period of time that go beyond the DLT period. Then allow you to look at what happens over a longer duration of therapy. I was wondering if you could comment on that.

Diane Wang: That's actually very good comments. For the chemotherapy that we used in this trial, actually, there is a cumulative effect on neutropenia. When I said we observed 67% of the patient with neutropenia actually that's not what we based on the first cycle. The first cycle is for DLDT determination, however we collected ... We look at the data we observed at the time, even at later dosing cycles. For example, some of them were lowered at cycle two and cycle three, and then those still count as our dose reduction criteria that we are looking for. Also, actually fortunately, normally the neutropenia grade actually can stabilize at first two cycles in most of the cases. There's no cumulative effect after that. The early prediction actually works out pretty well.

Gabriel Helmlinger: Hello everyone. First I'd like to thank the organization very much for this opportunity to present here in quantitative systems from oncology model on immuno-oncology. My name is Gabriel Helmlinger. I am within the quantitative clin-pharm department, led by Don Stanski in early clinical development at AstraZeneca. My background is in biomedical engineering, applied math, and I've been developing this type of systems model for the past 16 years in a pharma context. In research development settings, as well as across therapeutic areas. In the earlier years, it just didn't have that fancy QSP label. The work today, I've developed, in conjunction with our collaborators. Eminent Decision is a modeling group based Moscow with whom we have had a long history of collaboration on this type of models.

On this slide ... Basically what our systems pharmacology models, or I like to call them drug disease models, basically it's nothing more than an evolution and expansion from PKPD modeling. In PKPD modeling we strive to put together dose and exposure in a PK model, and then relate exposure to response. Here, with drug disease modeling, systems pharmacology modeling, we strive to insert what we know about the underlying biology, and possibly physiology, and how that relates to an endpoint of interest, for example, [human dynamics 00:40:59], into response. All of this keeps being driven by the pharmacology and the PK. I will present some more details about how we do this today.

First, a reminder of the cancer immunity cycle, which is of interest to you today. Typically, this is a cartoon stolen from one of the [mailman 00:41:20] reviews. We start describing it at the bottom here, with cancer cells dying. They release tumor cells, specific antigens that are picked up here by antigen-presenting cells. This leads into T-cell activation, T-cell trafficking all the way to, some of these, including cytotoxic T-cells infiltrating the tumor tissue. Then recognizing those cancer cells in the tumor tissue, and proceeding with further tumor cell kill. Hence, you have this circle cycle. It looks like a very nice linear process that folds onto itself into a nice cycle, but as we know, the situation is much more complicated than this. We have numerous feedback and feed-forward loops in here, and different strategies are being implemented to address some of this processes, either via inhibiting mechanisms or activating mechanisms. Complementary strategy's also to come in with treatments that directly affect tumor cell kill, such as radiation, chemotherapy, all the targeted agents to enhance tumor cell kill, to enhance, indeed, antigen release. The situation is further complicated by the fact that over time, immuno-suppressive cells set up and multiply in the [inaudible 00:42:47] environment. Other strategies are targeted at these mechanisms. I will present, again, some of the work we did in this context.

Here is our model development strategy for the systems pharmacology of immuno-oncology. The core of the model is here where we represent immune cell types, such ass effective T-cells, regulator T-cells, immuno-suppressive cells, for example, NDSE's, Myeloid-derived suppressor cells. How they accumulate for you to identity them, how they interact among themselves, and how this leads overall to tumor cell growth, tumor cell kill, which then translates to tumor dynamics. We are today at this level, and we have built, calibrated, and used this model in different mouse experiments, mouse that it. We're in the process of transposing and re-calibrating this model in human. Bare in mind, this model, which is really the underlying biology physiology, is all driven by the PK, so the dose exposure of the different therapies that modulate the system that we are studying.

Here's actually a study cartoon of what we have in the model to date. Basically, in blue you have the different cell types that are relevant. Effective T-cells, regulator T-cells, of course driven by mature dendritic cells. We also have immuno-suppressive cells like MDSC's. All of these interact in different quantities and kinetic terms to modulate tumor growth and annex here, which is represented as an output function. In red font you have the different molecular targets that are affected by the agents, or the drugs. We have full calibration to this date for anti-PD1 therapies, anti-CJA4, OX40, and anti-CXR too. We have all the targets represented in the system, but not fully calibrated to date. The calibration has gone through multiple stages. To date we have about 20 independent mouse experiments that are represented in here and captured in about six different mouse tumor models.

