6-13-16 FDA-AACR Oncology Dose Finding Workshop – Session 2 Transcript

Todd Palmby: It looks like some people are still filtering in but maybe I can get started then provide a little introduction to the next session. Welcome back to the afternoon session. This is session two. Session two will be on nonclinical models used for go/no-go decisions. My name is Todd Palmby. I'm a pharmacology-toxicology supervisor in the Division of Hematology-Oncology-Toxicology in OHOP at FDA.

Those of you who are familiar with the drug development process maybe asking why FDA is interested about nonclinical models for efficacy for oncology products especially given the risk-benefit for patients in the setting of early phase oncology trials. The level of activity in nonclinical models that is required from a regulatory perspective is generally low. Historically, the frequency with which activity in traditional nonclinical models such as in vitro cell proliferation or cytotoxicity assays in mouse xenograft models using human tumor cell lines translates to efficacy in patients. Translation to efficacy in patients has been low. Nevertheless, available nonclinical models and the understandings about how to use them has evolved in recent years.

The current focus on clinical dose optimization strategy is to improve efficiency and success of oncology drug development with newer generations of targeted agents presents an opportune time to readdress current approaches in translating information from nonclinical models to predict the best ways to use a therapy or combinations of therapies. In addition, there are different strategies and considerations for translation of nonclinical models to predict human efficacy for small molecules as compared to biologic therapies. The success of improved immune oncology therapies and recent wave of development in this area presents a unique challenge for applying nonclinical efficacy models since these products generally target the host immune system rather than the tumor cells themselves.

The goals of this session are to discuss the state of the science in nonclinical efficacy models, to discuss best practices and target validation in lead selection in nonclinical efficacy models, and to discuss the state of science in PK-PD modeling to translate nonclinical tumor accumulation characteristics and identify a minimum target plasma concentration in patients.

First, I'd like to introduce the first speaker in this session, Dr. Darren Cross who is a principal scientist with Oncology Bioscience in AstraZeneca.

Darren Cross: Good afternoon everybody. Thank you very much for the invite to this great workshop today. It's a pleasure and a privilege to be here and talk about our nonclinical discovery of osimertinib. Could we input them? What I'm going to do over the next 20 minutes or so is give you a brief overview of the nonclinical development of osimertinib.

Osimertinib is also known by its trade name Tagrisso and its code AZD9291. The discovery program of osimertinib was very fast. Actually, from project initiation to first time in humans was only four years, and that only took over two and a half years to actually then get approval from the FDA. The whole development and discovery

process of osimertinib was very rapid.

What I'm going to try and do over the next 20 minutes or so is give you a reflection of our experience from a nonclinical discovery perspective of how we use the nonclinical models to drive that success and speed of the osimertinib discovery program. I'm really focused on themes of really defining the specific design goals, underpinning the design goals really strung nonclinical models and assays all the way through from lead selection all the way through to helping predict [inaudible 00:04:11] humans.

Before I get into the project, for those who are not familiar with the area, I'll give you a very brief introduction to the EGFR mutant area. Back in 2004, activating mutations in non-small cell lung were identified. The most common of these being, shown on this diagram here, deletions in exon 19 and the L858R. Only with these activating mutations found to be oncogenic in the surface of non-small cell lung. They were also found to sensitize these tumors to EGFR ties in kinase inhibitors, and this subsequently led to their approval in the segment.

However, as we've all [inaudible 00:04:54] agents, unfortunately, these are not curative and patients will then firstly digress due to acquired resistance within time. It became quite clear that the predominant resistance mechanism to these TKIs was just gained at the second mutation in EGFR T790M gatekeeper mutation in the end EGFR. This rendered the EGFR receptor refractory to these TKIs, and this occurred in approximately 60% of cases so it is a very predominant mechanism.

Back in 2009, when we think about this program, EGFR T790M was still an area of high unmet need. As I've said the first generation reversible TKI such as gefitinib weren't potent against T790M. There were of second generation of TKIs developed such as afatinib which were irreversible. Despite having increased potency because of their irreversible action, they are still unable to achieve the adequate clinical exposure to target T790M effectively because of the work limited by anti-EGFR toxicities. There were no approved therapies available. This has made us think: could we actually develop a new class of EGFR or TKIs to actually target T790M more effectively? This is what led to the osimertinib program being started back in 2009.

Really, as we think about the project and what we did in that nonclinical discovery phase, the real first major thing we did actually which set the project up for success, i think, it was designed from the [inaudible 00:06:28] of age-specific design goal. We knew we had to develop a profile of an agent which had potency against deactivating mutations as well as potency against deactivating mutations in the context T790M double mutant and also drive a margin against anti-EGFR. This is a profile we needed to know to drive against to differentiate it from the earlier generation TKIs I've just talked about.

This real focus allowed us to be very specific in terms of thinking about the chemistry mechanistic hypotheses to be able to achieve this. We knew we have to find lead scaffolds that could target that both in methionine preferentially. Then in turn, this led to us to be able really [inaudible 00:07:10] rapidly generation approach. In fact, we

only screened 40 compounds from our collection that met those kinds of design criteria using biochemical assays as shown here. That very rapid early lead generation allowed us to prove the principle that this is possible because could find these scaffolds that did show this differential activity towards T790M and what type of films of EGFR and also very quickly identify attractive lead scaffolds take forward.

This is the first real breakthrough, if you like, in the nonclinical approach that allowed us to really get moving very rapidly. This has enabled us to move much more rapidly then kind of a historically in high throughput approaches we have done.

The second key decision we made, again, having underpinned the success and speed of the project was very rapidly then moving away from the biochemical assays. We took a decision to actually focus the chemistry testing using nonclinical in vitro cell line assays. The premise for doing this was really is that we thought these cell-based assays which would be in a much better position to define their target disease biology to drive that chemistry.

Furthermore or moreover, actually we're in very fortunate position of having a lot of well-characterized, validated cell lines available that covered the various different chemical-relevant mutation spaces as shown here. This put us in a very advantageous position. I think, again, this is, i think, a theme of this program is having a very good preclinical disease models that can really mimic the biology of your target are necessary to really drive the whole chemistry and predictive modeling, et cetera. This put us at great advantage.

Really, the main workhorse assays I'll talk about as we go through the story [inaudible 00:09:04] H1975 which represents the double mutant T790M with the L858R activating mutation and PC9 which represents the activating mutation exon 19 deletion.

Really, one of the first examples of the impact taken the strategy of focusing on using the cell-based assays to drive the chemistry is shown here. Although very early on where our lead scaffold showed very potent activity in biochemical assays, I guess, this is double mutant T790M enzyme assay. We were consistently seeing that these leads were not able to carry their activity through the cell-based double mutant T790M assays so this is H1975 cell line. We saw a consistent drop off in cell potency and this is due to the high ATPKM barrier, the T790M confers.

This is using these cellular models allowed us to actually design [inaudible 00:10:03] concept molecules that showed that if we added an irreversible covalent warhead to these molecules that was the structure that allowed us to maintain the potency of these compounds against the T790M cell lines.

Again, this very early on allowed us to be very focused in our design and approach going forward. I think another key for me, a very key thing we did decision we made very early on in the project was actually building in vivo xenograft models very early into the discovery program. Although xenograft models have got the limitations in terms of being able to mimic perhaps [inaudible 00:10:45] more complex extrinsic

treatment biology, they can be very good for mimicking more intrinsic aspects of treatment biology, cell biology.

Again, we were fortunate because of these models were very well characterized and validated so we knew the likes of H1975 modern PC9 were very good at mimicking EGFR mutant disease context. Because as well, we know we had to work in irreversible covalent space; again, we thought applying these models much early on would be much more powerful in building that more complex PKPDF relationship and therefore be in much better position to help drive the lead selection for optimization, prioritization, et cetera.

We really made a decision. We wanted to build these models into the cascade much early than perhaps you normally do historically in projects. We took a lot of time actually to optimize these models and be able to reduce the duration of the efficacy time course and also to decrease the number of [inaudible 00:11:51] in each cohort therefore be able to kind of much put these models in a position where we can use it with a much higher throughput manner. This is, I think, for me was like a key step in being able to then drive the lead selection optimization much more rapidly and robustly and also be able to very early on drive the understanding of the more complex PKPD efficacy in the program.

Similarly, we use these models in the similar manner to build a very early pharmacodynamic PD data. Again, the same rationale because of these irreversible compounds and because we have access to these very good xenograft models which did recapitulate the disease very well and the target very well, we wanted to build, to use these models much earlier to help drive the PKPD efficacy relationship and therefore help drive the lead selection optimization much more rapidly and robustly. Also, by doing this, it was able to also generate a much more robust status that actually also supported the [inaudible 00:12:53] human modeling and later on which we'll come and talk to later on.

