Assessing the Credibility of Computational Modeling and Simulation in Medical Device Submissions, Final Guidance
January 11, 2024

Moderator: CDR Kim Piermatteo

CDR Kim Piermatteo: Hello everyone, and welcome to today's CDRH webinar. Thanks for joining us. This is Commander Kim Piermatteo of the United States Public Health Service, and I serve as the Education Program Administrator in the Division of Industry and Consumer Education in CDRH's Office of Communication and Education. I'll be the moderator for today's webinar.

Our topic today is the final guidance titled, Assessing the Credibility of Computational Modeling and Simulation in Medical Device Submissions, which was issued on November 16, 2023. This guidance provides a general risk informed framework that can be used in the credibility assessment of computational modeling and simulation used in medical device regulatory submissions.

During today's webinar, we will discuss how manufacturers can use the guidance to show the computational models used to support regulatory submissions are credible, provide recommendations for using the FDA recognized standard from the American Society of Mechanical Engineers, V&V40, titled Assessing Credibility of Computational Modeling Through Verification and Validation, Application to Medical Devices, to show that computational modeling and simulation is credible, and answer your questions about the final guidance.

Before we begin, I'd like to provide a few reminders for the webinar. First please make sure you've joined us through the Zoom app and not through a web browser to avoid technical issues. Second, the intended audience for this webinar is industry. Trade press reporters are encouraged to consult with the CDRH Trade Press Team at CDRHTradePress@fda.hhs.gov. And members of national media may consult with the FDA's Office of Media Affairs at FDAOMA@fda.hhs.gov. And lastly, we look forward to interacting with you during the live question and answer segment of today's webinar. If you have a question, please wait and raise your hand at the end of today's presentation to get into the queue.

I now have the pleasure of introducing our presenter for today's webinar, Dr. Pras Pathmanathan, Senior Scientist in the Division of Biomedical Physics within CDRH's Office of Science and Engineering Laboratories. We'll begin with a presentation from Pras and then field your questions about our topic.

Thank you all again for joining us. I'll now turn it over to Pras to start today's presentation.

Pras Pathmanathan: Hello, everyone. Welcome to this webinar. I'll be talking about this final guidance--Assessing the Credibility of Computational Modeling and Simulation in Medical Device Submissions.

The learning objectives for this webinar are first to define what we mean by computational modeling and simulation in the context of this guidance, state the scope of the guidance, describe key points and the approach taken by the guidance, and outline a framework provided in the guidance for performing credibility assessment of computational models.

So let's start with what we mean by computational modeling and simulation and the scope of the guidance.

This figure illustrates that there is a spectrum of different types of mathematical model. On the left-hand side is the case of where there's a lot of data available on the system being modeled and it's possible to fit a statistical model or use modern machine learning techniques.

On the right-hand side is the case where there's knowledge available about the system being modeled, and that knowledge can be converted into governing equations. So, there's a picture of Isaac Newton, who postulated that force equals mass times acceleration, which forms the basis of many of the governing equations used in physics. And these can be called first principles-based models and include physics-based models and mechanistic models.

On the left-hand side, going back to the left hand side, those are sometimes called data driven models. But this isn't a clear categorization. There's no way of clearly delineating between these two types of modeling methodology, and there are hybrid methods such as first principle-based models that have data-driven subcomponents or the case of developing a first principles-based model and then training a machine learning algorithm to the results of the first principles-based model.

So what do we mean by computation modeling simulation? Well, basically, anything that's first principles-based. So, not statistical methods, not machine learning or artificial intelligence, those are not within the scope of this guidance, but anything such as physics-based models mechanistic models or first principles-based components within hybrid models.

Next, let's talk about how computational modeling simulation can be used in regulatory submissions. One typical application that's very well-established is modeling a device to perform virtual testing to address a safety or effectiveness question to generate data to be used in a regulatory submission. So as an example of this, for example, simulating the stresses and strains in a metallic implant under some kind of loading to answer the question of how safe is the device in terms of potential fracture risk.

That's the case whether a device is modeled and simulation results are submitted to FDA as part of the submission. Alternatively, a modeling algorithm could be implemented in medical device software. So for example, a typical example here is where device software takes in patient data such as imaging data, generates some kind of patient-specific model, and then simulates the patient in order to provide information, model-derived information, to a clinician.

Then there is an emerging method that's sometimes referred to as in silico clinical trials. This is the case where you have a model of a device like this, but it's coupled to models of patients and a range of model of models of patients representing some kind of virtual cohort. And that cohort is supposed to represent the indicated patient population in some way.