While you don't have that many players in the system, so to speak, so the blue and the red fonts, you can appreciate how the dynamics. The arrows are relatively complex despite the few number of players that we have chosen to put in the

system. Just one illustration, for example, we had several expendable data with different mouse tumors and different antibodies against CTLA4. TO capture that kinetics in a system, was actually quite challenging. You have to use your mechanisms that everybody knows of, where anti-CTA4 will simulate regulator T-cells, effector T-cells. Represented by these two green arrows. We could not describe the collective set of evidence, the collective set of data. If we went up to add this inhibition arrow on this other red arrow, which is itself an inhibition from T-regs to T-effectors. Inhibition on inhibition gives an activation. Another mode of activation of effector T-cells was anti-CJA4. That's just to illustrate the stages in a building process in this kind of a model.

Another word I like to say is about NDSE's, for example. Often these models are described as, depending on who you talk to, too complex or not complex enough. The reality is, you apply the parsimony principle at every stage of building this model. Parsimony principle means as you bring in an additional piece of calibration, an additional set of therapies, for example anti-6R2, you must capture the minimum components to make it work while everything else still has to work with the previous sets of calibrations. NDSE's for example. As we started with this, the biology of NDSE's is much more complex than what you have represented here, of course. We first did a biology map of NDSE's interactions with the other cell types. A Q rating of about 100 PEPs in the literature. From there we chose the three, four key interactions, kinetically speaking, that needed to be in this model to calibrate everything against these therapies, what we had data.

What can these models be used for? They definitely provide value in supporting decision making in different areas, and really along the big four, including those regiments, duration, how you sequence combinations. It can provide elements of support towards why do certain individuals respond or not respond for the same dose, or dosing regimen, or same combination that's applied. It can support combination priorities. Ion and ion, but also ion and non-ion mechanisms. Whether it's small molecules, or biologics. It certainly provides a benchmark for various compounds because we have here, one integrated model that acts as single frame of reference. That really allows us to compare. Let's say for a given therapy, you want to look at what bio-marker one and two are doing. For the next therapy, you can compare what these two are doing in the same frame of reference.

I'm going to illustrate now some of the work we have done. I start with anti-6R2 and anti-PD1 treatments. This is based on literature data from Dan et al. 2014 in a rapto myosarcoma mouse model. Top panel shows expendable data. You have control panel, you see control values in black font, red font are the anti-PD1 treatment. Next panel shows the anti-CXR2 treatment versus control. Then you see the combination treatment effects versus control. The model was able to describe this kind of data adequately, in fact. If we look at the lower panels, the model-based simulations. The dotted red line represents the average of the expendable data above. The continuous red line is the average tumor size simulated by the model. You have a fairly good match here. Again, this is not data fitting. You first run the model in forward mode, then you see how the expendable data fits onto it.

Middle graph shows the same good match between model predicted average tumor size versus expendable data, and same thing for the combination.

What the model also tells you at the same time, you can simulate all the elements within the model. For example, all these different cell types we have represented. T effector cells, you can see what they do, and anti-PD1 treatment. They go up in time, reach a peak, and go back to a lower baseline value. You would expect that from the mechanism of action of that drug, of course. You see here, in relative quantities, what is happening. At the same time you see what NDSE's do. They'll actually pick up quite impressively, and are not much modulated by this particular anti-PD1 treatment. Anti-6R2, the mechanism is to inhibit recruitment of NDSE's into the tumor micro-environment. What it does ... Indeed, there was some inhibition in terms of a delay in the quantities of NDSE's, but they pick up later. That translates into a slightly delayed tumor growth, but not that impressive. The level of T effectors and anti-CX2 does not have so much of an effect.

Now in a combination setting you see very interesting, very effective T effector response. A much delayed NDSE response. Again, that makes sense given the kinetic arrows we have in the model, and it explains the tumor growth analysis that you see in the combination. An initial slow down, a low plateau phase, and then it picks up. We're very happy with this and we also achieved anti-CXR2 calibration with two other sets of data not shown here. Notice here what's very prominent is a large variability in response from one animal to another. The point I want to make is these kind of models can also capture this type of variability. It's not just about projecting average responses, you can also see what happens at individual levels. That's shown in the next graph. The way to capture inter-individual variability in response, we go though a so-called model-based sensitivity analysis.

