FY 2026 GDUFA Public Workshop – Day 1: Edited Transcript

Compiled from captioner and AI-generated transcripts.

Session 2: Using AI to Address Practical Challenges for Generic Drugs

---

Co-Moderators:

Meng Hu, PhD
(Lead Chemical Engineer, DQMM/ORS/OGD/CDER/FDA)

Qing Liu, PhD
(Deputy Director, DB I/OB/OGD/CDER/FDA)

Meng Hu, PhD: Hi, hello everyone. Welcome back. Let's start session two. So, first of all, my name is Meng Hu. I feel very privileged and co-moderate with Doctor Qing Liu. We are moderating this exciting session named "Using AI to Address Practical Challenges for Generic Drugs". So this session will have a live presentation, consisting of speakers from industry, technical providers, academia, and also the agency. So with that, let me introduce our first speaker, Commander Diana Solana-Sodeinde. She serves as a lead program coordinator for FDA Inactive Ingredient Database, IID program under the Office of Generic Drugs. With that, let's welcome Commander Diana.

Presentation 1: AI and FDA's Inactive Ingredient Database

CDR Diana Solana-Sodeinde, PharmD *(Lead Program Coordinator, OGD/CDER/FDA*)

Thank you very much for that welcome introduction, Doctor Hu. Thank you, everyone. I hope you're all doing well. How was lunch? Was it good? Yeah? Okay, don't fall asleep now. All right. Okay, so as Doctor Hu said, my name is Commander Diana Solana-Sodeinde, and I am a pharmacist by training. And today, I'm going to be presenting on FDA's inactive ingredient database program and the application or consideration of artificial intelligence powered tools. And I work in the Office of Bioequivalence under the Office of Generic Drugs. And let's begin.

So the outline for my presentation is basically going over an overview of the inactive ingredient database, what we fondly like to call IID. Then I'm going to give you an overview

of the IID's three-prong process approach and present a little bit about the consideration of AI-powered automation tools.

So what is the inactive ingredient database? This is a system, and this is a system that is retrievable on the FDA website. So it's an external facing website. And the IID provides information on excipients that are present in FDA approved drug products. So specifically abbreviated new drug applications and the new drug applications, all the drug products associated with those applications, the goal is to find the information on excipients in the IID. And so when you go into the FDA website or you just Google the IID or inactive ingredient database, it will bring you to this page. And this is our homepage. And you can easily search for an inactive ingredient by just entering the first three letters of the name. And you can also do the same by just clicking on the first letter as it's listed down there from A to Z of the specific excipient that you're interested in.

And why did we even have this, right? So can you imagine, in 1987, FDA first made available a guide. It was called the Inactive Ingredient Guide, and this was available in hard copy paper format. Then in 2003, FDA started making the information, excipient information, accessible online through this database. So here's a snapshot. And so for more information, if you are interested, this is very critical information that our pharmaceutical industry uses when they're manufacturing drug products. You can refer to the helpful resources. We have a frequently asked questions document that explains what an inactive ingredient is. We also know that an inactive ingredient has an interchangeably terminology, the excipients. You're going to hear me using both terms. And then you can also learn more about what this database entails in the guidance for industry. And that's been very useful for our industry counterparts and for the public as well.

So that's just a brief presentation on the inactive ingredient database, which is a system that has excipient information. Next, what exactly populates the IID? Well, like I said earlier, we have information coming from the newly approved products and information from older previously approved products. And the goal is, based on our GDUFA three commitment, is to publicly publish updates on a quarterly basis. And these updates include excipient information as it relates to the maximum daily exposure, or maximum daily intake, and the maximum potency per unit dose information for each record, which consists of the excipient name, the route of administration, and the dosage form.

And so in order to make these updates on a quarterly basis, we go through a three-prong process. And as you can see here, step one is determining the maximum daily dose value. Then we calculate the maximum daily exposure, and then we update the IID database system. So when it comes to determining the maximum daily dose, we'll obtain this information from approved RLD labeling and reviews. Once we gather this information, we

plot it into a formula to calculate the MDE. And once we verify the information, we then upload the files to the FDA IID internal systems and eventually into the external facing website. I just want to point out here that all this has been going on through manual tasks. And so we did want to enhance the way we were doing the processes.

Here's the formula to calculate the excipient maximum daily exposure. And it simply is equal to the maximum number of dosage units that's recommended per day based on the product labeling that's approved, multiplied by the dosage unit level of the specific inactive ingredient or excipient. So let's say, for instance, you're taking a tablet, I'm not going to say the name of the tablet. And the maximum daily dose is 2 tablets per day. That's the dosage unit. So once you figure out, once we determine what the excipient level is in that specific tablet, let's say it's 20 milligrams, we'll just simply multiply both numbers and we'll arrive at 40 milligrams. That is the maximum daily exposure, right, for non-oral products, or in this case, a tablet, the maximum daily intake for the tablet. And that's 40 milligrams per day.

So as we were trying to think about what do we need to do, how do we enhance our program and the processes involved, we looked at the current state and where we wanted to go to. And we identified the gaps and the critical areas that needed attention. When we identified those, we were able to now determine the needs as it relates to the resources that we need, specifically the IT system, right? And we looked, we said, you know what, we have a goal. Let's create an action plan, which is to consider AI-powered automation tools as part of the IID program enhancement.

So, of course, we know AI-powered automation tools significantly can transform IT systems and programs across multiple dimensions by reducing the burden of repetitive, manual, labor-intensive tasks that staff currently have to do. In addition, if we're able to figure out a great tool, we can help relieve the staff so that they can focus on more higher value strategic work. And thus, this will help create a more proactive, intelligent, and efficient approach to addressing different types of challenges that we've identified in the systems. So we took that concept and we applied it to the IID program. And of course, we did this because we know, and we've seen that AI tools, and heard, that AI tools hold a strong potential to modernize and streamline review processes, and also to modernize our internal stakeholder interactions. AI tools can reliably extract inactive ingredient information directly from the approved RLD labeling and drug reviews. Thus, again, it will help reduce the manual data entry, repetitive task, and associated human error and meet our goals. In other words, increase our outcome, right? We've set a target that we want so many MDE values populated in the IID, because that's our main goal. And all this to say is we're moving towards an automation input, which is favorable so that we can apply a

proactive, intelligent, and efficient approach to address challenges within the IID program and systems involved.