These first three examples are where modeling simulation is performed to support a regulatory submission for a single device, a specific device. So it could be used in a 510(k) submission, or a PMA, or IDE. Alternatively, we also have a pathway called the Medical Device Development Tool Pathway, and these are tools that could be used to generate safety or effectiveness data for a wide range of devices, and one possible type of those types of medical device MDDTs is computational modeling and simulation.

So, the scope of the guidance, as I mentioned, it's first principles-based models or, for hybrid models, first-principles model components. Out of scope of the guidance, standalone statistical or data-driven

models-- so, no pure machine learning models, artificial intelligence, is not within the scope of this guidance. Also outside the scope of the guidance is models that don't involve any simulation at all-- so, purely anatomical models, for example. And then as we will discuss, the guidance is focused on credibility assessment in a general sense. So it doesn't talk about how to perform modeling studies itself. It doesn't-- because it's general, and we'll talk about this in future slides, it doesn't talk about the specific technical details on how to perform credibility assessment. And it doesn't provide specific information on the level of credibility needed for any particular type of regulatory submission. That is something that has to be determined on a case-by-case basis and probably in discussion with the appropriate office within the regulatory side of CDRH.

So, some key definitions-- first of all, before we get into the structure of the guidance, credibility is defined as the trust based in all available evidence on the predictive capability of a computational model. The context of use, which is another key definition, is defined as the role and scope of the computational model in answering a question of interest. And then there these two terms-- verification and validation. These have specific interpretations for computational modeling simulation that are different to the definitions of verification and validation for software or for processes.

So, for mathematical models, so for modeling in general, verification basically asks the question, is a computational model, which is a software implementation of a mathematical model and a mathematical model is basically governing equations and parameters written on paper-- verification asks the question, has that been implemented correctly in the software, and what is the numerical error? Because typically, with these types of mathematical models, they can't be solved exactly. There's some discretization that's required, introduction of time steps or space steps, and there's some numerical error introduced by performing that discretization.

So verification is generally a mathematical or a software question. Doesn't require any real-world data, real-world observations. Validation, on the other hand, takes a computational model and asks, is it an accurate representation of the real world? And thus, this one does require real-world observations of some kind.

Moving on to the second part of the talk, we'll talk about some key points on the guidance and the approach taken by the guidance for providing a framework for credibility assessment.

So the guidance builds upon a standard that was published in 2018, the ASME V&V40 standard. That's American Society of Mechanical Engineers, Verification and Validation 40 standard. That is a standard that was developed for credibility assessment of modeling and simulation for medical device applications. That standard is also a general standard. It's not specific to any particular type of medical device. The standard provides a risk-informed approach to credibility assessment, which the guidance uses, and the standard has an emphasis on question of interest, context of use, and model risk, three concepts that we'll discuss more in following slides.

And the guidance also has the same emphasis. But the guidance also has additional recommendations on information to provide in a regulatory submission and generalizes the approach of the standard in a way that we will discuss. The guidance is intended to be applicable to a wide variety of models-- basically, all first principles-based models. And it's intended to be applicable to all applications of those models, all types of devices, and all types of regulatory submission. So it's a very general guidance, and as a result, it's not prescriptive.

Why do we need the guidance, or what did we try to provide in the guidance that goes above what's provided in ASME V&V40? Well, the framework in the guidance extends the approach of the ASME V&V40 standard. The guidance is motivated by the fact that the V&V40 standard implicitly assumes that when you evaluate a model, you will be performing validation involving prospective well-controlled bench tests.

So as an example, models are sometimes used to assess the question of how much a metallic implant will heat if a patient with that implant undergoes MR imaging because the metallic implant can absorb energy from the radiofrequency coil from the MR machine. Obviously, those types of models which try to predict the absorbed energy and thermal damage, very difficult to validate those models using clinical data. So those models can be validated using a bench test experiment such as in this figure.

Those are the types of validation that's implicitly assumed by the standard, but there's a wide range of possibilities for validating a computational model or a wide range of ways of using computational modeling, as we saw in an earlier slide. And that's part of the motivation for the guidance.

So, the third part of this talk is an overview of the framework provided in the guidance.

This is figure one taken directly from the guidance. Don't try and read the text, obviously. It's much too small to see. But I just want to emphasize that this figure is taken directly from the guidance. That's easier to read, and we'll go through these steps. There are nine steps, and we'll go through them one by one. So, for starting off with these preliminary or initial steps, stating the question of interest, taking the context of use, and assessing model risk.

Step one of the framework is state the question of interest. The question of interest is defined as the specific question, decision, or concern that is being addressed using the computational model, modeling and simulation. This should be a question about the real world, not a question about the model-- so, typically a question about the device or maybe a question about the patient if it's computational modeling implemented in medical device software.