This is just examples, some example data we generated for osimertinib showing the PD data both in terms of dose and time, of course, against the target in [blue 00:13:08] EGFR, phosphor-EGFR, and also downstream signal molecule AKT. This is the data we were able to very easily and quickly generate across both early lead series and as well as for osimertinib.

After three years of how we use these nonclinical models both in vitro and in vivo in this kind of manner, we're able to very rapidly [inaudible 00:13:32] osimertinib as the candidate drug after three years. This just summarizes some of the key data for osimertinib.

You can see in this table, it didn't meet all the design goals in terms of in vitro pharmacology having potency against both the single and the double mutant T790M cell lines whilst retaining less activity against wild-type cell lines. This is in contrast to the early generation compounds. Also, we could show using these xenograft models of [inaudible 00:14:00] how osimertinib is able to deliver very strong efficacy at low doses

in both the 1975 T790M representative model as well as a PC9 activating mutant model. Also, osimertinib is very selective kinase inhibitor, a very good DMPK properties. We, again, like I said, we used these data to drive early those predictions, and I'll come on to that in a moment.

We didn't stop there. We really want to continue and build confidence in osimertinib as our candidate drug. We did do support from further in vivo studies. One cool experiment actually did with this and xenograft models was in a chronic treatment study. Again, taking the 1975 as the representative T970M model and PC9 as the representative activating mutation model, we could show that very low dose of 5 mg and 25 mg/kg/day was sufficient. Actually, it's tried both for these models into a rapid and complete sustained regression and actually complete response which was maintained upon continued dosing.

This data actually got us very excited about the opportunity of the osimertinib could actually bring. This data is very exciting and it gave the impetus to move this compound forward as quickly as possible into the clinic. However, because we were aware that a lot of our work have been predicated on these xenograft models and cell lines, we wanted to make sure that is activity and efficacy levels we were seeing also translated across into other disease models. It wasn't just an artifact of a xenograft. I wanted to make sure the xenografts were representative of disease models.

One model we did, we did this in the genetic-engineered mouse models. For a very good collaboration with [William Powell 00:15:59] when we was at Vanderbilt, he used these models. These models because they spontaneously form tumors in immuno-competent mice in the appropriate tissue, in the lung organs of mice. It can arguably they better recapitulated more holistic human disease and tumor biology, although they still do have disadvantages, some of listed there.

I think with all these models there are pros and cons so whatever study you want to do, you have to aware of the pros and cons of each in all of these models when you're designing your experiment and your objective. We did use these models because they do recapitulate the tumor biology very well. Again, we're very, very well established and validated. We could show that in these models as with the xenografts at very low doses, this is 5 mg/kg. Again, osimertinib was sufficient to drive very rapid and a sustained tumor shrinkage both in the single mutant L858R model shown here and also the double mutant model shown here.

This, again, gave us increased confidence the overly exciting data we'd seen in the xenograft models wasn't just the xenograft artifact or outlier and actually that the activity was able to be translated in other disease models and actually in more maybe appropriate disease models. This, again, gave us increased confidence and excitement to take osimertinib forward into the clinic.

More recently, we've actually … There will be a talk about this later in this session so I won't dwell on this. We are more and more focused on using patient divide [inaudible 00:17:41] models. These models although like xenografts they do suffer from

disadvantages in terms of not being able to mimic more complex and extrinsic aspects of tumor biology such as immune microenvironment. They are very powerful in terms of being able … because they are directly from the clinic. They are very powerful now to better mimic the heterogeneity of clinical tumors; therefore, also very excitingly can provide a co-clinical platform.

An example of how we're using these models now is here. This is a very good collaboration with Pasi who's here today and speaking later. This is an example of such a co-clinical study. The PDX Pasi took from erlotinib-resistant patients who had T790M and this patient who's being treated with osimertinib. What we can do actually parallel to the clinical study, we can then do a nonclinical study in parallel and use these models then to investigate drivers of resistance and response to osimertinib in a much more detailed manner and therefore help also inform potential combination opportunities.

This is example of some really nice data that started to come out using these models. This is showing that, again, a low dose, this is 25 mg/kg. Osimertinib can … in this PDX can drive a complete and sustained tumor response shown in the blue. Even after a treatment with gefitinib as the tumors become much bigger on retreatment with osimertinib. Again, you can drive that very strong durable regression. I think this is a very exciting area going forward and so we'll hear more about that later in the session how these models can be used to better represent the heterogeneity of clinical tumors and be used in this mode of co-clinical studies to really say help drive the understanding and the biology of resistance and response in combinations, et cetera.

In my last few minutes, I'm going to switch gears slightly and talk about how we used all these nonclinical data together in terms of preclinical modeling and simulation. I'm not a [modelist 00:19:49] so you'll have to bear with me; I'm a bioscientist but I'll do my best at the modeling piece.

Really, the idea was can we actually take all the data we've got from the PK, the PD, the downstream treatment pathways and phenotype I've talked about from these models? Can we understand the relationship and build a model to understand that relationship and therefore be able to simulate ultimately how the dose impacts efficacy?

With osimertinib, we had a couple of extra challenges with this molecule to build these predictive models. Firstly, we knew that osimertinib could form active metabolites in rodent studies. One [inaudible 00:20:26] metabolite is shown here; this says he's at 5104. This is [inaudible 00:20:31] because like the parent, it is irreversible and potent inhibitor of mutant EGFR and also it can inhibit wild-type EGFR. We have to incorporate the potential of these metabolites in any of these modeling studies which obviously added an extra layer of complexity.

The second complexity for the osimertinib because [inaudible 00:20:52] irreversible mode of action compound, after that the PK-PD relationship is much more complex. Actually, the level of target inhibition is probably more of a factor of protein turnover

because of the irreversible nature of these compounds rather than PK and exposure. Actually, it's probably the PD and the target inhibition is probably more relevant for thinking about biological consequences. Again, this is an added layer of complexity when we're trying to think about these modeling and simulation.

The modeling guys, the [inaudible 00:21:26] modeling guys [inaudible 00:21:27] going to take a modular approach to take all these different complex pieces of biology and information and build a much more holistic model to incorporate these pieces. For instance, [inaudible 00:21:40] modeled the PK and metabolism of both the parent metabolite, and as you can see here their predictive model was very good at predicting the PK of both for those components. The model here shown in the solid represents the data point very closely.

Similarly, they built a module to understand the kinetics of target inhibition. Again, shown here this model was very well at describing the kinetics of the PD data showing here for these solid bars versus the real data showing the points.

Finally, the [inaudible 00:22:18] to a module to look at the tumor growth and inhibition to kinetics so by integrating all of these complex pieces of nonclinical information from PK through to target inhibition and pathway inhibition and tumor growth kinetics. They're able to form this complex modeling piece and therefore, by using this integrated approach going to anchor a dose and schedule to the target in biology and therefore, then uses models to inform ultimately how to the PK exposure relates to target inhibition and efficacy. This is a very good piece of modeling work given the complexities of osimertinib.

We then went on to use this model to simulate what would be the likely potential clinical dose to applying the predicted human PK both the metabolite and the parent to this model. We found … actually, we predicted there were simulated the clinical dose required to achieve the same levels of tumor shrinkage as seen in the mouse models was actually very low in a very low mg/day level.

This was quite exciting because this aligned very well with the first dose level cohorts that, I think, Pasi will talk about later of 20 mg so it gave us confidence actually that even at that first dose level, we may expect to see efficacy which again was very exciting and again gave us confidence to move forward rapidly. Also, being able to model the potential contribution of metabolite also gave us the confidence to move forward.

I just put this slide in to show you that. Actually, the modeling was correct and we did see efficacy at this very first level for 20 mg. One of the first patient's dose did see a strong shrinkage at 20 mg so it did actually support that those modeling was actually very good.

Then finally, just to finish, the other mode we used, the modeling was actually [inaudible 00:24:25], I guess, reverse translation helping form phase two dose. I think it's Pasi will be able to talk about a bit more into detail because we saw efficacy at that

very first dose level at 20 mg. We actually ended up having efficacy from 20 mg all the way up to 240 mg cohorts. We haven't reached an MTD, so really the challenge was what was the optimal dose then to take forward into phase two. Again, by then applying the known human PK for both the metabolite and the parent back into that model, we're able to again simulate how that real clinical PK align to the efficacy we saw in the models.

Using this model, we actually gained confidence at 80 mg. This is probably saturating the dose response curve in the clinic. It did give us confidence that 80 mg is probably sufficiently optimal dose to move forward with, although it wasn't used for decision making and search for selecting dose, it does certainly helped provide confidence in that dose decision-making process.