And with that being said, you're going to hear more from Dr. Wang. She's going to talk to you about an AI automation, AI-powered or generated automation tool that was used to calculate and determine the maximum daily dose value for the drug products. And I just want to say a big thank you to the Office of Generic Drugs. That includes all the offices. We've been really hands on deck. Our Office of Bioequivalence, IID team and review staff, the Office of Research and Standards, the Policy Shop, and ORO. And then we also want to say thank you to the Office of Pharmaceutical Quality, the Office of Business Informatics, and CDER's Office of Communications. So now I hand it over to Dr. Wang, who's going to go into details regarding the AI-generated automation tool. Thank you.

Meng Hu, PhD: OK. Thank you. Thanks, Diana, for the nice introduction about the programme of ID and the roles of AI to facilitate this programme. So, next speaker, I would like to introduce Doctor Jing Jenny Wang. She's a currently working on the division of a quantitative method and modelling of Office Research and Standard in Office of Generic Drugs. She has a focus on utilizing quantitative analytics and artificial intelligence to support regulatory review efficiency. And she will share a research work on development of AI automation tool to facilitate maximum daily dose determination.

Presentation 2: Development of an AI Automation Tool to Facilitate Maximum Daily Dose Determination

Jing (Jenny) Wang, PhD *(Staff Fellow, DQMM/ORS/OGD/CDER/FDA)*

Hi, good afternoon, everyone. Thank you for being here. Thanks for the introduction from Meng. And today, yeah, I'm very glad I got this chance to share our development of an AI automation tool to facilitate the maximum daily dose determination. So, this is a disclaimer.

In today's presentation, I will first introduce some background information, and then I will share the details of the Agentic AI maximum daily dose tool regarding our prototype workflow and a case example, and how the tool does the batch processing and our current validation results. And, at the end of this presentation, I will give a summary and an overview of our future work.

So, the MDD information is very important in excipients evaluation and determining the MDE. So accurately identifying the MDD information is a crucial step to avoid potential

toxicity or adverse reactions caused by the excipients. So the MDD is usually determined based on the information from drug labeling and FDA internal documents. So, our objective is to develop an AI automation tool to facilitate the efficiency and consistency of the MDD determination.

So, here I will share some details regarding our prototype workflow. So when the users start to use our tool, they can start this process by providing a drug labeling to the tool. So within this tool, we have 4 AI agents developed. They are information retrieval agent, extraction and curation agent, calculation agent, and the summary and evaluation agent. So when this process starts, the information retrieval agent will retrieve the relevant information from this drug labeling and pass the format and then give it to the next agent, the extraction and curation agent. So the second agent will analyze the retrieved information and extract all the potential dosing scenarios from this information and curate them to a list with the source information. And then pass this information to the calculation agent. So this agent will calculate the MDD value for each dosing scenario based on the standard principles established in FDA internal procedures, and then give this list of MDD information for this whole list to the summary and evaluation agent. And this evaluation agent will review all this information and give a final summary, which is a final MDD value for this product, and shares the calculation process, the thought process, and gives this final response to the user.

Here are some like detailed information for each agent. Here, we would like to emphasize that for our calculation agent, the calculation would be based on the standardized approaches and clinical principles established in FDA internal procedures. So, in the next part, I will share how each agent will work through showing a case example.

So here is a case example. This is like part of a drug labeling. And this drug labeling will be given to the tool. The first agent, information retrieval agent, will handle the format parsing. As we know, like sometimes the format of the drug labeling can be complicated, involving like multiple columns and multiple tables. So this agent will handle this part, complete the format parsing and retrieve the relevant information, follow the principles of the FDA internal procedures, and then this retrieved information will be given to the second agent. This is an extraction and curation agent. This agent will digest the retrieved information and then extract a list of all possible dosing scenarios. So for each scenario, it will include the source information in the drug labeling, and also give like the condition, the population, this information, and curate them to a list.

This list will be given to the next agent, the calculation agent. So following the principles in FDA's internal procedures, this calculation agent will calculate the MDD value for each dosing scenario. As we can see here in the output of this agent, for each scenario, it will

include the source information from the drug labeling. And then it will share the language model's thought process, like how this model digests this information, and then shares a very detailed calculation steps. The model will show what kind of steps, what kind of calculation it did to get the MDD value for this dosing scenario and list all the information here.

So this list will be given to the last agent, the summary and evaluation agent. This agent will review all this information from the calculation agent and then give an overall summary, summarize like for each dosing scenario, what's the MDD value. And based on this information, this agent will give a judgment, like, what should be the MDD value for this product, and this MDD value is from which scenario. And for this chosen scenario, this agent will also give the detailed calculation steps, thought process, the source information, and what's the condition for this dosing scenario.

So this is a case example showing how our MDD AI tool works for a single drug labeling. In the next part, I will also introduce how this agentic AI MDD tool does a batch processing to process multiple labels in one batch. So here, this is the batch processing workflow for our tool. So this batch processing starts from a list of RLD application numbers, can either be in an Excel file or a CSV file. So this list of RLD application numbers will be given to the tool. And then this tool will download the labels automatically from the Drugs@FDA in a batch, and all these downloaded labels will be saved to one folder. Then this folder will be given to our tool. This MDD AI tool will process all the labels in this folder for the batch processing. So all the MDD results for all of this like RLD products, like each RLD product will have a one row of this MDD information in our Excel file and all these results will be saved in one final Excel file, also including like the detailed thought process, calculation, the source information for each product.