The question of interest should be as precise as possible based on what the modeling is going to be used for. So, not overly broad such as, is the device safe, but more specific than that. So, in terms of-- as an example related to one of the figures that was shown earlier with the metallic implant and the question of fracture risk, a question of interest could be, is the device resistant to fatigue fracture under anticipated worst case radial loading conditions?

Step two is to state the context of use. This is defined in ASME V&V40 and in the guidance as the role and scope of the computational model in answering the question of interest. So this should be some text describing what is modeled and how the model outputs are used to answer the question of interest, the type of modeling field, key inputs and outputs, where key inputs come from, and critically whether other information such as bench, animal, or clinical data will also be used to answer the question of interest.

So, an excellent example, continuing the previous example, a context of use statement could be combined computational modeling predictions and empirical fatigue testing observations to estimate device fatigue safety factors under anticipated worst case radial loading conditions with additional

information on the type of modeling used, the key inputs and outputs, and especially what outputs will be used to answer the question of interest.

Step three is to assess the model risk, and this is defined in V&V40 and the guidance as the possibility that the computational modeling simulation results may lead to an incorrect decision that will lead to an adverse outcome. We follow the approach of V&V40, which breaks down model risk into two components-- model influence on one axis of this figure and decision consequence. So, considering model influence first, this is how much the modeling and simulation results will be used to answer the question of interest compared to other sources of information, such as bench, animal, or clinical data. So if both are going to be used or, for example, if animal test data is going to be used together with some computational modeling and simulation results to answer the question of interest, then the model influence might be assessed to be low.

But if the simulation results are the only source of information, then it might be assessed to be high. This is the case of modeling simulation used to generate safety or effectiveness data for a medical device. But if instead, it's simulation software that's implemented in medical device software, then, it could be the case if the simulation results are provided by the device as output to a clinician but it's intended that the device output is provided as an adjunctive information and the clinician should use other sources of information as well to make a decision, then the influence might be on the lower end. Whereas if it's a device that simulates the patient and the model derived output of the device will be the only source of information that will determine what therapy the patient undergoes, then the influence would be high.

The second axis here is decision consequence, and this is the consequence on what typically the patient of an incorrect decision. This could range from minor patient inconvenience, in which case, the decision consequence might be assessed to be low, or serious injury or death, in which case, the decision consequence might be assessed to be high. And we recommend using a figure such as this to convert the assessed model influence and the assessed decision consequence into an overall model risk. A few comments about decision consequence-- this is essentially risk as defined in the ISO14971 standard. And therefore, we recommend manufacturers consider both probability of occurrence and severity of harms when assessing decision consequence.

Moving on to the next stages of the framework, these are planning stages in advance of performing credibility assessment. They borrow from the approach of V&V40, but there's an additional concept that is introduced in this guidance.

That new concept is something that we called credibility evidence. We define credibility evidence as any evidence that could support the credibility of a computational model. And in the guidance, we provide a categorization of eight different types of credibility evidence. These are not intended to be completely comprehensive. There may be other possibilities that we didn't include in the guidance that arise and can be defined in a regulatory submission as needed. The categorization is also not intended to be a ranking in any way. The categorization is really a communication tool.

So we defined eight categories. I'm going to talk about these first five to make these ideas a bit more concrete. So, the first is code verification results. Code verification is basically testing that a computational model is an accurate implementation of the underlying mathematical model. So this is testing basically to try and identify software bugs.

The second is model calibration results. So, calibration is where you tune or vary parameters to fit experimental or clinical data. This is typically done as part of the model development process, but good calibration results can sometimes support the credibility of a computational model, though often, it might be in a complementary way to validation.

Validation is testing a computational model against data independent of the data used to generate a model. And so the next three categories are basically different types of validation distinguished between different types of comparator-- different types of data that the model is compared against-- either bench data or what we call in vivo validation, where it's compared against animal data or clinical data or population-based validation, where you have a-- this is for the case of a range of computational models of different patients representing, say, the indicated patient population. And each of those patient models is not validated individually, but you can validate the range of models against population level data from the real population.

The guidance provides recommendations and examples in these two sections, not provide any specific recommendations on exactly which categories of evidence should be provided for any particular regulatory submission. But we do recommend in the guidance providing, first of all, some evidence related to code verification-- that falls under category one-- some evidence relates to calculation verification-- that's basically information about the numerical error. And these numbers are the categories that could fall under. And some evidence related to validation or other evidence pertaining to the ability of the model to reproduce real-world data-- and those are the categories that are relevant.

The next step is most easily understood using an example. So let's suppose that bench test validation is performed. Then, the approach of the guidance and also the approach of the V&V40 standard-- oh, the V&V40 standard defines something called credibility factors and credibility gradations for bench validation, and the approach of the guidance is basically similar but allows the user to define their own credibility factors or gradations but provides recommendations for what to use.