I can finish because that's my last slide. I'll just leave you with a few key summary points, I think, to take away that I have taken away from this nonclinical experience. The whole success and speed of osimertinib was really based on the strong, robust, nonclinical models we had. I think it is key that you need models that do represent the appropriate disease context and target context. If you have got that which we had for this program, then it does really support a very rapid and successful lead generation and lead optimization, et cetera and underpin very strong modeling studies.

Each model does have pros and cons so from xenograft, to GEMs, to PDX, they all do have pros and cons. I think you have to take those pros and cons into consideration. Pragmatically, you can use this simple model such as xenograft to drive a chemistry and then you can use maybe the more complex models to drive more complex questions such as we are now doing PDXs. It can be very pragmatic.

As you're shown in this experience, the PK-PDs, the efficacy modeling was very powerful to help and support human dose and understandings. There are experiences showing that already a drug exposure itself is not appropriate. You really need to anchor the dose of PK to the biology. Now, we are going forward as we've heard about the small and the key challenges are … as the modeling tools are getting better so how would then utilize in nonclinical mode for support and combinations. Also, I think another challenge is how can we then also back-translate to ensure we can validate those nonclinical models using the clinical data. I think that's another challenge going forward.

In the project now, we are carrying on using this in the various aspects of nonclinical modeling to help inform resistance, how we can best tackle resistance, and also help inform how we best inform combination strategies. I'll finish the [inaudible 00:27:41] over time just to leave on the acknowledgement slide, so just thank everybody who's being involved in the discovery phase as well as the development team. I know obviously with the study patients and investigators who have helped us [inaudible 00:27:53]. Thank you very much.

Todd Palmby: Again, I think we'll take questions during the panel discussion to follow all the talks. The next speaker in this session will be Chandni Valiathan, associate principal scientist

as Merck Research Laboratories.

Chandni Valiathan: Hi everybody. My name is Chandni Valiathan. I would like to start off by thanking the organizers for inviting me to speak today, and thank you all for being here today. Today, I'm going to talk about a translation model that we developed, established dose range selection for pembrolizumab. The work I'm presenting was done by a large group of people and I'd like to acknowledge them upfront.

This is a brief outline of my talk. I'll do a very, very brief introduction of immuno-oncology. We've heard a lot about some of these programs and compounds. I won't [inaudible 00:29:12] but I'll do a very brief introduction. I'll also do a very brief introduction about the mechanism of action of pembrolizumab. Then I'll go into some of the details of the criteria and challenges in clinical dose selection and translational model development. The meat of my presentation will be around the translational physiology-based PK-PD model development. Once we get through that, I'll describe some of the model simulations that we did to suggest doses for further studies and then I'll summarize and touch up on some key points.

It's actually a very exciting time to be in this space in the immuno-oncology space. Just in the past five years or so, there's been an explosion of clinical trials, approvals, breakthrough designations. I think this slide is already outdated because we do have at least one antibody that has been approved. This field is just moving really fast and there's a lot of momentum and it's almost like a medical renaissance and it's really very exciting to be in this field at this time.

Today, I'll be focusing on pembrolizumab. Pembrolizumab is a humanized IgG4 antibody that binds to PD-1, programmed death-1. As we heard earlier this morning, PD-1 is one of the receptors that's present on T-cells that is one of those immune checkpoint protein so it inhibits that killing activity of T-cells. What pembrolizumab does is that it actually binds to PD-1 and inhibits the inhibitory signal thus reactivating the T-cells that align them to kill tumor cells.

Thinking about dose selection based on preclinical data, whenever we look at this question of dose selection, there are two components that I think we try to integrate. One is the PK-PD efficacy relationship and oftentimes, we have a lot of preclinical data that helps us characterize this relationship.

The other component is population variability in humans. Many times, we don't necessarily know what this variability would be either for PK or PD. In any case, what we try to do is we try to pull together information that we have and define a target PK or PD or efficacy cut off in a target percentage of the population that we're looking for. To do this, this is a difficult thing to do in general but in the immuno-oncology space, it's actually a little more difficult. The reasons for that are that we don't necessarily have very good preclinical models for immuno-oncology, and I'll discuss that a little bit later.

Also, we don't necessarily know or understand fully how efficacy that we in mass

models, for example, translate to the clinical setting. Some of the biology that we deal with in the immuno-oncology space is actually very new and so we might not necessarily have any PD biomarkers that we know or understand very well. There are also still questions about the appropriate PK parameter that correlates with efficacy. I think we had some of that discussion earlier this morning as well. Again, this is a very difficult question in general and also specifically in immuno-oncology. At the end of the day, what we're striving to do is to identify the appropriate dose level and frequency at which together form this appropriate dose.

Coming to the challenges that we face in preclinical species selection for transition methods in immuno-oncology, traditionally, xenograft models have been used to look at tumor growth inhibition [NC1 00:33:52]. With immuno-oncology as you all know, the immune system is an important component of the response. The traditional xenograft models, the mouse immune system is weakened or ablated so that we can actually inject a human tumor. This doesn't really work for immuno-oncology.

There are a couple of options that we can go for. One is the syngeneic mouse where the mouse immune system is kept intact and we study a mouse tumor. In that particular scenario, we then have to study a mouse antibody because many times, the human antibodies do not cross-react with the target in the mouse.

The other type of model we can use is humanized mouse models and in that case, the mouse immune system is ablated and replaced in some sense with a human immune system and then we study a human tumor in that setting. All of these models have their limitations, but these are the tools that we might use to study immuno-oncology compounds in the preclinical setting. The model that we used for this particular model that I'll describe is the syngeneic model.

Because we're doing something that's really froth with uncertainty and unknowns, it actually really helps to integrate data from various sources as much as we can. Wherever possible, I think it is beneficial to incorporate data from clinical studies, literature, in vitro experiments, in vivo experiments or ex vivo experiments, and any other data that you can use. In the example that I'm going to present, we actually integrated all of these [inaudible 00:36:01] components so you can see that we integrated data from various different sources.

I'll actually skip this because this is a repetition of the earlier slide and jump right into the meat of the presentation. All of the model development work that I'm presenting here has been summarized and written up in a manuscript that hopefully will be published soon. If you're looking for details, you can refer to that manuscript again once it's available.

The story actually starts in a very different place than what we would think of traditionally. I've put an icon at the top of my slides to help you follow where the pieces of data and models came from for that particular slide. Here, the story starts with human data. At the beginning of all of this, we had already started phase one studies, so just recapping what [Dinish 00:37:16] went through in the morning, we'd

actually tested three-dose levels in the dose escalation paradigm. This is basically [part E 00:37:25] study that Dinish described in the morning. We had seen promising results with these three doses on a once-every-two-week schedule.

The traditional way of selecting the maximum tolerated dose in oncology would have suggested that we go with the 10 mg/kg dose once every two weeks. Now, the pharmacokinetic properties of pembrolizumab followed the standard antibody kinetics, and we estimated a half-life of about 26 days which, again, as Dinish mentioned this morning led to proposing once-every-three-week dosing schedule.

Additionally, there were some PD marker data so the IL-2 data that Dinish presented this morning, that led us to believe that there was some saturation occurring at 1 mg/kg. In some sense, we started off with quite a bit of knowledge. We had seen efficacy, we knew some PK parameters of the compound, we had some sense of population variability minimal information about that in humans, and also we had a PD marker. We started off in a really good place.

To actually develop this translational model, what happened was we went back and we designed these mouse experiments to specifically understand a dose range and specifically characterize dose response relationship. Again, preclinical studies were properly designed with the range of doses we tested somewhere between 200 and 300 mice, and we had data separated out to build a model versus validate the model. The parameters that we measured were plasma PK, blood and tumor receptor occupancy, and tumor growth inhibition.

Once the data was available, the model development process began. It ended up being a monster of a model and it had these different components that try to capture all the information that we had. We started off with the PK model and this is a standard empirical compartmental model and [inaudible 00:40:07] of some details in the next slide.

That particular plasma PK model was coupled with a physiology-based PK model to describe tumor PK. The main structure of that physiology-based model was taken from a publication and components of it were fine tuned for our dataset. We also then linked that tumor PK model to receptor occupancy which we had measured in mouse. There was some target-binding constant that we had from in vitro experiments that we fed into the model and again, the other parameters were estimated from the mouse data. Then that receptor occupancy actually drove tumor growth. Again, this tumor growth model was taken from literature. Again, there were some structural changes that were made depending on how the data fit to the model. Basically, we ended up with this huge, again, monstrous model, I guess.

This slide here provides a little more detail of the various modeling components so you can see that the plasma PK component was a standard two-compartmental model with both linear and nonlinear clearance. The tumor compartment was made up of various sub-compartments and it also included the lymph flow binding to FcRn and partitioning and so on. The receptor binding model was a standard binding model. I don’t have

details of the tumor growth model but basically at the end of the day, after the model was developed, we got really good fits to the mouse model. All of this was mouse data. Sorry. Until now, I've basically described model development for mouse data. The next step was to really do the translation.