Here, I also would like to share some current validation results. So the scope of our validation is to test how accurate the current MDD AI tool can follow the principles in FDA's internal procedures regarding extracting MDD information from drug labeling. And we also want to refine our structure and setting of the MDD AI tool through the validation with Subject Matter Experts' verified MDD values, so this will serve as a foundation for our development for the pediatric MDD AI tool in the next step. So our tool generated MDD results for 113 RLD products. Within this result, there are 12 RLD products for outside the scope of current testing, including 10 products with the SME verified MDD determined from sources other than the drug labeling. And also there are two products with the SME verified MDD based on the pediatric indication. So of the 101 RLD products with the SME verified MDD determined from labeling, for adult indications, the tool matched 92, reaching an agreement rate of 91.1%. So, here are some reasons for our total nine discrepancy cases.

And some cases were attributed to the absence of default standard weight for the obesity population, or some related to the execution of the rare conditions.

So, our MDDAI tool is a collaborative effort among the Office of Research and Standards, Office of Safety and Clinical Evaluation, and Office of Bioequivalence in OGD. So, this tool is designed to strictly follow the standardized approaches and clinical principles established in FDA internal procedures for MDD determination from drug labeling. Please note that the MDD determination is a complicated procedure which may involve multiple review assessments with internal and external documents other than drug labeling. So, the current MDD AI tool only focuses on processing the information in the drug labeling, and the final MDD is subject to the further assessment by the SMEs. So the utility of this tool in the review process is under discussion. And we are working on extending this tool to identify the pediatric MDD to support an FDA-NIH collaborative project on pediatrics. Here's our acknowledgement. We really want to give great thanks to our team from ORS, OSCE, and OBE in OGD. Thank you.

Meng Hu, PhD: Yeah, thanks, Dr. Wang. Next, I'm pleased to introduce Miss Arti Patel. She's a global regulatory affair leader at the Cosette Pharmaceuticals with over 15 years of experience driving end-to-end regulatory strategy, approval, and life cycle management. Ms. Patel, please**.**

Presentation 3: Artificial Intelligence: Key to Address Practical Challenges for Generic Drugs

Arti Patel, MS *(Sr. Manager, Cosette Pharma)*

Good afternoon, everyone. And thank you for the introduction, Meng Hu, and thank you, FDA, for the opportunity to present the industry perspective here. Myself, Arti Patel, and I'm talking about the AI strategies to modernize IID and MDD for generic development.

General disclaimer that this presentation reflects the view of the author and should not be considered to represent Cosette's view or policies. The views expressed in this presentation do not necessarily reflect the official policies of the regulatory agencies.

In the subsequent slides, I am going to present about why Inactive Ingredient Database, that's IID, and Maximum Daily Dose, that's MDD, are bottlenecks. What initiations have been done by FDA for these parameters? What are the key challenges and what are the root causes? And finally, how artificial intelligence bridges those critical gaps.

Let's check more about why IID and MDD are bottlenecks. As part of formulation development, important parameters to be considered are safety, efficacy, and quality. As

part of safety, evaluation of IID and MDD are required because it serves as a critical threshold for establishing safety. However, (..) the evaluation process for IID and MDD is often not straightforward. In many cases, questions arise due to incomplete or unclear information within product labeling, complex dosing instruction, or challenges associated with certain inactive ingredients. When such ambiguity arises, applicants frequently need to pursue clarification through the controlled correspondence pathway. Currently, the response timeline is approximately 60 days for the IID related correspondence and up to 120 days for MDD related correspondence. In the highly competitive generic drug market, where development timelines are critically important, even a relatively small delay can significantly impact submission readiness, approval timeline, and at the end, patient access to affordable generic medicines. Therefore, IID and MDD have become bottlenecks, affecting overall development efficiency.

FDA has already initiated several important steps to mitigate or to improve the IID ecosystem, which includes populating maximum daily exposure information, updating the IID database on an ongoing basis and publishing a quarterly notice of updates, searchable quarterly change logs, and adoption of new data standards. This resulted in the number of entries in the database collapsing from 14K to about 10K.

With all these initiations, there are still some key challenges. Starting with the IID database, there is a lack of context for specific use. Example, we cannot find detail about whether this IID is related to pediatric population or not, whether it is for the chronic or acute use, and the specific context of use. Absence of maximum daily exposure for many entries, approximately we can say around 50% yet, and inconsistent naming conventions. Hence, it is important to recognize that an IID listing should not automatically be interpreted as direct evidence of safety.

On the MDD side, manual and time-intensive MDD calculations, which are prone to inconsistency. Ambiguous dosing instruction across multiple labeling sources that further complicate the MDD calculations. In addition, there is currently no automated mechanism to cross-reference dosing assumptions with approved reference standards or historical precedents. And if we look into the root cause of these challenges, the following parameters could be considered as playing an important role: unstructured historical data and heavy reliance on PDF-based labeling, which makes automation extraction very difficult. Manual interpretation of dosing instruction leads to inconsistent assumptions during MDD evaluation and lack of system integration to manage a high volume database.

With these challenges in mind, artificial intelligence presents a significant opportunity to modernize and strengthen the IID and MDD evaluation process. Before we dive into an AI

system, I would like to be very clear that based on my experience and understanding of the process, AI is to facilitate the process. Final control will always be with the human.

Conceptual future state vision. The following example is a conceptual mock-up intended to illustrate how AI-enabled regulatory intelligence may support IID and MDD evaluation workflows. These examples do not represent actual FDA systems or regulatory decisions. Based on the challenges discussed earlier, we conceptualize an AI-enabled framework called IID MDD Navigator Tool. This workflow begins with the injection of a label. This will be considered as an input window. Based on the stage of product development, applicants should be able to upload formulation details and can perform specific tasks like IID evaluation.

Next step includes AI-powered extraction using OCR and NLP tools, which identifies key information from the label. The extracted information is then processed through an automated MDD calculation engine. This engine needs to be trained for the calculation formulas for the critical parameters of all different dosage forms. Some of them are summarized here. That could be body surface area, fingertip unit, standard criteria of drop size for the ophthalmic products, or standard body weight criteria, and there are many more when we come across many more products.