So in the example, let's suppose bench test validation is being performed. Then what can be done is go into appendix one, that recommends credibility factors to use for that type of evidence. And specifically, it recommends going using the factors that are defined in the V&V40 standard. So, these are some factors that are defined for the V&V40 standard that are relevant to bench test validation.

And the next step is for each of these factors to define what's called a gradation. So this is a series of different levels of rigor for that particular factor. And so for this test samples factor, which is basically the number of samples that they used in the experiments that will be performed to validate the model, this is a range from a single sample to multiple samples to a large statistically relevant number of samples.

So, given that gradation, the next step [AUDIO OUT] appropriate level based on the risk-- so, the assessed risk from earlier. So if the assessed risk from earlier was medium, then that could be used as a justification for choosing the kind of intermediate level of the gradation.

Getting towards the end of the framework, we now move on to something called adequacy assessment. So there's to adequacy assessment steps in the framework.

I'll talk about the first one. First, which is part of the planning process, we recommend that as part of this planning, manufacturers generate a justification for why the planned credibility assessment activities would be adequate to support the computational model for the context of use given the risk assessment. So this is the rationale for why the planned evidence and planned results would be sufficient. And following this, we recommend coming to FDA with a Q-Submission to basically present the plan and present the entirety of the plan, including the question of interest, context of use, risk assessment, the planned credit assessment activities, this justification, and then getting feedback from the agency on whether we think it will be sufficient.

So this is this optional-- that was the recommended step of coming in with the Q-Submission at that point. After that, if everyone decides to go ahead, then there's the execution of the plan, the developer-- generating the credibility assessment activities, which is not really within the scope of the guidance. And then the final steps are to-- which I'll discuss in the next slide, something that we call post-study adequacy assessment and the preparation of a report.

So, whereas the previous adequacy assessment ask the question, will the credibility evidence support using the model, here we now ask the question, does the generated credibility evidence support using the model for the context of use given the risk assessment? This is a decision to be made based on all available evidence and requires clinical and engineering judgment. So we recommend providing this rationale, and some considerations are provided in the guidance, such as have all relevant model features been tested-- if some credibility goals from the plan were not met, it might be possible to adjust provide a rationale for why the results are still sufficient-- how do predictions compare to safety thresholds or decision thresholds and discuss limitations.

Finally, we recommend providing a self-contained report on the credibility of the model. So this will be distinct from the simulation study results itself. And appendix 2 provides a recommended structure for such a report. And appendix two also provides a recommended structure for Q-Submissions that might be provided earlier on just after the planning.

This slide just provides links to the two resources that I've mentioned-- the guidance and the standard.

And finally, to summarize the talk, this guidance is relevant to computational modeling and simulation, which is essentially first principles-based models, such as physics-based models or mechanistic models or first principles-based components of hybrid methods, hybrid models. The guidance provides a general framework relevant to all modeling fields or first principles-based modeling fields and all submission types. And therefore, it's not a prescriptive guidance, but it provides a process that can be followed. That's a nine step process. The first three steps define how the model will be used, the question of interest in the context of use, and assess the risk of using the model. The next three steps walk through the planning of the credibility assessment activities, lead up to a possible Q-Submission to discuss that plan with FDA. And the last three steps involve the execution of the credibility assessment activities, the justification in that post-study adequacy assessment, and the development of a report.

Thank you for your attention, and I will pass this talk back to Kim.

CDR Kim Piermatteo: Thank you Pras, for your presentation. We'll now transition to our interactive question and answer segment for today.

We have gathered a panel of subject matter experts, in addition to Pras to help you better understand and get clarity on what we intend in this final guidance. So joining Pras today on our panel is Brent Craven, Senior Science Advisor in the Division of Applied Mechanics in CDRH's Office of Science and Engineering Laboratories, Kenneth Aycock, Interdisciplinary Engineer in the Division of Applied Mechanics in CDRH's Office of Science and Engineering Laboratories as well, and Finn Donaldson, Team Lead of the peripheral interventional devices within the Office of Health Technology number two in CDRH's Office of Product Evaluation and Quality. Thank you all for joining us and for participating on today's panel.

Before we begin, I'd like to go over how we will manage the segment and a few reminders. First, to ask a question, please select the raised hand icon, which should appear on the bottom of your Zoom screen. I'll announce your name and give you permission to talk. When prompted, please select the blue button to unmute your line and then ask your question.

When asking your question, please remember to limit yourself to asking one question only and try to keep it as short as possible. And we appreciate that you may have a very specific question involving your device or scenario. However, we might not be able to answer such specific questions, therefore, we'll try to frame a broader response based on what's described in this final guidance. After you ask your question, please lower your hand, and if you have another question, please feel free to raise your hand again to get back into the queue, and I will call on you as time permits.