Here we had to become very creative. I'll try to walkthrough the various steps that we took. With the PK model, we had PK data. We had developed a PK model with our clinical data. We just plug that in and replace the mouse PK model. The way that the tumor physiology-based PK model was structured and set up, it was pretty much species and variant or independent. Basically, all the parameters depended on volume of the species. As long as you change those volumes, you are okay. You don't have to re-estimate the parameters. That's how we did the translation for that tumor, PB-PK model.

With the receptor occupancy model, we didn't have any human data, so we relied on the model structure from mouse. We substituted out the binding constants from in vitro measurements of the human antibody. Remember, with the mouse experiments, we used a mouse antibody. We had to change out those binding constants for the human antibody. Then that particular component tied into the tumor growth model. Now, the tumor growth model, this is where things got interesting because the tumor growth you have is the standard growth rate of the tumor and a kill rate.

The tumor growth rate was actually tricky. What we did was we actually went and read a whole bunch of papers and tried to guess or collect information on what the typical growth rate was for the particular type of tumor that we were studying. We basically identified a few papers that showed there was actually a range of growth rates for melanoma. We ended up categorizing those rates into three classes. There's a slow growth rate model, a medium growth rate model, and a fast growth rate model.

The second component of this tumor growth model is the kill rate. We had estimated that kill rate using mouse data, but then we wanted to translate that and so there were two ways that we use to translate that kill parameter. One was using allometric scaling and the other was actually scaling it proportionally to the tumor growth rate that we had. At the end of the day, we had multiple scenarios that we thought captured the variability that we might see in humans. We took all of those different scenarios and we simulated out what the tumor responses might look like in human.

Again, just to summarize, we tried to use whatever information we could get to compile and put together this translation model. The physiology-based PK model had a parameter structure that was species-independent and that was really very helpful in doing the translation there. We plugged in clinical data and models wherever they were available. When that was not available, we used allometric scaling and other assumptions to do the translation. In order to capture the variability that we might see in humans, we looked in literature and identified parameters that we could use from literature.

Just a detail that I think might be interesting to focus here, literature sources often

report linear tumor measurements, and our model was done with tumor volumes and so there was a conversion factor that we did. That's shown in this next slide here. Basically, what we did was we made a few assumptions about the linear measurement. We assumed a spherical shape and we assumed a range of one to five tumors and that the tumors were similarly sized, and then performed some algebraic manipulations to basically come up with a very simple conversion factor that helped us go from linear tumor size that was reported in literature to the volume-based sizes that we were using in our model. Then, also at the end of the simulations, to go back to a linear tumor size and growth rate.

Once the translation model was set up and we had all of these scenarios, we performed simulations and calculated the probabilities of falling into the resist classes for a range of doses once-every-three-week dosing schedule. Basically, what you see is that the probability to achieve partial response actually plateaus off after the 2 mg/kg dose. Based on all of this, we recommended that the 2 mg/kg dose was a saturating dose.

As it happens often when you make these decisions, go/no-go decisions or other decisions, we actually have to have evidence from multiple sources. In this particular scenario, we had evidence from the ex vivo [IL-2 acid 00:48:30] that Dinish talked about this morning that pointed to maybe 1 mg/kg dose and this translational work pointed to a 2 mg/kg dose. These various points of information got integrated into the doses that were tested in the clinic of 2 mg/kg and 10 mg/kg once every three weeks.

At the end of the day, after all of the trials and studies, we ended up with the exacted dose of 2 mg/kg once every three weeks. This is actually very exciting because it was a very new paradigm for dose selection in oncology where traditionally you would have selected the 10 mg/kg once every two weeks and now we ended up with 2 mg/kg once every three weeks.

That's pretty much the end of my presentation. I just wanted to highlight two things really. One is that in this particular scenario, we actually went back to preclinical studies design them properly so that we could characterize the dose response relationship and do these simulations after we had clinical data and hints of clinical efficacy.

The second is that we used whatever information we could find. We came up with new ways, innovative ways of incorporating that information into the model and we're still learning. Some of these techniques might be applicable to other programs; some might not. In this particular case, I think we are considering it a successful translational model. I'm going to stop there.

Todd Palmby: The next speaker is Alan Korman, vice president, Immuno-Oncology Discovery, Biologics Discovery, California, and Bristol-Myers Squibb.

Alan Korman: Thanks, Todd. Thank you for the invitation and I'd like to spend some time on mouse models which we have performed which helps us inform the isotype choice for the

selection of immuno-oncology antibodies and what the implications are for dose selection.

We begin this work quite a long time ago focusing on CTLA-4 which the first negative regulator that was well-studied, interacts with the B7 ligands, but we also made a pair of antibodies to PD-1 and PD-L1. Here, actually, the Fc choice, I won't address this; it's quite simple. You really don't want to touch Fc receptors. I will show you we discovered accidentally that at least in the mouse, CTLA functions for antibody's function because of their ability to bind Fc receptors.

Some points I'd like to make about syngeneic tumor models and Fc selection. The antitumor activity of many of these kinds of antibodies is highly dependent on the Fc isotope. That shouldn't be surprising to people that work on antibodies but there continue to be surprises with many of the targets in this field. Through our work and work of a number of other groups, [inaudible 00:52:45] and Novartis and others, this isotype choice is very important for both agonist antibodies and some antagonist antibodies such as CTLA-4. I think the difficulty that we have in translating between mouse and human really is due to the fact that there isn't a true concordance between the mouse Fc receptors and the human Fc receptors. I'd give you an immediate translation. I'll show you how we're trying to navigate that.

Although I want to address that today, there are some instances where human and cynomolgus Fc receptors are also … their concordance is limited in many cases. Again, we need to understand how various antibodies interact with Fc receptors in both mouse and human systems. In the mouse, there are number of activating receptors that bind to different isotypes, the IIB and IIA isotypes, bind well to the activating receptors whereas mouse IgG1 binds to the inhibitory receptor and frequently, antibodies that bind to the inhibitory receptor function well as agonist antibodies.

This is the experiment that set off a lot of thinking about how CTLA-4 functions. When we tested blocking CTLA-4 antibodies and the MC38 colon tumor model, you see that tumors grow very rapidly and here's the antibody that we had used for a number of years. It's a mouse, anti-mouse CTLA-4 that's the IIB isotype. It's similar to some of the antibodies that had been used that were made in hamster. You could see that we see some tumor rejection but tumor growth inhibition. This antibody binds to activating receptors as compared to inhibitory receptors with a ratio of seven. This is a ratio that was defined by Nimmerjahn and Ravetch for antibodies that actually are depleting antibodies similar in concept to say Rituxan for depletion of B-cells.

There are two surprises in this work. One was that when we made the antibody as an IgG2a isotype, we know see that CTLA-4 works as monotherapy in this mouse model completely eradicating all tumors, again, because it binds to the activating receptors more tightly. Equally interesting was the fact that just a simple blocking antibody actually has no function at all. If you recall, the function of CTLA-4 and B7 interactions are inhibitory and just in the tumor model in this case, you do not see any activity of just that simple blockade.

Here's the only PK I'll show. Each of these antibodies has similar PK in these models. There's an allotypic difference in this strain where the IIB ends up becoming immunogenic.

How do we explain this? In the periphery, when you block CTLA-4, you're blocking the tonic inhibition of B7, CTLA-4 interactions and that can be observed in Tregs by their expansion in the periphery. You'll see that each of the isotypes causes the Tregs to expand. When you look at what's happening at the tumor site, you see that the IIB and more important the IIA is very efficient at removing the Tregs which are defined by Foxp3 that suppress the effector responses. In concert with that, we see activation of CD4s and CD8s resulting in very high T effector to Treg ratios. This occurs because the only cell at the tumor site that has a lot of surface CTLA-4 are the Tregs. The other cells have CTLA-4. They are expressed in intracellularly small levels. What happens is that the Tregs are depleted. Blockades of CD4s and CD8s are important but those cells are spared leading to a good antitumor response.

The CTLA-4 IIA is highly potent. It's much more potent than the IIB isotype. Here, we could see good antitumor activity perhaps equivalent to that IIB at 140 at the level. You could see the dilemma here of looking for efficacious doses when you have different isotypes.

How can we understand this a little bit better? What we've done is to look at the function of the surrogate antibody with human Fcs using a mouse developed by the laboratory of Jeff Ravetch in which the mouse FcR genes have been removed and the human FcR genes have been replaced. As I mentioned, there's not an exact concordance between these Fc receptors. You have activating receptors, FcR3A, the most dominant one but you have two FcR2, one is inhibitory, one is activating. One needs to understand how various antibodies bind to these different Fc receptors.