The next layer introduces context-aware analysis, where AI evaluates critical regulatory considerations such as pediatric use, chronic exposure, sensitive route of administration, and high-frequency dosing. Based on the updates in the regulatory understanding or criticality of the products, this layer needs to be updated for the consideration of parameters. This information is then connected to IID intelligence through comparison against the IID database, historical precedents, and similar approved products using a defined integrated tool with the IID evaluation process.

Based on these inputs, the system performs AI-driven risk scoring by evaluating concentration risk, exposure risk, population risk, and cumulative risk. Finally, the platform performs a reviewer-ready regulatory summary, including key findings, identified risk flags, justification needs, precedent preferences, and a complete audit trail to support regulatory traceability and scientific decision-making. As I cleared in the start, this framework is designed as a human in the loop system, where AI assists the reviewer, but final scientific and regulatory decisions remain under expert human oversight.

This is a little bit clumsy but I just presented to see how the output might look. Here I would like to demonstrate the potential final output of the IID MDD Navigator platform. In the example, the system has processed an oral capsule product label and automatically extracted critical information. Using these extracted inputs, an automated MDD engine

performs the calculation to estimate maximum daily exposure of active ingredients. The platform then performs context-aware analysis. Next, the tool compares the proposed excipient levels against the IID database. Here the system creates a tabulated summary file which describes if any excipients have a value above the IID database and presents on the dashboard an overall excipient risk. And that way, it will score the final risk scoring, and based on this outcome, the reviewer will come to know if any additional justification is required or how to proceed.

Definitely, we are talking about AI, but challenges are there. High dependence on source data quality, standardized criteria for the AI tool teaching, limited availability of contextual regulatory data, validation and continuous maintenance, integration with the constantly updated published data. And if we are talking about novel excipients and new formulation, that's again a challenge.

Key takeaway, the status quo is unsustainable. With the scale of many products, if we go manually, chances of error are high. AI can bridge the critical efficiency and consistency gaps. Human oversight remains essential for regulatory compliance. With this, I would like to acknowledge Ravi Patel, Ganesh Pillarisetti, Rekha Chappidi, my fellow colleagues, and FDA. Thank you.

Meng Hu, PhD: Thank you, Miss Patel. Next speaker, I want to introduce Doctor Jeremy Evans. Doctor Jeremy Evans serves as the Director of AI Machine Learning Innovation for Pharmaceutical Manufacturing at the USP, where he applies simulation and modeling to advance cutting-edge drug manufacturing techniques. And he will talk about AI and machine learning data enhancement in generic drug manufacturing. With that, Dr. Evans.

Presentation 4: AI/ML and Data Enablement in Generic Drug Manufacturing

Jeremy Evans, PhD *(Director, AI/ML Innovation in Pharmaceutical Manufacturing, USP)*

Thank you, thank you, Meng. So, as he mentioned, my name is Jeremy Evans, and I work with the US Pharmacopoeia. I'm going to talk a little bit about how we're using AI and machine learning at USP, as well as our considerations for applying digital-driven techniques in a regulated or validated context.

At USP, our mission is to improve global health through public standards and related programs that help ensure the quality, safety, and benefit of medicines, dietary supplements, and food ingredients. Many of you may know us largely from the Pharmacopoeial Compendium, as well as our reference standards, but we have a broad number of offerings, including education services, supply chain and manufacturing

services, verification services, analytical solutions, and what's most relevant here, digital solutions.

We recognize that AI and machine learning are becoming more and more an integral part within both drug discovery and drug manufacturing because of the massive opportunities that exist to essentially put drugs in the hands of patients faster. Opportunities such as data quality, integrity, and consistency, by, for example, avoiding manual transcription errors and improving consistency from application to application. Operational efficiency, so essentially making our generic drugs more efficiently, cheaper, getting them to the market faster, and essentially getting them into the hands of patients and making them accessible to patients.

At USP, we have quite a number of applications of digital methods to generic drug development. So I've listed a number of them there. One of our key applications is the Medicine Supply Map, where we leverage millions of data points to help predict drug shortages and help to identify those risks and alleviate them. Digital reference standards, so the USP-ID program, where we essentially offer a digital alternative to some physical reference standards leveraging quantification through NMR. We have updated and modernized our compendium, and that has included updating the compendium to allow for digital standards in addition to physical standards. It's not listed here because it's been just recently announced, but we recently leveraged our compendium along with AI pipelines to deliver the Method Connect platform, which essentially allows compendial testing standards to be directly integrated with digital products. We've worked closely with the FDA in supporting the global substance registration system, and we have a number of in silico modeling services.

Within the industry, there's a large number of applications for digital-driven techniques. Foremost, particularly when we're talking about large language models and AI is document intelligence, authoring, and digitalization. There's opportunities for improving our quality processes, improving our retrosynthetic routes, analytic technologies such as chemometrics and process analytical technologies, and finally modeling processes and de-risking scale-up.

Related to the IID, we also are quite involved in the setting and guidance of nomenclature. So one of the things we do find is any sort of data-driven method relies heavily on consistent terminology, unambiguous terminology. And to that end, we support the Expert Committee on Nomenclature and Labeling. We provide nomenclature guidance, which we strive to remain lockstep with the GSRS. Similarly, we have a program for fostering novel excipients in drug products, including the Novel Excipients Knowledge Hub. We've recently released a stimuli article, which is currently open for comments on approaches to

excipient toxicity assessment. So for example, that includes a whole evidence approach where the IID can become quite useful, and a broad excipient standards catalog.