Now, as we wait to receive some of your questions, I'd like to welcome our newest panelists with a few questions we have gotten over the past few weeks about this guidance.

For our first question, I'll be directing that to Kenneth. Kenneth, the question is, the credibility guidance and the 2016 guidance reporting of computational modeling studies in medical device submissions both address physics-based modeling and recommendations for documentation and regulatory submissions. Can you clarify the differences in scope and use between these two guidance documents?

Kenneth Aycock: Thank you for the question, Kim. The 2016 FDA guidance primarily addresses how to report technical details regarding computational model setup and associated modeling study results. So essentially, the guidance focuses on what you modeled, how you modeled it, and what the model predicts.

The new guidance that's been discussed today instead focuses on assessing the credibility of your computational model through various credibility evidence gathering activities as Pras summarized, and the purpose is to establish trust in the predictive capability of the model. We anticipate both guidances will be used together as companions in regulatory submissions, and as Pras mentioned, for recommendations on developing a credibility assessment report, we recommend referring to appendix two of the new guidance.

CDR Kim Piermatteo: Thank you, Kenny. Now, for our next question, I'll be directing that to Brent. Brent, the question is, for what types of CM&S uses do you expect the new guidance to have greatest impact?

Brent Craven: Yeah, thanks, Kim, for the question. As Pras mentioned, the ASME V&V40 standard extensively covers credibility of modeling for non-clinical bench testing. I think the areas that will have the greatest impact-- this guidance will have the greatest impact is those areas where bench testing is

not practical-- for example, in vivo applications, areas like clinical studies or animal studies. In the new guidance we provide a lot of recommendations for this type of evidence that I think will be very useful.

CDR Kim Piermatteo: Great. Thanks, Brent. Now I'll be coming to Finn. Finn, the question I have for you is, how do you anticipate the new guidance will impact CDRH premarket reviewer's expectations for the content of marketing applications?

Finn Donaldson: Thanks, Kim. I guess in general, evidentiary requirements really shouldn't change too much, especially for the quite established uses of modeling. There's some differences in formatting, and particularly the test reports I expect would be required to align with the guidance, particularly being explicit about how the credibility was assessed. I think that will end up being more prominent as people follow this guidance more closely.

Really, I think the difference is that the guidance provides a lot more structure and language to support engagement between review staff and industry in demonstrating credibility. So hopefully, there's a lot of assistance there. Just upstream as industry are thinking about how to go about doing modeling, what would be involved in that, and how to support it. And I think the biggest impact in a sense is really going to be on novel uses of computational modeling and simulation and just take this chance to promote early engagement through Q-Submissions and Pre-Submission process for those.

CDR Kim Piermatteo: Thanks, Finn. Alright, we will now take our first live question, which is coming from Amir. Amir, I have unmuted your line. Please unmute yourself and ask your question.

Amir Sama Rezaei: Hello, everyone. Thanks for your presentation. It was great. I had a quick question regarding the step three that assesses the model risk. I wanted to know if you're actually assessing the robustness of the model per se or we are assuming the model-- if the model goes wrong, what are the risks associated with that? Which one are we questioning in this step?

CDR Kim Piermatteo: Thanks, Amir, for your question. Pras, would you like to provide a response?

Pras Pathmanathan: Yeah, sure. Thanks for the question. It's a great question. There's a lot of conversations that have been had about this step in V&V40. We've tried to, in the guidance, provide some extra context about what we mean by this assessment of model risk. So it really is, I think, what you said as the second option. If an incorrect decision is made, what is the consequence of that on the patient-- basically weighted by the consideration of model influence, how much you're using the model to answer that question. There's two parts of it-- decision consequence, and we emphasize in the guidance that decision consequence really isn't anything to do with the model.

All you have to consider when you try and assess decision consequence is the question of interest that was stated in step one. Consider that and consider what would be the result of answering that question incorrectly based on all the evidence that you use. Then there's the model of influence, how much is the model used compared to other sources of information. That requires what should be written down in step two, the context of use statement. And then through the two, do you end up with an overall assessment of model risk that basically assesses the overall consequence of making an incorrect decision using the model. Thank you.

CDR Kim Piermatteo: Thanks, Amir, for your question, and thank you, Pras, for your response. Our next question is coming from Lane Desborough, I have unmuted your line. Please unmute yourself and ask your question.

Lane Desborough: Good afternoon. Great presentation, Pras. Great guidance, team. Thank you very much. I have so many questions, but I think I'll go for this one. At the end of the day, how would one go about assessing the credibility of a simulation composed of both data-driven and first principles models? in other words, a simulation is a composition of models. What happens if there are a combination of those?