For this experiment, we took our mouse surrogate variable region, grafted it on to human IgG1 which, as I told you in the case of Rituxan, is a depleting isotype. We also made two different versions which actually bind to CD16 better. This is non-fucosylated derivative. These Fcs which have low fucos are known to bind to FcR3 or CD16 more potently than G1 or these amino acid variants that increase that binding as well as increased binding to other Fc receptors.

What you can see in this tumor experiment in MC38 is that if we use the G1 variant [HER2 01:00:17] dose titrations, you see that you have some antitumor activity but the antitumor activity is superior with those non-fucosylated or amino acid variants at these two different doses. This mimics what we're seeing with the IIA versus the IIB isotype except there are the human Fc regions.

I won't spend too much time here but the same concept applies here in this tumor model. You see reduction of Tregs at the tumor site expansion and the periphery and expansion of 4s and 8s. Again, if you plot out some of the tumor growth inhibition versus the doses that you use, you can see there's about 10-fold difference in potency; again, with the fact that the human IgG1 will never reach maximum potency the way

the higher FcR binding.

How do we understand this in man? We actually don't know at this time. There's anecdotal data that ipilimumab depletes Tregs; however, we do have a clinical experiment. Ipilimumab was made as a human G1 and there was a competing antibody, tremelimumab, developed by Pfizer at the time; that's human G2. G2 antibodies do not bind well to Fc receptors but interestingly, we actually thought that one could view this positively in favor of ipilimumab and that can bind to Fc receptors. If you look at the clinical trial data and response rate, the early data, actually both of these antibodies are capable of inducing responses and durable responses in melanoma. They both cause the adverse events they're seeing with blocking CTLA-4.

Perhaps there isn't as much difference between these two antibodies as we first imagined. That can be seen in some comparison of long-term survival data from ipilimumab. These are some 3,000 patients from the expanded access trials showing long-term durability. From the failed phase three trial of tremelimumab, an analysis of some of those patients also shows long-term durability. What this suggests to us is that perhaps what we thought might be a depleting isotype in solid tumors, the human G1 is perhaps not very effective.

Consistent with that, our early observations that we made where we looked at polymorphisms in the Fc receptor and so let's just focus on FcR3A or CD16 where there are known polymorphisms that enhance the binding to these Fc receptors. These have been shown to be important in the activity of rituximab. What you can see here is that there really is no difference in survival in these metastatic melanoma patients treated with ipi from the pivotal trial of ipilimumab when we look at these different polymorphisms. I think this data is consistent with our vision of thinking about the role of human IgG1 in solid tumors and its depletion capability.

I'd like to move on a little bit to just address some data with regard to the co-stimulators on nearly every molecule on the surface of the T-cell is now the subject of clinical investigation. Many antibodies against co-stimulatory molecules have been to the clinic. Bristol and Pfizer have developed CD137 co-stimulatory antibodies, [Gitter 01:05:07] by ourselves and Merck and many groups with OX40. I'll show some example with a co-stimulatory antibody but without disclosing the target.

However, if you look at tumor models with respect to the expression of these co-stimulators [inaudible 01:05:34] and you look at them compared to their expression on effector cells versus T regulatory cells. What you find, as I showed you for CTLA-4, is that most of these co-stimulators are actually higher on Tregs than they are on the CD4s and CD8s which is different than many negative regulators, the important one being PD-1. LAG-3 and TIM-3 share that property of being highly expressed on CD8s but not on Tregs whereas some other negative regulators are more similar to CTLA-4 such as [TGF 01:06:18], for example.

Here, we've taken an antibody to a co-stimulatory agonist and we've made it as two different isotypes, either the depleting isotype, the IIA. We're now back to mouse Fcs

or once isotypes that bind well to the inhibitory receptor and are co-stimulatory. If you look at this tumor model, CT26 and look at dose responses using these two isotypes, what you find is that the IIA isotype is superior in activity, many complete regressions, very good tumor growth inhibition on this particular day of the growth.

However, you do see activity just of agonism. Tumor rejection, tumor growth inhibition do solely to agonism of that target. If we look at the tumor site, we see in comparing these two isotypes that … Let's look at the CD8s again across this dose response, you do see activation of the CD8s that appears to be dose-dependent in some cases. As I showed you for CTLA-4, you do see a potent Treg depletion at certain concentrations. Again, we're looking at the tumor Foxp3-positive cells depletion using the IIA isotype but not using the activating isotype even though each of these antibodies can cause tumor growth inhibition.

We understand that this T effector to Treg ratio would be superior to this one; however, there is still antitumor activity. Again, if you look at conventional CD4s, of course, they are higher because of the laws of the Tregs. Again, these Tregs are more activated; another marker of their activation.

We can also look in the periphery and see some pharmacodynamic effects so for example, looking at activated PD-1, CD8 cells, the two way agonist which, again, will deplete but also agonize causes activation of cells in the periphery as does the pure agonist antibody in a dose-dependent manner.

One of the other issues though that confounds the choice of efficacious doses is that it is also dependent on the tumor model that you've chosen. In the previous model, we had very good antitumor activity that was CT26. CT26 is a colon adenocarcinoma that's known to have a high frequency of tumor antigen-specific T cells, is highly Treg dependent, it is clearly a more immunogenic model than MC38. If you look at the activity of these two different isotypes in MC38, at least at these tumor doses at which when we begin treatment, you see that there's very modest activity of the agonist antibody.

One could conclude from this that the efficacious dose here would be much higher that than of C226. If you look at the IIA, there is clearly more activity because in general, these subcutaneous models are very responsive to Treg depletion but you still do not see activity that's equivalent to that observed for CT26 thereby leading to a different calculation for the efficacious dose.

We're left with this emerging knowledge about how Fcs [function 01:11:07] at the role of these antibodies, how to convert them to human activities. Again, what we know very clearly now is that the isotype matters. I think the epitope also matters because depending on the epitope, you may have different efficacies of antibody-dependent cytotoxicity. Again, the choice of the tumor model will give you a different efficacious dose calculation making it difficult to predict what would be the case in man.

Again, we need to understand how these different isotypes function in man with

respect to their immunomodulatory activity. I think for PD-1, this is not difficult. You don't want to engage Fc receptors. We've published that in mouse models if you make PD-1 with a IIA isotype, you actually lose activity. There does remain the question of what the human G1 does in solid tumors even though clearly for hematopoietic lineage depletion, human IgG1 is an efficacious depleting isotype.

Again, the comparison with mouse also leads to the question of with the subcutaneous models where you're, in a sense, creating an inflammatory [inaudible 01:12:51] where you get a lot of infiltrating FcR-positive cells. The question of what kind of effector cells are present in human tumors is relevant to the mechanism of action of these antibodies.

With that, I'd like to thank my collaborators, long-time collaborator Mark Selby and John Engelhardt who worked on the CTLA-4 work and the co-stimulator, the work of Mike Quigley, Bill Barnhart, [Patrick Enuldom 01:13:27], Maria Jure-Kunkel so thank you.

Todd Palmby: The final speaker in this session is Juliet Williams, Head, Oncology Pharmacology Drug Discovery at Novartis Institutes for BioMedical Research.

Juliet Williams: Thank you and thank you for the invitation to speak. On behalf of Novartis, I'm going to present to you today data from the work that we've been doing in mice to try and predict human clinical trials trying to model inter-patient response, heterogeneity in PDXs.

What's the problem? The problem is as we all know here that less than 10% of drugs that we actually put into the clinic actually get approved. This is the worst failure by therapeutic area. The cause of failure is largely that we see a lack of efficacy. Why are we so bad at predicting what drugs we put into clinical trials actually will have an effect in man. I'm sure you've all got your theories and here's just a couple.

One, cancer is very complex as we can see by this diagram from Hanahan and Weinberg. Of those many complex mechanisms, they are often an aberration of normal function so we have very small therapeutic windows. There are also lots of different mechanisms in which tumors can evade therapy. Another reason that we're possibly so bad in predicting if our compounds are going to be efficacious in man is the limitations of preclinical cancer models, and we've touched upon this a little bit already today.

We grow cancer cells in dishes. These are cancer cells that have been selected because they like to grow on plastic. We put these cells in mice so we take bits of tumors directly from humans and we put them into mice, we put them into immuno-compromised mice so that they'll grow. Here again, we have this system which doesn't fully recapitulate the human situation of human tumor growth.

In this particular xenograft models, of course, we don't have a fully intact immune system so then we go on to use mouse models so mouse tumors in mice. Here, we

have mouse cancer which doesn't fully reflect human cancer in an immune system which is a mouse immune system which doesn't fully recapitulate the human system.