One of our key viewpoints around leveraging digital tools is that they are just that, they are tools. They must remain anchored to standards and governance and support scientific justification rather than replace scientific justification. Essentially, these are tools to work hand in hand with the scientists to be used within a risk-based validation, monitoring, and governance framework. And in particular, when we're talking about a number of these models, considering both positive and negative validation, as well as input data validation. I think one of the challenges with applying some of these models is essentially ensuring that the model has the humility to essentially recognize when it has received inputs that are edge cases or inputs that it should defer to a human. And ultimately, we want to ensure that the outputs are grounded in reliable data and standards. So for the case of an AI application where we're, say, summarizing some input data, providing that traceability from the output to that source data, and traceability that is specifically visible to the end user so that they can follow that and validate that and verify that.

Within any sort of digital technology, it must exist based on a strong underlying data enablement platform. And there are a number of challenges to be addressed in establishing those sorts of platforms, challenges which I see the team here at the FDA is pursuing to address, such as unclear data provenance and prior transformations on the data. So being able to trace back exactly what was done to the data to get from the source to the output. Linking across data sets, so using, for example, controlled taxonomies, controlled foreign keys to associate with related data sets. So, for example, the presence of the UNII within the IID allows us to then trace back automatically to the rich data in the GSRS. Avoiding biased and censored data, and that connects pretty closely to being able to trace that provenance and knowing where the data comes from, and to the degree possible, avoiding as much as possible, manual ingest and cleaning steps, both for efficiency and also because that's where errors tend to get introduced. So, for example, moving away from manual export into program interfaces and APIs, or in the case of making some of these tools available to agents, providing a model context protocol interface in addition, and ultimately ensuring that those ingest pipelines are well validated.

As we apply these considerations to the IID and the maximum daily exposure, there's quite a number of advantages that we see, such as automated extraction of input data from the new drug applications and the labels. This is one area where artificial intelligence models really shine, is being able to kind of pull out and extract that data and move it from an unstructured into a structured form. That allows direct calculation of derived data, such as the maximum daily exposure to be calculated from that structured data. Being able to

integrate directly those various data sources. So for example, being able to integrate the pipeline with the GSRS, which I see Arti already included as well within her mock-ups. And then being able to present that combined data for a human in the loop review.

Of course, that means that there's going to be quite a number of challenges to solve as well, and things that will be encountered as the teams move forward in developing these ingest and model pipelines, such as ensuring that reliable extraction across a variety of formats and documents is created and validated, resolving data challenges such as unit and naming mismatches, ensuring consistent definitions of calculations such as the MDE. So, you know, a model isn't going to be any better than a human at resolving definitions that are just inconsistent. Ensuring that that human in the loop user experience is well developed. We have seen within the history of software development that people do become very used to just clicking through message boxes, so we want to make sure that those users are qualified and the user interface is really supporting a robust review. And finally, recognizing and handling edge cases that need to be elevated directly to a human. And with that, I would like to thank the FDA for their invitation to speak, as well as all of you.

Meng Hu, PhD: Thank you. Let me introduce the next speaker, Mister Tushar Nitave. Mister Tushar Nitave brings over five years of data science and software engineering expertise to his current role at the Vivpro Corp. Mr. Nitave.

Presentation 5: GenAI Opportunity to Derisk and Accelerate Generic Product Development

Tushar Nitave, MS (*Staff Software Engineer, Vivpro Corp.)*

Thank you, Dr. Hu. So I'm Tushar. I currently work as a staff software engineer at Vivpro. And today we are going to see how can we accelerate and de-risk the generic drug approvals using AI.

So, the goals for today, the agenda that I'd like to talk about are what are some of the practical uses of AI, especially in addressing key pain points facing complex drug development? What are some AI playbooks for a winning strategy? For example, decreasing the risk, increasing the think time, how can we make the submissions faster with increased quality, and how can we build a CRL risk radar? And then discuss a practical use case where you can use AI to expedite the filing, reduce the risk of rejection, and achieve the first round of approval. And that pretty much means you can get faster access

to the market. And then we can have a look at a call to action, where how can we use knowledge as a data set, eliminate redundant tasks, and increase traceability by design.

So, if you look at the GAO data, we find that about 12% of about 2000 generic drug applications that are reviewed by the FDA, only 12% get approved during the first review cycle, and about 88%, which is roughly about 1800 drugs, do not get approved. So what are some of the key pain points in the generic drug industry? So there are about 462 drugs approved in 2025, generic drugs, out of which, let's say, 870 received CRL letters. 56 are about complex generic drugs, and 37 approved generic drugs never get launched. So how can we solve some of these key pain points that the generic industry faces at the moment?

So let's see how can we increase the think time and reduce the risk via intelligence. A couple of speakers before me have already talked about using the unstructured data, using the extraction methods. So there is about, there's a goldmine of data, unstructured data, that lies in about, let's say, millions of pages of FDA SBAs, about thousands of pages in the product-specific guidances. So the way we can leverage AI is to look what has worked for others, what mistakes others have done. And as you see here, there are about thousands of pages of product-specific guidance available. All we have to do is collect the data, and as was mentioned by other speakers, we put the intelligence on top of this data. We synthesize this data, we identify patterns, and then let AI do the intelligence stuff. So that uses an interface where we can query all this data instead of having someone to do a manual review. The way we used to do so far, we just simply query using natural language that we are much more familiar with, and that pretty much increases the access to the knowledge, which helps us to think more and helps us to rather than doing a guesswork, have a precedent-driven strategy.

How can we increase the consistency using automated writing with the help of artificial intelligence? Some of the mundane work of copy-pasting can be easily transferred to AI. One of the speakers before me showed an agentic workflow. Agents are pretty much capable of doing repetitive tasks at an extremely fast pace, and with increased efficiency, that of course we can never match. So that's how we can have access to lightning-fast document creation, which is where AI helps to generate the first draft at an accelerated rate. For example, typically it takes about 5 to 45 days to generate a draft for a clinical study report. And I'm talking about a draft, not even the final submission. On an average, let's take 17 days. With expedited and automated writing, that time can be reduced to less than one day, which means, let's say, within 24 hours, you can have the first draft ready to review and work upon until it's ready for submission. That's a pretty huge saving in time, rather than someone manually writing or copy-pasting the data.