I tried to summarize it, whether we do it in terms of sensitivity analysis. For each therapy, we take each parameter in the model, and you vary the value of that parameter gradually, in very small steps, plus or minus X percent, for example in this case, 60% around the median value of that average parameter value. You do that in X number of steps, 300 steps in this case. Each time you re-simulate the outcome of your model, which here is the tumor growth dynamics, you see how the outcome changes or is sensitive versus that one parameter you're changing. There are certain parameters that don't do much in that respect. Then you can actually identify, given the wiring of the diagram, you end up identifying some parameters that are extremely sensitive and are very important in driving the overall response, tumor growth dynamics. In this case K and one, which represents on that arrow, the [chisel 00:52:53] influx in the tumor tissue. If we look at this treatment, anti-6R2 alone, there is not much sensitivity to K1. Also makes sense because this is not, presumably, related to the mechanism of action of that compound.

How do we see that? The dotted red curves represent the envelope of the individual tumor growth dynamic responses. The continuous red curve is the model predicted average of this data. In green and blue, you see what happens as you

vary K and 1 in individuals. For each individual, what happens to their tumor growth curves. It stays very much along the average response. Look at the bottom panel, combination treatment now, anti-PD1, anti-6R2. Again, the dotted red curves represent the envelope of the individual experiments of data. The continuous red curve, the average. Now you vary K and 1 around that average, and you see tremendous variability in response. To the point where some animals respond completely, a full tumor growth control, and others respond not at all. If you take an arbitrary cut. Let's say at day 80 I want to look at tumor size, at day 80, across the animals, and you bin that into tumor size responses. You see a lot of responders, very effective tumor growth control, a lot of non-responders, nothing much happened, then some individuals between. It's not your nicely Gaussian distribution of responses. It's like an all or none type of phenomenon. 40's one, again for this one parameter that we're looking at.

Not all of them make the model behave like that, but there are some that do. That's the point here. Of course, K and 1, if you talk to a biologist or oncologist, T-cell influx into tumor tissue for treatment, of course that's important. Everybody knows that. It'll tell you in what relative quantities and kinetic terms it is important, and what it can actually do in terms of response versus non-response. It can drive it from a response to a non-response, and not in a small way.

If I go on, another form of therapy we looked at in a combination setting is radiation with anti-PDR1. One set we used for calibration, this is again, mouse data from the literature by Dan et al. They looked at two tumor types, we'll report the results here for MC38 tumors. The same applies to what they did on tuber tumors, and what the model does in that context. Upper panel we look at tumor size, so expendable data. The average expendable response in each group is shown in dotted black curves. You see that the dotted red curves, the average model response follows adequately the expendable data. We have an adequate sense that the model is doing what it's supposed to do here. Again, this is not data-fitting. You run the model in a forward mode, with the expendable conditions set up in that paper, and you see where the model goes, and then you plug onto it the expendable data. It's not only a pure match, but the trend is always confirmed.

With the model, as I just explained, we can also simulate individual animals. This is what we did here. Interesting here, so why you match the average behavior from the experiments, you can do that in individual animals. Anti-PDR1, what you see is, again, a bit of this all or none type of response in this particular experiment. While on average, you have an initial tumor growth control, and then an escape. At the individual level, it's either animals respond very well, with full tumor growth control, or they escape rather fast. On average, it gives you the impression that there is something happening here and then escape later, but at the individual level it's a different story.

Radiation therapy on day eight. This was actually a massive dose of 20 ray. You see more of a uniform response from individual to individual, and the average data is well-described as well. Now you go into combination setting. On average, very

effective, full tumor size control. Anti-PDR1 plus radiation. However, a the individual level again, you see something very interesting. There are four animals in this simulated experiment that do completely escape tumor growth control in this combination setting. On average, 96% of individuals are under control here in terms of the tumor growth. Now, I can look at the same time, per usual now, at the cell counts in the model for the different forms of therapies for the parameter conditions set forth by Dan et al. What I see is well under control conditions. Yes, you have a certain level of NDSE's that set into the tumor in micro-environment over time. You have a certain amount of T effector cells, and a certain level of regulator T-cells. That's under baseline control conditions, where the tumor just keeps growing. No treatment.

Notice also, I forgot to mention when I showed the model cartoon, what we found out is very important, is not so much the absolute values of T effectors and T regulators, it's the ratio. That ratio T effectors to T regulators has to be high enough to achieve sufficient tumor growth control in whatever combination setting you are looking at. On the anti-PDA1 alone, you do see a massive T effector response with that treatment, some control of NDSE's before they escape again, and a bit of a T regs response, but the ratio of T effectors to T regulators is way high enough to get some tumor growth inhibition in the animals around here. The radiation alone, here you see that there is some effect on NDSE's, partial control over time until it picks up again, some T effectors increase, but not too much. Presumably here, the T effector to T regulator ratio is not high enough. In a combination setting you get all this desirable attributes.