The added issue that we have with this in generic models is we don't actually have very many of them. Why is this an issue? This is called stuck. This is a problem because as we all know cancer isn't just one cancer. Cancer is many cancers. In this box here, it should say soft tissue sarcoma. This is just an illustration if you do pick a cancer you can then subdivide that particular cancer into many more cancers. Here, histologists has looked at what is labeled as soft tissue sarcoma and come up with 19 different labels for that particular cancer, each which could potentially respond differently to a different drug or a subsection of each could potentially respond differently to a different drug. Really, the aim of what we had been trying to do is understand the pharmacological response in a range of models.

At Novartis, we have several tumor banks. We have a cell line bank which has over a thousand cell lines very well characterized. We have our PDX tumors which I'm going to discuss today. These are tumors that we've taken directly from patients and placed into mice so they haven't touched plastic, and they've enabled this mini clinical [inaudible 01:18:11] which I'm also going to describe today.

We've been trying to humanize the immune system for IO studies in these mice, again, to try and recapitulate the whole range of tumors that we get from these models and then we have these in generic allograft initiative which obviously enables this [inaudible 01:18:31] and immune system studies in intact host which other models don't allow.

We started collecting these tumors in 2006. We started collecting them for developmental pathway projects were we knew if we placed tumor lines that we're expressing, the hedgehog pathway or the Wnt pathway. They would lose their dependency once on plastic.

We started collecting many more models and we thought we would need for these projects and we were able to in the event of time very rapidly characterize them so the [genomeX 01:19:16] using DNA-seq, RNA-seq, DNA copy-number and also look at the histopathology.

Do our large collection of PDX tumors actually represent what we see in patient tumors? We found that genetically they really do. Here, we're looking at somatic mutation frequency in PDXs and just like you would see in human tumors, you will have low rates of mutation in soft tissue sarcoma. You have higher rates of mutation in colon cancer reflective of microsatellite instability. The typical base substitutions that you see in lung cancer, for example, C to A, typical of tobacco use and in melanoma C to T, typical of you being exposed to UV.

When you look at the genetic level, when you look at mutations in genes particularly genes and also on the pathway level, you see that the PDX is very closely mimic what you would see in patient tumors. The cell lines do as well; however, the PDX do have a

better presentation of what you'd see in patient tumors. You'll find that in cell lines you'll get a higher amount of mutations in cell cycle genes.

How did you then use such a vast array and collection of PDX tumors to actually look at the effect of drugs? Obviously, you can't use them in the typical manner of any closed 8 or any closed 10 in particular studies. What we do is we use what we call the 1x1x1 approach. This is where you have one tumor model in one mouse that's treated with one drug or a combination drugs or regiment in drugs. You look at a whole population of models and you collect data much like you would in a human clinical trial where you look at responses, so this is best average response. You also look at a sort of [pseudo 01:21:34] survival curve. This is probability of being progression-free.

We screen in an unbiased manner and the reason we do that is, first, so you have a negative control. Secondly, we always think we understand the biology but certainly we've seen throughout these trials that there's a lot we don't know about biological systems and we do get surprises in patients or xeno-patients which respond to drugs which we would not have predicted.

Obviously, with our xeno-patients we can enroll the same patient into multiple trials and have a look to see how they may respond to different drugs and different combinations of drugs which you obviously can't do in the clinic. In our first two trials, we looked at 62 treatments, single and combinations across six indications.

Just briefly, how do we pick the doses to go into these mouse clinical trials? We pick the doses based on work that we've already done in workhorse models much like what have been shown before with PK-PD efficacy response relationships in a handful of models. Once we've worked what efficacy we'll likely to see with a certain PD response in a few models and also typical schedulings that we might imply, then we also question what does this look like when we put it across a whole population. We can ask different questions so we can ask the biological question, what happens if we hit the pathway as much as possible so 100%?

Often, we do have PD assay limitations so sometimes that forces us to say, "Right, let's test this at the highest dose possible where we're inhibiting the pathway as much as possible." That's more of a biological question. Then we also look at doses and scheduling that we either know is occurring in the clinic although we think we're going to achieve in the clinic.

After we did a couple of these trials, the first thing we obviously ask was how reproducible are these trials? This is looking at reproducibility with the same trial. These two compounds here are structurally distinct PI 3-kinase inhibitors. If we look on a tumor-to-tumor basis here, you can see that we're able to get the same response with both of these agents even though, remember, we've just got one tumor here being represented.

We're not actually that interested in what happens into one individual tumor. What we're interested is the population response. When you look at the population level you

see when we look at our best average response and also our probability being progression free that the reproducibility is very good. We also look at reproducibility of cross-trials. Here, we have a combination of a CDK4 inhibitor with an M2 inhibitor. If you look at the individual tumor response in the alpha trial which was done a whole year before the beta trial, you actually see that it is very reproducible on a tumor-to-tumor basis.

Again, we're not that interested in an individual response. What we're really interested in is the population response and on a population level both by best average response and also by probability for being progression free, it's very reproducible.

As I said before, to find the drug, to locate our lead optimization phase, to understand what dose we will possibly go into the clinic with and scheduling, we'll use our workhorse models. Then we'll take that dose or doses usually and different schedulings that we want to try out and put them across our panel of tumors. This is an example with BYL; this is a PI 3-kinase alpha inhibitor.

What you'll see when we put this particular inhibitor across a whole range of models, you'll see that not all the models responded but we did have a subsection of tumors that responded in each of our indications. You'll also note that the response isn't that great. It's stable disease maybe a little PD at best. Then the question is can we predict what tumors will respond and because the efficacy isn't that great, we also ask the combination question, how can we increase the efficacy by doing combinations?

Working cell lines had showed us that if you have a gain of function mutation in PIK3CA that was a positive predictor response. If you had a P10 mutant, that was a negative predictor of response and that's exactly what we see in our mouse clinical trials where you have more probability for having efficacy in tumors that have a PIK3CA gain of function and less of a probability if you have P10 mutant. You'll see here that we also had quite a few wild-type tumors that responded and we believe this is due to some RTK activation.

Talking about how to predict what tumors might respond or might not respond, I'll just show you an example here of a CDH6-ADC. This is an antibody that has a cleavable linker with a DM4 payload. Here we were trying to look out what threshold of expression we needed to have to see activity. You can see the power of the mouse clinical trial when you have a whole range of tumors incredibly heterogeneous for expression of CDH6. From here, we were able to predict what kind of expression even though it's heterogeneous expression you might need to see an efficacious effect, and this is being tested in the clinic as we speak.

Back to BYL, I showed you before that as a single agent, there is some efficacy but it's not really that impressive when we look a whole population. Then we decided to do combinations to try and increase the efficacy that we see. We picked nodes in the pathway in the vertical pathway here that if we combine with BYL, we may potentially be able to increase the efficacy so with an M2 inhibitor or an RTK inhibitor and even with CDK inhibitors. You see that we were able to increase efficacy. LEE has proved to

be a very good combination partner for many pathways.

How translatable is this? At the moment, we've just got data showing that retrospectively for targeted agents, it's actually we believe very translatable. This is just an example here where we have a RAF inhibitor. What we're seeing in our mouse clinical trials is an overall response rate of about 70%, so about 70% of our BRAF mutants and the patients will respond to a BRAF inhibitor. That's roughly what you see in the clinic. You also see that none of NRAS mutant patients or xeno-patients are responding and that's what you see in the clinic.

Then as you combine with MEK inhibitors, you can increase that response rate and in our mouse clinical trials, we see that you've increased that response rate to nearly 100% overall response rate in the clinic when you have a combination of a RAF and a MEK in BRAF mutated melanoma patients that's what your overall response rate is. Similar with CRs and PRs, we get about 70% response rate which is also typical of what you generally get in the clinic.

We were then interested to see if we could expand the combinations that we did with our BRAF inhibitor and this is in in vitro. Here, we looked at 500 combinations across 42 melanoma cell lines and we created these synergy schools and hits are basically cell lines in which scored a synergy score of greater than two. Then we plot it the number of hits to try and rank order what our combination partners would be.

You can see that our RAF inhibitor with a PI 3-kinase inhibitors scores pretty high up here but you'll see there are RAF inhibitor with a CDK4 inhibitor didn't actually score very well. We were intrigued by that because rationally it should. When we went to our mouse clinical trials, we decided that we would add a RAF inhibitor with LEE or CDK4 inhibitor. First of all, here you can see our RAF inhibitor with a PI 3-kinase inhibitor increase the efficacy that you can see in melanoma models. We also saw the LEE CDK4 also increase the efficacy in our melanoma models which was not predicted by the in vitro work.