We can build a CRL risk radar. Once, let's say your submission documents are ready, you can upload your documents, and since, as I showed you previously in previous slides, we already have a tremendous amount of data with SBAs, product-specific guidances, the radar can always look for gaps and deficiencies. It can look for labeling inconsistencies, and it can guide authors in the submission and find the gaps in the submission document to make it more foolproof, so that it passes the inspection and it has all the data that FDA needs.

So since we looked at the practical use case, let's look at the strategy that we can use to make sure all the things that we discussed so far can be integrated into the first is design selection. Again, starting with the product-specific guidance, looking at the precedents that the data that we have before us can help us prepare an accelerated regulatory strategy. Once we have the strategy with us, we can use AI-powered automated content writing. Again, that gives us faster access to draft versions, which means we can have the final drafts ready within no time. And then post-lock integration, we can generate more tools and CMCs that align with the labeling. So all these things, they reduce the rework cycles that we have, they give faster internal sign-offs, and they increase the collaboration between internal reviewers and increase the readiness of the data.

So again, how can we accelerate and de-risk the generic drug approvals? Transparency, FDA transparency, meaning if we can try to publish, continue to publish aggregated themes, having similar patterns, shared learning with anonymized CRLs. Having an AI-first culture, like treating documents as data. One of the speakers also mentioned that the data that goes in is the same data that comes out. So having a more structured data, having traceability into authoring. Human in the loop was a major point by other authors as well. And then growing skills in the input design, output data validation, and cross-document consistency. Thank you, thank you, FDA, for inviting me.

Meng Hu, PhD: OK, thank you. Let's welcome the next speaker, Doctor Sandip Tiwari. Dr. Sandip Tiwari is a pharmaceutical scientist and industry leader with over 26 years of experience in formulation science and drug product development. He currently serves as the head of technical service at the Pharma Solution at BASF. Let's welcome Dr. Tiwari.

Presentation 6: AI-Assisted Generic Formulation and Nitrosamine Risk: Algorithmic Workflows

Sandip Tiwari, PhD *(Head of Technical Services, BASF)*

Thank you, Doctor Hu and FDA, for this kind introduction, as well as invitation. What I'm going to speak today is on AI-assisted generic formulation and nitrosamine risk, how we

can use algorithmic tools for risk assessment. Again, I think I'm sure by now we know that nitrosamine is not necessarily an impurity issue now. It is a broader CMC challenge. And that broader CMC challenge, if not addressed properly, it has the potential to in fact put your product at risk from commercialization, right? So, I'll try to make it very simple and straightforward.

What we will look here is a very brief intro on nitrosamine. Many colleagues have already talked, so I'll keep it very brief. When we talk about nitrosamines, you know, it is not about API. It is also about the composition as a whole. What other drug components are there? What manufacturing process is used? What packaging components are there? All there is an interplay of all of these contributing factors that determine ultimately the nitrosamine risk for a drug product, right? We'll also briefly discuss about the digital tools, the machine learning-based tools that can be used to predict the risk to a drug product. Again, remember that these digital tools are tools. They are not a substitute for actual testing, right? I think this has to be made clear, and I'm sure many colleagues have already clarified that point. We will look very briefly about the research gaps in this segment. And finally, conclusions.

We know that nitrosamines are now, you know, at the interplay of very complex structures that need to be identified at very low levels, right? If you look at, you know, certain impurities, the limitations that FDA has set out is like 18 nanograms, you know, very low. So, we need to have very sensitive analytical methods as well, right? And I think that also creates challenges. Nitrosamine formation in a drug product is very simple and straightforward, right? What you need is three things: you need a nitrosating agent, you need a vulnerable amine, either secondary or tertiary, and then you need a favorable condition. Now, when you look about a drug product, what is a favorable condition? A favorable condition is nothing but pH, moisture, temperature, processing conditions. You know, how intense is your process? Are you introducing heat or stress during your manufacturing process? And what packaging components are there, right? That is going to ultimately determine if nitrosamines are getting formed in your drug product.

I think my colleague in the morning from industry already, you know, very nicely explained about how do you de-risk a formulation, right? Now, when you look at the components of a drug product, there are many things you need to look at. You know, besides API, you also need to ensure that you're using excipients that will not contribute to a nitrosamine impurities, right? And I think you need to also take into account that different suppliers may have different impurities, right? So that needs to be taken into account. And once we identify that, we need to use appropriate scavengers, inhibitors, or, you know, if we know the source that it is pH or moisture, we need to mitigate those, right? And ultimately, I think

manufacturing process, it's very important to make sure that we use a manufacturing process that will not accelerate translation of these impurities into nitrosamines, right?

So I think when you look at nitrosamine impurities or NDSRIs, nitrosamine drug substance related impurities, the question that a formulator gets most of the time is this, you know, okay, what nitrosamine related drug substance impurities are possible if I look at a structure of a drug product? The next question he has is, can I calculate the acceptable daily intake dose of those NDSRIs, because that is a requirement, and also the maximum daily dose. Am I able to calculate that? And then lastly, if the risk is identified, you know, how do I reduce the amount of those impurities or the risk to a drug product, right? Can we use any digital tool to predict? Again, prediction is for risk management and risk mitigation, not a substitute to actual testing. So if we are able to mitigate and predict the risk, we can plan and strategize better for early resolution, right?

So, keeping that in mind, BASF has developed a digital tool which is AI-aligned. And what this tool does is basically it predicts the chemical structure of nitrosamine drug substance related impurities based on the SMILES code of an API. And it also helps calculate the maximum daily intake limit. And most importantly, third, it also incorporates the considerations for excipients and also the manufacturing process. What excipients' impact will have on a drug product? What happens if you change your manufacturing process? So we are able to predict that, and we are able to now use that prediction as a tool to strategize your risk mitigation strategy.