What's interesting here in this combination setting, radiation anti-PDA1, we had an independent type of experiment. In house in the pre-clinical domain. Some of this data has actually been published by Dovete et al. cancer research 2014. Different tumors in mice, and here we also have the luxury of comparing two different dosing regiments in the same combination setting. We have no treatment here, again, these are tumor size data, radiation alone five times, two rays daily over several weeks. Each trace is a different animal. Then you have radiation with anti-PDA1. Now, here the anti-PDA1 is applied just one day after the initiation of radiation therapy. For most animals you have very good tumor size control.

Here's another setting. Same antibody for anti-PDA1. Radiation therapy. The same radiation therapy regimen, but anti-PDA1 is applied 19 days after initiation of radiation therapy. Here you do not have such a impressive tumor growth control. The model here can not only reproduce the average tumor size data, again same principle. Dotted line are the expendable curves. Red lines are the model-based tumor sized curves. You can also follow what happens to the different cell types, and molecular marker. For example, 3PDA1 in the tumor tissue. That model applies only to tumor tissue. Interesting phenomenon again, radiation allows you to control NDAC's to some extant, before it picks up again. In the combination setting, the first treatment where anti-PDA1 comes in right after radiation, we do have a PDA1 decrease tumor growth control, and you have a very effective T effector response here. Also, a very effective suppression of NDSE's, green curve.

Same combination setting, different regimen. You come in with anti-PDA1 much later. Not good in terms on MDSE control. Some T effector response but not enough compared to the blue curve here. That can possibly then explain why this dosing regimen is not as good. This was for the calibration and validation piece. We have used this model extensively to simulate various experiments or scenarios in a pre-clinical setting. We're going to show you just a set of results for this. We wanted to simulate tumor size response and the corresponding cell type responses for various scenarios. To do this, this is now really playing around with the simulator, we defined a tumor type, I put it in quotes because what we mean here by tumor type is a set of baseline setting conditions of cell counts. At this baseline setting we allow for some inter-individual variation. For example, here we have a tumor that, we call it immunogenic tumor, because we do allow for the tumor a high influx of T-cells, but we also set a baseline, relatively high levels of immuno-suppressive NDSE's. This is a real setting in mouse tumors such as MCA25, or other myosarcoma. Tumors like this, they grow of course, and control.

You have a certain amount of T effector cells. Some T regulators as well. The ratio of these two not very high relatively speaking, and you see the variation we allow for the different individuals. Each trace for different mouse in terms of baseline quantities of T effector cells and T regulators. Same story for NDSE's. Now you bring in a test compound. We have the PK model for each of these, and you apply it to this hypothetical tumor. We take a test antibody, we say let's put in .2 milligrams at days ten, 20, and 30, and we have the PK model for that antibody. We apply directly into the QSP model. We simulate the tumor size and cell count responses. We can also compute a response rating quotes again, for that experiment. Responders being an animal that ... Which is a tumor size less than X hundred micrometers at a given point in time, her we chose day 50. Here are the simulated results for such a tumor. We bring in anti-PDA1, where do we see the tumor size level, there is again, impressive split almost, so to speak, between responders and non-responders. What do the T effector response look like? Given that small variation of T effectors at baseline, some individuals end up having a very robust T effector response under that one treatment, mono-therapy treatment. Others don't. NDSE's, same story. Some do not show any control or suppression of NDSE's, other individuals do.

Now we look at anti-CXR2 alone, mono-therapy. Here too, you have responders and non-responders. In terms of key factors, there's some response, but not as impressive as under anti-PDA1 alone. Again, that makes sense from the mechanism of action point of view. In terms of NDSE's, much stronger control with anti-CXR2. That also makes sense if you consider the mechanism of action. Now, these percentages are these artificial responder rates in the experiment at day 50. 18% of the animals have a significant reduced tumor size by day 50. In this one, only 34%. In a combination setting, you have much more impressive results. You see what happens to tumor cell growth on average. Now notice, at day 50 there is a tumor growth inhibition, but it doesn't go without the growth phase initially followed by regression later on. Dynamically, this is very different response from this or this.

From one individual to another as well. If you look at T effectors, very, very effective response. Also from individual to individual you can have variations on a long time, so important to consider this inter-individual variability. NDSE's fully under control, and very active dentritic cells.