CDR Kim Piermatteo: Thank you, Lane.

[INTERPOSING VOICES]

CDR Kim Piermatteo: Go ahead. Sorry—

Pras Pathmanathan: Hi, Lane. Thanks for the question. Yeah, that's another great question. This is going to be an emerging technology. We're going to see more and more of these hybrid methods, probably more and more sophisticated models that use very advanced physics-based models coupled with very advanced machine learning models. And we hope that in the future, there'll be more specific guidance that we can release on that.

At the moment, we recommend kind of decoupling as much as possible the two different types of modeling using this guidance for the physics-based components parts of your model and then using relevant resources provided by the CDRH on the data-driven parts. And in the guidance, we in particular recommend, again, speaking to the regulatory team early to see what recommendations can be provided about the statistical or machine learning parts. I will pass that-- I will open that question to the other members of the panel in case any of you have anything you can add cause I know that this is a topic that could be discussed a lot.

Brent Craven: Yeah. Pras, yeah, I would just like to add that this is a very timely question and something we hope to be addressing more in the future. But for the time being, I think a good approach for industry is to, if you have specific questions, to come in with a Q-Sub so we can start to discuss how these different modeling modalities are being combined. So I guess coming in early with the Q-Sub for discussion, I think, is a really good approach.

Pras Pathmanathan: And I guess I can add that the Q-Sub can cover everything. It doesn't have to be-- we provide in the appendix a recommended structure for a Q-Sub related to first principles models computation modeling simulation. But you can come in with a Q-Sub that covers that part as well as questions about the, let's say, machine learning side.

CDR Kim Piermatteo: Great. Thank you very much, Pras and Brent, and thank you, Lane, for your question. Our next question is coming from Charley. Charley, I have unmuted your line. Please unmute yourself and ask your question.

Charley Taylor: Hi. This is Charley Taylor from HeartFlow. Thank you again very much for the presentation. You had a slide where you talked about kind of the four categories of CM&S in regulatory

submissions. The fourth was for you really specifically focused on devices. But do you see that this guidance document will also be relevant, for instance, for pharmaceutical or therapeutics where simulation technology CM&S is used for pharmaceutical development, drug delivery?

Pras Pathmanathan: Hi, Charley. Yeah, that's a really interesting question. The guidance is-- because it's a general guidance, there's certainly-- people can look at this guidance and think, yeah, it could be applicable to drugs. At the moment, I want to emphasize that this is a guidance that's come out of CDRH only. The clearance process and the authors of this guidance are entirely CDRH. So it hasn't been reviewed by CDER, the drug center. And I recommend speaking to CDER for the plan for computational modeling for any regulatory submission for a pharmaceutical product.

I should point out, though, that there's interest in CDER on the approach taken by V&V40 and therefore in the kind of approach taken by this guidance. They have published papers talking about how V&V40 could potentially be used for drug related applications, and I would hope that they'll be continuing conversations, and ultimately, we could-- very much down the line, we could imagine a guidance that's applicable to a wide range of medical products that are regulated by FDA.

Charley Taylor: Thank you.

CDR Kim Piermatteo: Thank you, Pras, and thank you, Charley. Our next question is coming from MArun. I have unmuted your line. Please unmute yourself and ask your question.

Mike Arun: Hi, this is Mike Arun from Alphatec Spine. Very nice presentation. Really appreciate it. So, my question is, so the computational models can be used as two different forms, as I can understand. So, one is like a supporting tool to back up like a hardware submission or something like that, and the other one is the computational model as a tool on its own, like a standalone tool. So, is there any differences in FDA's expectations when a computational model is submitted as a supporting tool versus when it's submitted as a standalone tool?

Pras Pathmanathan: I think I need to ask a clarification question to answer that. By supporting tool versus standalone tool, are you talking about the distinction that I made in an earlier slide, where a computational modeling could be used, one, to generate safety or effectiveness virtual test-based data on a physical device versus where the modeling is implemented in software within a device?

Mike Arun: Or something like a HeartFlow, where the computational methodology on its own is like a tool-- so, where the decisions are made directly from the output of the model itself without any hardware interference. Yeah.

Pras Pathmanathan: Yeah, great question. Yeah, we do try and cover both in the guidance by providing this general categorization of the evidence. I'll emphasize that in the first case, the FDA reviewers will be reviewing the test data that's generated by the computational model whereas in the second case, where the model is part of the device and the model is going to be used by clinicians, say, like the example of HeartFlow, FDA has to review the initial test data but also assess how reliable the model is expected to be once it is used in the field.