So there are two tools here. One is degradation product of active ingredient, and what it does is that you simply input the SMILES code, which is nothing but a linear text that a computer understands as a structure of a compound. And once you enter that, it considers multiple machine learning models or chemical reactions based on the nitrosating agent and predicts the chemical structure of NDSRIs that are possible. Now, once those NDSRIs are identified, what you do is input them into the next module that is called as Nitrosamine Risk Assessment and Mitigation, where you input those impurities together with the excipients you are using and also your manufacturing process, the total dose of your tablet. And then it predicts based on that information what nitrosamine drug substance related impurities are possible. It also takes into account the impact of manufacturing process. And I think that is the beauty of it.

So this is just an illustrative example of an out workflow. My goal is not to go into the details, just to show you a snapshot. What you need to do basically is simply input the information about the API and then on the right-hand side you will see that it is already predicting the impurities there, you know, the structure based on the SMILES code. And then the next step is you are able to input, you know, on the left-hand side, you will see the

name of the excipients, the total dosage form, and then it predicts what is the impact on those acceptable daily intake doses. You will see that, you know, on the top and the bottom slides there, that, you know, based on the manufacturing process, if it is direct compression versus wet granulation, and the excipients, you know, the nature and the chemistry of excipients, you are able to predict the acceptable daily intake dose of those impurities.

So I think if you look at the research gaps now, you know, this is a complex interplay of multiple factors, right? It is not a simple impurity issue, as I discussed. We need to take into account various interplays of API, its interaction with excipients, and also the manufacturing process, and also the packaging components. So how do you integrate that all together in one setup or one tool that will help you predict? That is a challenge and there is still a gap. Excipient nitrate measurements, you know, different vendors use different methods. There is no standardized method. The quantities are very small, in nanograms, and I think the expectations and that is where the challenge is. And I think, of course, industry, regulators, and colleagues like Jeremy, we need to work together to standardize it. Process risk relationships, you know, I think when you have different manufacturing processes, how do you link that risk and how do you develop a predictive model based on that risk? And again, you know, like I think when we have this interplay, how do you ensure that the potency prediction and acceptable daily intake dose will have the confidence of regulatory agencies? That is the challenge. And most importantly, how do you integrate these multiple factors into one prediction? That is the bigger challenge.

So our goal is basically to shift from reactive testing to some tool that will help you prioritize your strategy for risk mitigation, because you can put your entire product and portfolio at risk if not handled correctly. And again, I think this is my concluding slide, and you clearly see that, you know, NDSRI risk is a CMC issue that needs to be handled early. And again, we need to develop scenario-based models so that we can prioritize the experimental work and mitigation strategies. And only when we have standardized data sets and validated models and harmonized expectations, we will be able to tackle this using these digital tools. And I think I'm sure that will help prioritize the strategy for mitigation. Thank you again, and I really appreciate FDA for this opportunity and look forward to the discussions.

Meng Hu, PhD: OK, thank you. Let me introduce the next speaker who will present online, Doctor Brooke Langevin. She is an assistant professor in the Department of Practice Science and Health Outcome Research at the University of Maryland. Her work focuses on applying quantitative modeling and simulation to inform drug development in data-limited settings. Let's welcome Dr. Langevin.

Presentation 7: Using AI to Address Practical Challenges for Generic Drugs: Using Digital Twins to Serve as RLD

Brooke Langevin, PhD (Assistant Professor, University of Maryland, Baltimore)

Hi, thank you very much for the opportunity to present. So I'm going to be looking at using AI to address practical challenges for generic drugs, specifically what to do when the RLD is unavailable.

So when the reference listed drug is unavailable, what ends up happening is that there are fragmented evidence sources that lead to an inability to determine BE because there's an inability to conduct the test between the test product and the reference product. Overall, this can lead to delayed generic entry and most importantly, reduced patient access. So when this occurs, there are a few potential solutions that have come up. One is when available, sometimes there are reference standards that can be used, such as different dose strengths or formulations. However, in some cases, that's not available. So we can use the available data such as innovator or legacy studies using statistical or model-informed approaches in order to try and bridge this gap. However, some of these alternative approaches rely on assumptions about the reference that can impact the bias, type 1 error, power, and require explicit evaluation that can be somewhat difficult in certain cases.

Because of this, one thing that we've looked at is a qualitative framework to outline the process of what could occur when the reference listed drug is unavailable. So the first step of the framework would be to construct a reference bank. So you'll need reference data in order to compare the test product to. This would be a virtual reference bank. The next step is to characterize the data from the reference bank and the test data from the test product. And then you would use a statistical or a model-based approach to be able to conduct the BE assessment. And then you would draw a BE conclusion. And one factor about this entire process is that each individual step can be, can take a very long time to occur and can require very specific, precise technical expertise. So what we're proposing is the ability to summarize these steps within an overall AI framework in order to operationalize the process and to speed the ability to make this BE conclusion.

So I'm going to go through one of the steps in the proposed AI framework a little bit more in detail. So looking at the first step, which I think is the most intricate or complicated step, which is actually creating this virtual reference population. And looking through it step by step, we would first start with an AI-assisted evidence extraction layer. So what does that mean? I mentioned that fragmented data that occurs throughout the available information.

So you may have things like literature information, the FDA may have access to certain aspects of the raw data, you may have things like the dose, formulation, sampling schedule of both the reference product and the intended test study that you're going to do, as well as data within your test population. What are the covariates? Is there a food effect? Is there a renal or hepatic effect and other study design features? So the goal would be is you could put all of this information that is coming from the different sources into this AI-assisted evidence extraction layer. From this layer would be a very standardized, detailed format indicating and summarizing all of these characteristics, including things like the AUC, Cmax, and all of the covariates through the subjects.

You would then be able to put this extracted piece of information into the reference layer. And the idea is that the reference layer is the layer that can generate these reference profiles. So it could use a popPK or semi-mechanistic approach in order to select a method that would allow for the generation of these profiles to capture all of the exposure metrics needed for the BE assessment. And it would simulate these reference subjects with covariates, different PK profiles, and all of the important BE metrics. Most importantly, you could do this many times and get many different sets of subjects. So here you can see the output is color coded to represent, for example, patients of different covariates, and you can see different simulated tests performed.