So what we did is ... Here is just a summary table for the different cases we have played with. These three columns, effector T-cells, regulator T-cells, NDSE's, others, the conditions we set at baseline. Case one, as already explained, we put some amount of T-Cells along the baseline. Which is reflective of certain tumors like this. Very few regulator T-cells. High levels of NDSE's. These are the different treatments we applied and an over response wait for the experiment. Also, we can look at individual responses. Case two we tried some level of effector T-cells, strong regulator T-cells presence at baseline, and more or less, some presence of NDSE's. Here, with the different therapies, again, can do. Just are in mind, this is all in mouse and it's in fixed doses of hypothetical antibodies that we just throw in a .2 milligrams in the simulator with the corresponding PK model. We can actually do this with realistic dosing regiments for actual antibodies that are either in the pre-clinical or clinical domain.

Case three, that's another type of immunogenic tumor with some effector and regulator T-cells on baseline, a lot of NDSE's. A bit different from this case one. You see how the different treatments play out here. I'm not going to comment too much on this. Case four, that's the typically very hard to treat scenario where the tumor is purely immunogenic. You don't have much effector or regulator T-cells at baseline. You have some amount of immuno-suppressive cells. You do need minimum double, if not triple, combinations to ... If you just look at ion mechanisms, in combinations, to get some effective response. You can also do this for ion and non-ion combinations as we have shown for anti-PDA1 in radiation therapy.

In summary, we looked at the feasibility and utility of such systems models. In terms of attributes that are quantitative, that contain multiple targets, so in that respect they are integrative. They're mostly mechanistic, doesn't mean they're entirely mechanistic because some of this arrows that you see in the model, sometimes they're empirical functions because we do not know enough of the mechanistic biology that's underlying. They're as parsimonious as we can make them because we want to include essential biogen [inaudible 01:08:33] to reproduce the existing set of data, if not go beyond that. We calibrated and validated this through multiple mono-therapies and combination therapies. We did that for small molecules, biologics, other modalities such as radiation. To this point in experiments of mouse models. As I mentioned at the beginning, we are re-calibrating this for human. We are also working on the addition of direct tumor cell kill affects via targeted compounds to support this portfolio of non-ion plus ion combinations.

Last slide, acknowledgments, I think Dr. Kim mentioned in the opening that this is very much a multidisciplinary area. I think that point was also stressed by Joe

Biden's talk at ASCO, multidisciplinary. Why we have many moderators with backgrounds in mathematics, physics, an engineering. It doesn't go without all these different functions that are also represented here. Quantitative clinical pharmacologists, oncologists, and other medical experts. Bio-marker and bio-informatics experts. We work a lot with pre-clinical biologists and pharmacologists across AstraZeneca and MedImmune, also pre-clinical modelers. This is a lot of work hand-in-hand, with literally, daily interactions. Thank you.

Dinesh De Alwis : Dinesh [inaudible 01:10:18]. Gabriel, it's an excellent presentation, and actually some very interesting points there. My question is really around the kind of platform you have and it kind of simulates all these hypotheses based on obviously, a knowledge-based model. It's coming up with some very interesting hypotheses based on the pre-clinical animal models that you have. The problem know is how do you actually test this out in the clinic. What you're actually advocating does require some very different way of developing drugs because you actually need to match up the histopath of the patient in terms of what's actually happening in terms of T-cell activation, in terms of what the state of the tills are, et cetera, et cetera. How would you actually put this into practice, otherwise it'll just be lost in translation.

Gabriel Helmlinger: That's, of course, the big question. The purpose of starting pre-clinical was to catch a lot of the wiring in that diagram that I showed. The big assumption is the underlying wiring remains, but it's the relative quantities and kinetics, which means re-paramaterizing the model, that is important, and that's going to be fundamentally different in human. That's what we're looking at right now. While clinical data sets are typically not as rich, because we have, let's say, 80 to 90% of the underlying wiring we can use with this [inaudible 01:11:56] human to re-calibrate the model. If we call for other adjustments as well, for example, there's a basic tumor growth function in that model, which in mouse is very different from human. There are lots of again, literature data on human tumor growth data that we can use to again, build that function. Where some of the mechanisms would be un-parroted driven, not possible entirely mechanistically.

Dinesh De Alwis : Just another question. Can you comment on the sort of synergetic models versus sort of the immuno mouse models and how informative they are in terms of a platform such as this. So the sort of humanized mouse models versus sort of the syngenetic.

Gabriel Helmlinger: Basically we ... Whatever the experimental model was, we would capture the experimental conditions and calibrate the baseline of the model accordingly. It was not much of an issue.