Generally, the framework of the guidance applies in both cases. So we don't expect too many differences. The types of evidence that would be submitted will very likely be different. For the first

case, those kind of models would often be validated against bench data. The second case, where it is a clinical tool, it often or many cases, it would only make sense to validate against clinical data or animal data. But overall, yeah, I think the-- but I think to be honest, I think the answer will be very case-specific I don't know if any of the other panelists want to provide any comments on this question.

Finn Donaldson: Yeah, this is Finn. I suppose I echo your comments there. I think in general, it's going to be an applicable approach here. In the case where the modeling, if you like, is part of the software as a medical device, there's additional things to consider there where the device has all the software requirements, software as medical device type components that need to be evaluated, and we have other guidance out of scope of this discussion that would need to be thought of that wouldn't be relevant, say, if it was just a computational model to evaluate as a surrogate for a bench test, for example.

I think the other thing that might be different is just if the model is the device, if you like, then the model influence is much higher in general, and the credibility expected might be higher as well as a general statement.

Pras Pathmanathan: Kenny?

Kenneth Aycock: Yeah, that's a really good point, Finn. I'll also just add briefly that throughout the guidance, especially at the beginning, we have a few bullets kind of distinguishing between encyclical device testing and models used within device software. So when reviewing the guidance, some of those bullets may be helpful in kind of understanding nuances with respect to the question of interest, context of use, and how those end up being framed a little differently.

Mike Arun: So as a kind of a spine company, so as to speak, say, if the software output is going to kind of predict some kind of a, say, spine stature, so would it be more appropriate in terms of validation to retrospectively validate the model with the real world-- some kind of a follow up data? Do you think is that a reasonable way of validating the software?

Kenneth Aycock: Yeah, I think that's getting into a quite specific question that's probably best handled in the review time team here.

CDR Kim Piermatteo: Thank you, Mike, and thank you, everyone, for providing a discussion on this question. We're going to move on to our next stakeholder, and that is Kristian. Kristian, I have unmuted your line. Please unmute yourself and ask your question.

Kristian Debus: Yeah, hello. Thank you. This is Kristian from [INAUDIBLE]. Yeah, great to see this presented today. So, I've been excited for it for a while now. I just want to ask a very simple question. So, does the FDA have specific plans to promote this guidance document not just internally but also externally. So basically, for internally, how are you planning to roll this out to reviewers? Because I'm a big friend of use cases and just example cases and just really working together and promoting that and externally, internally, will be a-- bring this to a success I, think. So, do you have any plans for that or any timelines?

Pras Pathmanathan: Hi, Kristian. Thanks for the question. Yeah, we have already provided some initial training on the guidance internally for reviewers. We hope to provide more detailed training in the

future. We also hope to collect lots of information about how the guidance is being used in regulatory submissions and share that.

One of the things that we're planning on doing over the next few years in OSEL-- that's the Office of Science and Engineering Laboratories within CDRH. That's the regulatory science office, not the review office. One of the things that we try and develop are something called regulatory science tools, which are tools that industry can use to support the regulatory submissions in any way.

And one class of those kind of tools are examples, and we hope to generate examples of this guidance that we would make public. Example of the guidance being used, certainly of the process of the guidance being used that we would make public, and that would be hopefully extremely useful both for industry and for reviewers. So we have various plans to, yes, internally and externally promote the use of this guidance.

Kristian Debus: OK, great. So you're kind of touching on the point of, like, offering a tool to ease that communication because it's different than submissions used to be, right? So that sounds pretty interesting, though. Thanks for sharing.

Pras Pathmanathan: Thank you.

CDR Kim Piermatteo: Thanks, Pras, and thanks, Kristian. Alright, we're going to move on. Our next question is coming from Naveen. Naveen, I have unmuted your line. Please unmute yourself and ask your question.

Naveen: Yeah. Hello, everyone. I'm Naveen from Stryker. So, my question is how to appropriately select credibility goals for a model risk which is between two risk levels like low-medium or medium-high?

Pras Pathmanathan: Hi, Naveen. Yeah, good question. So, if you have a model risk that's assessed to be something intermediate, let's say, like the second or fourth one on a five level scale-- medium-low, medium-high-- and when you write down these gradations, V&V40, they provide a list of gradations for the credibility factors. And most have three, but some have four, some have two.

And I think if you have a three level gradation, then by default, if it's medium-high risk, then the gradation level that's the highest one that's greater than the risk is the most appropriate one. But basically, you can try and justify any level within that gradation and try and explain why it's sufficient. We do allow that flexibility. It's something that I know that the V&V40 subcommittee is trying to clarify in future versions of that standard. Brent, would you like to respond?

Brent Craven: Yeah, I'd just like to add that I think we provide-- one aspect of this guidance that I think is really helpful is we have a section on adequacy assessment, and we provide a lot of discussion on how a lot of the justification for how much modeling or how much credibility assessment you've done really comes down to a justification exercise and is somewhat subjective sometimes. You have to make justifications for different aspects of the model that may or may not affect the overall credibility. So I would recommend checking out that section, and there's a lot of different considerations.