We started to get interested to look at that difference between the predictions made in vitro to in vivo. One that stood out very clearly was IGF1R. IGF1R inhibitors have typically been very successful in the preclinical space scoring very high and working very well in in vivo models and in vitro models. In our hands again, IGF1R inhibitors do perform very well in our preclinical models as you can see here scoring top when we were looking at combinations in colorectal cell lines. This is an IGF1R inhibitor with a MET inhibitor. However, when we took this into clinical trials, it really failed to show any benefit at all.

We saw the same issue or the same results when we went in vivo in our mouse clinical trials in pancreatic cell lines although in vitro in our pancreatic cell lines, it scored very high as a therapeutic in combination. As we all know, there is data in vivo as well if you take select models. You can show the IGFR inhibitors with MET inhibitors can have some benefit. As you also all know when these have gone into clinical trials, there's only select indications where IGFR inhibitors have done very well. They haven't done

well, for example, in colorectal cancer.

To summarize, I've shown you the PDX models largely recapitulate patient heterogeneity histologically and genetically. We believe that this mouse clinical trial approach does model the inter-patient response, heterogeneity, and could potentially improve our ability to predict response of patients. We have shown that it's highly reproducible. Retrospectively, it's very translatable. It can be used for the validation and discovery of predictive biomarkers. It facilitates the evaluation of the same treatment across patients. It also enables evaluation of multiple treatments in the same patient or xeno-patient. Obviously, we haven't yet shown prospective translatability but that's what is currently being assessed.

With that, I'd really like to thank a very large team in particular Hui Gao who is my lead PhD on this and Bill Sellers, the head of the research department who is a great advocate for these trials. Thank you.

Todd Palmby: I'd like to ask all the speakers and panels for this session to come on up and grab a seat at one of the tables up here. I think while they're piling up here, I'll just make a comment that last year when we were devising the themes and the idea behind this workshop including this nonclinical session, a lot of it was intended to bring all these different disciplines into the same room in the same conversations around the problems with oncology dose finding. It actually ended up working out very well.

One of the themes that came out of the nonclinical session from the safety perspective last year was going backwards once you've obtained some clinical data. You've done your initial safety studies in animals and you've gotten into clinical trials. You've gotten some pharmacokinetic data and maybe some pharmacodynamic data. You've gotten into patients and then you run into some issue that really needs to be addressed. There were a number of cases last year where development programs were able to go back to nonclinical models to try to address those and have managed those and move forward in the clinical setting.

It was actually very fascinating to see the same theme emerge in a lot of these presentations and these development programs that were just discussed with the efficacy models and going back to address specific questions and specific issues around dose finding from an efficacy perspective as well. It actually brought to mind one of the little nuances coming to FDA in the pharmacology –toxicology division that we work. We always refer to these studies as nonclinical studies. The clinicians especially often refer to them as preclinical and they ask, "Why do you call them nonclinical?" and we say, "Because a lot of them are actually done concurrently with clinical trials. They're done to address specific issues. They're done during the clinical development. It's not all just …"

The standard thought is really that you do all these preclinical or nonclinical studies as a requirement to get into your clinical trial. You do your efficacy studies really as proof of concept and then once those are completed, you're done. As we've seen, that's not the case in most cases actually. I think more recently that those nonclinical models are

being flushed out and evaluated more carefully how they can be used to address specific questions are being evaluated very, very closely so that was really interesting to see.

I'd like to thank all the speakers for excellent talks setting up this discussion. I think in an effort to try to save time, I know we were supposed to have 30 minutes, maybe we can see how we're doing after about 20 and try to catch up some time because we're running … We're supposed to be done by now.

I'll get us started and then maybe we can take some questions from the audience and see how that goes. I thought I'd actually start by asking some of those on the panel who represent those other disciplines not necessarily the preclinical or nonclinical translation side of things but maybe the clinicians and the statisticians and maybe the patient representative to give their thoughts on the use of these nonclinical models as was described in some of these development programs.

From the clinical perspective, do we have confidence at this point in time that we've learned more how to use these models appropriately and that it really can add benefit to the clinical development which may not have been the case 10 or 20 years ago? Then from the statistical side of things, are strategies that can be employed? We heard some of those last year around safety end points and including prior information obtained from nonclinical models in statistical designs for early phase clinical trials rather than the standard three plus three MTD-based phase one trial.

Then from the patient perspective also are there specific issues or questions in the use of these drugs specifically around dose selection and the standard approach is more is better and that's often hard for the patients because that also means more toxicity and if that isn't necessary. Are there specific issues from a patient perspective where these nonclinical models, you can see they might add value. I'd like to just kind of let anyone volunteer or I can call … yes. I'd also like to ask, in an effort to save time, I won't do introductions for none speakers so if you could just introduce yourself the first time you speak, please.

Mark Ratain: I'm Mark Ratain from the University of Chicago. Obviously, pre, peri, post clinical models are important in drug development. I think the question now is what's their role? Obviously, they're critical to demonstrate that a drug has efficacy but I think from the standpoint of quantitative work making quantitative predictions, I didn't see any data today that would help me from quantitative perspective to use any of the drugs. I was most interested in the EGFR inhibitor because it's the most complicated. It's got two pharmacologically active metabolites. It's supposedly selective for the mutant, the parent drug selective. One of the metabolites isn't selective at all; one is less selective. We have issues of how interspecific differences in metabolism, interspecific differences in protein binding. I'm just not sure how one can make a prediction as to what the right dose is given all the complexities.

Todd Palmby: Okay. Thank you. Any other perspectives from the clinical side before we move on? Maybe the statistics representative can address?

Thomas Jaki: Thomas Jaki, Lancaster University. I'm talking about the statistics background. I just want to say that basically from what we've seen in the talks, I think it's clear that there is use to these nonclinical models. Just in terms of understanding how the drug candidates and combinations work, for example.

The other use that has been illustrated was confirming that the dose that should then be taken forward to further studies in clinic. The thing that I was noting and wondering why there wasn't an integrated approach that rather than looking at, okay, here is my clinical data that suggest I should use a certain dose, a certain schedule; here are my nonclinical data that hopefully confirm exactly the same dose, the same schedule, why on the data analyzed jointly in combination and thereby again gain even better confidence about what the right dose should be.

The comment was made earlier that if safety is the objective, one could think about forming [inaudible 01:42:26] from the preclinical nonclinical data used and then in the dose escalation study. The same thing exactly same idea could be used in the context of modeling and looking at efficacy but somehow this wasn't at least described here in the talks today.

Hans Loland: Hans Loland, patient. One of the things I wanted to expand on what you said was the overdosing or more is better. I have CML and the line of treatment has something like six approved drugs and then virtually all of those drugs are now being administered lower dosage than what was approved. I think it shows us there was a rush to get to a very high dosage for efficacy but then throttle that back when you're looking at long-term risk because all of these drugs were on very long term.

The other thing I would challenge is think about this differently as when you design a trial, instead of simply looking for dosage and efficacy is maybe think about a response dosage. If you reach a certain response, maybe you back off the dosage rather than having a one model fits all because I think we're seeing that at least in the CML community.

Karthick Vishwanathan: Karthick Vishwanathan, AstraZeneca. I just want to address one of your points, Mark, where you mentioned about why we haven't taken the metabolite information and everything else into the model that is taken into account. I think Darren did mention briefly with regard to how the metabolite information is taken into the … not only the PK model but as well as the PD model. From a human point of view, the metabolite is only 10% of the parent so it is not contributing as much in humans as it might be contributing in the animal species. There is some complexity involved in it. In the 15-minute talk, it is very difficult to bring in all of that element. We can talk about that.

With regard to the integrated piece that you were mentioning about, typically, that integrated piece is how we justify the dose so it is not just from an efficacy point of view or safety point of view or nonclinical or any of the cell line models or anything. It is always a combination of all of that to come up with an optimum dose, a biologically effective dose. I think [inaudible 01:44:45] talk this afternoon would be [inaudible

01:44:47] with regard to bringing it. It's not just one discipline or one action that is coming and some … It makes it very difficult to just focus only on nonclinical to come up with that particular element of it.

However, one of the things which was interesting with regard to that is four different areas, immuno-oncology, one as a small molecule area, one as an Fc area, and the PDX model development. All of them integrated with regard to how nonclinical models are mostly used after the human PK is validated. Then going back and retrospectively figuring out is this the right dose? Is there something else that we could do from dose or schedule or combination point of view that will help us understand what we see in the clinic and then how we can improve that further.

The days of how I have seen is more qualitative so initial evaluation of the nonclinical models is giving a qualitative idea on where we can see the efficacy, where it is likely to be target engagement is likely to happen, and then we do more and more. Once we get the human PK data, then you can integrate it then validate it and say, "Okay, is my human PK within the prediction that I anticipated? Is my PD data within the reasonable expectation or if it is modulating the target, et cetera." That's how I would use the nonclinical data as presented.

Amit Roy: Thank you. Amit Roy. I'm the head of oncology pharmacometrics at BMS. I wanted to actually follow up on Thomas' comment; first, some reflections on the talk. In terms of utilizing the data from nonclinical studies as a prior for efficacy, I think that's a very challenging concept or idea in part because if you're just thinking in terms of [Bayesian 01:46:48] theory, you think [inaudible 01:46:50] assumes that it really is information from the same system. In fact, we've seen a lot of data to day that suggest even within an animal model using one cell line versus the other cell line using an antibody, one maybe very close affinity to the receptor but which is the different Fc receptors can have quite different results.

I think in a [inaudible 01:47:19] I'm sure I'm on the same page with [inaudible 01:47:22]. I think we are … The best use of this information is largely qualitative but we're getting more and more quantitative as we go along. Rather than use it as a pure or [Bayesian 01:47:33] translation, mechanistic translation scaling up from what we understand from animal modeled perhaps with, let's say, density of receptors and such ratio of Tregs to effector regs, those kinds of things maybe more informative too actually making the translation.

Todd Palmby: Yes, please.

Thomas Jaki: I don't agree that information from a [inaudible 01:48:00] needs to come from the same system as you call it. Plenty of applications in the industry use prior data, prior information that may or may not have relevance for the question that they are trying to answer [bridging 01:48:18] studies, extrapolation from adults to children, things like that.

Yes, they may be relevant but they may as well not be at the same concept applies to

different system that maybe relevant but they may not be. I'm not disagreeing that mathematically. It's challenging but I think it's possible to do so the important part is going to be to make sure that if there is a conflict, if the data aren't relevant, you discard them; you discard them quickly. They don't influence the further decision at the end of your study anymore. At the same time if you find that they are relevant, you use them. You make use of the opportunities to get more precise estimation at the end of your study.

Todd Palmby: One of the things that actually again came up with regard to safety and came up in the morning session this morning is relating to schedule. We often have many questions about the schedule and how the schedule is decided from the FDA's perspective because that often occurs during that period of time where we may not have frequent discussions with a sponsor during development. Many times, we question whether an intensity of schedule is really needed and what data supports that.

I just had a general question about can nonclinical models really inform the clinical schedule that should be used from an efficacy point of view and then in particular or in addition, can inform the schedule and sequence of combinations? Alan.

Alan Korman: Yeah. I'd like to address that question. I think the subcutaneous models are really not very useful for scheduling. They grow too quickly. You have little opportunity to re-dose in a way that's effective once the tumor burden has gotten too high.

Similarly, for scheduling of different therapies, we've explored the combination of CTLA-4 and PD-1 in these models, and we only saw activity when they were giving concurrently and we lost activity when they were given sequentially. We know that in the clinic, these drugs can be given sequentially. In terms of safety, we also know very well that with respect to immuno-oncology therapeutics both the mouse and the xeno models have under-predicted toxicities nearly in all cases.

I think they give the flavor of activity of these molecules and can be used for exploring combinations, but it's important not to extend them beyond what they're capable of.

Todd Palmby: Just to continue on that too. One of the themes that has come up that we have discussed internally is a lot of these questions and concepts that we're talking about here may differ from a small molecule to a large molecule and then even within the large molecule may differ between if it's a tumor-antigen-targeting antibody versus an immuno-oncology. They may be very different responses or answers to these questions. Maybe if someone has any comments about a non-immuno-oncology answer about scheduling and whether that might be a more translatable approach to use nonclinical models for, say, a small molecule.

Darren Cross: I can just say from our experience using nonclinical model [inaudible 01:52:27]. We have used xenografts to explore different schedules like intermittent schedules and although there are limitations of those models that has given us the … It helps validate the whole preclinical modeling piece [inaudible 01:52:40] pressure test and with different schedules but also understanding that biology for those intimate in

scheduling studies has told us how we need to be hitting the EGFR target and therefore, those help us maybe the appropriate kind of schedules to take forward.

I think in our experience that the work we have been doing in xenograft models has been very informative. It has told as you probably need to have very constant coverage against receptor given actually the receptor, et cetera. I think in our experience these models have helped us in terms of … At the moment, due to time, we haven't thought about it but we are now going back again and based on the clinical information asking the question, "Can we now explore other alternative scheduling [inaudible 01:53:31] or intermittent schedules based on the data we're now seeing from the clinics?" and feeding that back into the preclinical models. These are pieces of work we are doing at the moment. I think there is some option [inaudible 01:53:43] models to help inform that.

Juliet Williams: Yes, I agree. We found the same thing. You can ask the biological question, do you need to have continuous target covering or can you just hit the target hard and then know that may have consequences of, say, cell death in which you can potentially use intermittently rather than continuous. I also agree that you can't get into details about a scheduling in these models but you can ask the broad question.

Mark Ratain: Ask a follow up question.

Chandni Valiathan: I think just to add on to that. I think in general the way to look at these models is, again, a piece of evidence that goes into the collective piece of evidence that ends up helping us make these decisions. Because of the limitations that we have with the preclinical models that we're using currently, there's only so much you can do in terms of your understanding the mechanism or understanding what is translatable.

The general sense is to get an idea of the overall picture some of the things that Juliet mentioned but a lot of the details end up changing when you move to the clinical space so it's one piece of evidence, I guess.

Mark The follow up question I had was a big challenge, I think, in the translation especially with respect to schedule is the difference in the dynamics of how the tumor in an animal model would grow compared to human tumor. I was wondering whether or not … You've looked at biomarkers as a way of bridging that difference in time scale of tumor growth. Let's say for example with the phospho-EGFR or phosphor-AKT if that could be used as a way of bridging the efficacy in animal model and humans.

Darren Cross: Again, in terms of …

Mark Ratain: Rather than looking at schedule and looking at alternative schedules and looking at the tumor response in an animal model to look at an intermediate biomarker so make the link between exposure and the biomarker and the biomarker and the tumor response to animals and use that, leverage that information in humans.

Darren Cross: Yeah. [inaudible 01:56:16] as we're doing at the moment. Again, like I did touch on that

showing how we're using not only the efficacy but again, the advantage for the EGFR these xenograft models do describe the biology of EGFR mutant disease very well so we can use those models to also look at the biology of the receptor [inaudible 01:56:32], and how that then links downstream to pathway inhibition feedback in such.

We are doing those more detailed PD studies at the receptor level as well as the downstream seedling and how that then links to efficacy. That's where I see I think, like we've said, that the understanding the biology at the receptor level even does allow you to then get insights into potential schedules that may or may not work. That's where we know for EGFR given the nature of the turnover and the rate of turnover, et cetera, you probably need to get that constant cover and then therefore on the back of that [inaudible 01:57:06] use those models to see can we actually move to a more high … approach but also building [inaudible 01:57:14] downstream target and PD activity and so I think … It's a very holistic approach.

Karthick Vishwanathan: I think from an efficacy point of view at least we are able to try to do something to understand the biology behind it and then whether dosing or schedule or how it may affect but where the gap is on the toxicology end of it when you combine two components which are two combination agents. We don't do any toxicology studies with the combination agent. We have the individual agents and then we go into the clinic and then they are safe and then you find out from the clinic that the combination is actually having some of the toxicities which we haven't seen before. That's when how can we have preclinical models that is actually predicting some of the toxicity concerns that we see that is making them not combinable.

That's the gap I see as a big one. From an efficacy point of view, at least we understand some biology behind it as to why we need to combine these two agents. The remaining portion of what is happening to some of the other organs at a higher dose or how it may translate that gap exist.

Female: This would be our last comment for [inaudible 01:58:27].

Female: [inaudible 01:58:31] from Merck and actually I'm in preclinical safety assessments. I like to respond to your comment. I'm surprised by this comment because we consider toxicology studies and most of the time, we can predict that there won't be either additive or synergistic toxicology based on mechanism based on what we know already from single agents. There is some even literatures showing that actually combo studies do not add much information in term of discovering new, unexpected toxicity; that's one aspect.

The other, toxicology studies are done in healthy animals and that's another very different model. I think we need to be very careful about proposing, doing more toxicology studies because it seems like the data do not support that we need much more. Of course, we need to know the [mechanistic 01:59:37] pharmacology and so on and so on. In some cases, perhaps it's needed but not generally.

Karthick That's exactly what I was referring to, say, especially when you combine an immuno-

Vishwanathan: oncology agent with a small molecule agent. We are seeing some toxicities which are not daily predictive previously, let me say. Also, when you change the dosing schedule, we have different schedule in toxicology studies compared to what we might study in the clinic and how to translate that.

Todd Palmby: Thank you very much for the excellent presentations and discussions. We need to [inaudible 02:00:12].

Female: [inaudible 02:00:15]