At the end of the day, even V&V40 does not prescribe a quantitative way of prescribing a value to the end credibility. At the end of the day, it's often a justification. So I would check out that section.

Finn Donaldson: I'd maybe just add as well I think that those are really helpful comments, and it's a very important question you're asking. I'd refer you to figure three of the guidance as well. It talks about the credibility factors. And you can see there that even though there might be a medium-high, say, goal, some factors are going to be above, and some factors are going to be below that in terms of how intense they've been evaluated. And that's fine. This has to be done in a practical way, and it's not going to be too precise about hitting exactly that model risk. But hopefully, that helps. Figure three is very relevant.

Naveen: Thanks, everyone, for your comments.

CDR Kim Piermatteo: Thank you, Naveen. Alright, I think we have time for one more question. I'm going to call on Divya Gowder. I apologize. I have unmuted your line. Please unmute yourself and ask your question.

Divya Gowder: Thank you, Kim, for taking my question. Can you hear me OK?

CDR Kim Piermatteo: Yes, we can.

Divya Gowder: Great, thanks. And this is a great and informational session. Thank you so much for that. So my question is very quick. So, in terms of the computational models and leveraging the data for regulatory submissions, so is it possible that for a class III device, which is obviously a higher risk class than a class II, if there is impractical reasons for not conducting the clinical testing, can the submissions for class III PMA be entirely done using the computational models?

Pras Pathmanathan: That will probably be very-- be a case-dependent answer, but I will pass it on to Finn as the reviewer to see if you want to say anything specific.

Finn Donaldson: Yeah, I think in theory, absolutely. In practice, I think you'd anticipate that it would be quite challenging. The model influence would be at the maximum, the decision consequence. Particularly for a class III device, which is the highest risk category we've got, is also kind of near the maximum too, if not the maximum. So the credibility required to support a model like that is high, and that's going to require quite a bit of consultation with the relevant review teams to figure out what those steps to demonstrate that high credibility are.

So in theory, it's possible. I think it's a hard road to travel. I imagine an easier path is some sort of hybrid approach. We've seen, I think, a mock submission in this space where there's a sort of Bayesian attempt to blend together real clinical data and then in silico clinical trial data. We can talk about that at length, but I suspect that's somewhat lower hanging fruit while still high up the tree, as it were.

Divya Gowder: OK. Thank you.

Kenneth Aycock: This is Kenny. Definitely agree with Finn and Pras. Just would also clarify that the impact of modeling for any class of device is going to be very dependent on the context of use, and definitely, there are opportunities to use modeling across all classes. It's just for the particular specific question that the bar would be pretty high.

Divya Gowder: Sure. Understood. Thank you so much.

CDR Kim Piermatteo: Thank you, everyone, and that concludes our question and answer segment for today. Thank you all again for your questions and your participation today. I hope you found it very engaging.

So I'd now like to turn it back over to Pras to provide his final thoughts for today. Pras?

Pras Pathmanathan: Thank you, Kim, and thank you, everyone, for joining. I hope you found this session informative and useful. This guidance is something that we're excited about because it's an important milestone for the center, the first guidance that we have released that's focused entirely on credibility of computational models.

But I want to emphasize that this is part of a process of where the community, the medical devices community, the modeling community, process that we've been working on for a long time. Been more than 10 years since the V&V40 Committee was developed, and I want to also thank everyone in that committee as well because we're building on the backs of everything that they've done. And there are a lot of people that have worked in this area for a long time now and would be obviously not possible to thank everyone.

So finally, yeah, thank you, everyone, for your time and attention today. If you have any further comments or questions that are not device-specific, then you can contact us at the email addresses that are in the first page of the guidance. And thank you, and I'll pass it back to Kim.

CDR Kim Piermatteo: Thank you, Pras. For your information, printable slides of today's presentation are currently available on CDRH Learn at the link provided on this slide under the section titled “Specialty Technical Topics” and the subsection titled “Regulatory Science Tools.” A recording of today's webinar and a transcript will be posted to CDRH Learn under the same section and subsection in the next few weeks, and a screenshot of where you can find these webinar materials has been provided on this slide as well.

If you have additional questions about today's webinar, you may also reach out to us in DICE at DICE@fda.hhs.gov.

And lastly, we hope you're able to join us for a future CDRH webinar. You can find a listing of all of our upcoming webinars via the link provided on the bottom of this slide at www.fda.gov/CDRHWebinar.

This concludes today's CDRH webinar. Thank you all again for joining us. Have a nice day.

****

END