You could then put this layer into the curated virtual reference bank layer. And what you can see coming out is the bank can be filtered and weighted or stratified to ensure that for a proposed generic study, the workflow pulls from a matched reference cohort from the bank based off of the intended test covariates, sampling design, route formulation features and uncertainty. So for example, if we imagined that patients in the blue colors represented the covariates that we saw in our test subjects, we could filter out the green profiles or weight some of the other profiles differently if we imagine those belonged to a different group. In this way, you could actually have this curated final reference bank layer that accurately represents both the reference product and the intended sampling design for your test product. And so this process leads to these final profiles with AI metrics, starting from this fragmented data across multiple sources and ending with these final profiles. These final profiles could then be put into a model master file that could be distributed for a variety of generic companies to be able to use. So this overall would allow for easier access to this new reference bank and would allow for companies to be able to develop generics based off of it.

Now, one factor to keep in mind is that, as I mentioned, not every simulated subject should automatically contribute equally to the virtual reference bank. So each digital twin or each digital profile can be evaluated for how well it represents the intended reference

comparison. So some things to keep in mind are, for example, the population similarity. So does this digital twin profile match the intended population on key covariates? The exposure plausibility. So are things like your AUC, Cmax, Tmax, and profile shape within the plausible reference ranges? So this is making sure, again, that check that when you generated these profiles, you're actually getting back reference profiles that match the original drug. Then looking, is the twin anchored to that observable data? So is not just the individual layer, but is the full population within the observed reference data. And then you can also factor in things like how much uncertainty is surrounding the model that was picked, looking at if the twin can support the proposed test BE endpoint, and the actual influence of each individual layer. So would one profile disproportionately affect the end BE conclusion. And the beauty of being able to use an AI approach here is all of these factors could be checked internally within the system and raise any flags necessary. The goal is to make the generation of this large reference bank something that can be easily operationalized within an AI framework and doesn't require a large amount of time and a large amount of effort from a human to put together this bank.

And then looking at how this agent fits into the overall quantitative framework. So once you have this reference bank and you have your intended test product study, you can then put it into the overall model. And that represents that first step of the reference bank construction and looking at the data characterization. So those first two steps are accounted for using the process we just went through. Then the AI agent can look at the best approach for the MIBE. So whether that be a model-based approach or more of a statistical resampling from this initial reference bank. And the AI agent would actually be able to test both approaches and compare the approaches in terms of a previously generated example where you would have been able to look at a known ground truth and been able to evaluate it forward. So you would be able to select the best method for the given product for the actual MIBE approach. And then you would actually be able to conduct an assessment. You'll notice that these first steps are within the AI overall agent, because as mentioned previously by several of the speakers, AI is not a substitution for the judgment and the human aspect that is necessary. So once the virtual BE test was conducted, humans could then evaluate each step of the process and the conclusion and determine BE virtually.

Some of the strengths are that it's an operationalized framework, so by being able to perform each step within...

Meng Hu, PhD: Hi, Dr. Langevin. Sorry to interrupt. It would be great if you can finish your presentation in 30 seconds because we have more speakers. Thank you.

Brooke Langevin, PhD: Yeah, sorry. So the strengths are that it's an operationalized framework and it evaluates the methods under controlled conditions and the limitations are going to depend on the available data. And the overall conclusion is that it allows for a timely generic entry and improved patient access. So thank you.

Meng Hu, PhD: Thank you. Then let me introduce our next speaker, Lieutenant Commander Geng (Michael) Tian. He served as a supervisory research officer in the Office of Pharmaceutical Quality within CDER. In this role, he collaborated with industry, academia, and government agencies on research to advance innovative manufacturing technology that ensures drug quality. With that, welcome.

Presentation 8: AI/ML Regulatory Perspectives: Challenges and Considerations

Geng Tian, PhD *(Supervisory Research Officer, DPQR VI/OPQR/OPQ/CDER/FDA)*

Good afternoon, everyone. My name is Geng Tian from the Office of Pharmaceutical Quality at CDER, FDA. And today, I would like to share with you some of the challenges and considerations surrounding AI and machine learning from a regulatory perspective.

OPQ's mission is to assure quality medicines are available for the American public. Since 2022, FDA has launched several key initiatives to support AI and machine learning. And we're actively developing guidances and policies for AI and machine learning. And we are also highly engaged in international collaborations on AI and machine learning with our global partners. And we have launched several AI tools to enhance our internal capabilities. And we have been hosting workshops and scientific discussions on AI machine learning to foster dialogue with our stakeholders. And last but not least, we're providing funding to support extramural research on AI and machine learning.

In OPQ, we began our AI machine learning science and research in 2020, and we have identified several key objectives. First, we're working to understand how AI and machine learning is being used to support product quality, particularly in drug manufacturing and complex product characterization. Second, we're identifying, assessing, and managing potential risks associated with AI and machine learning use. And third, we're supporting regulatory assessment of submissions that use AI machine learning for product quality assurance.

Now, I'd like to highlight several key FDA documents. We started with two discussion papers in 2023. First one on AI in drug manufacturing, and then another one on the AI in drug development. Those two discussion papers really laid out our initial thinking and

invited public feedback. In March 2024, FDA released a white paper on how CBER, CDRH, and OCP are working together on AI across different medical products. In January 2025, FDA issued a draft guidance on considerations for the use of AI to support regulatory decision-making for drug and biological products. In January 2026, FDA and the EMA published guiding principles of good AI practice in drug development. In February 2026, CDER announced its guidance agenda for 2026. So a new AI guidance focusing on pharmaceutical manufacturing is coming up soon, so stay tuned for that.

Our 2023 discussion paper identified four major areas where AI could potentially transform drug manufacturing. First, process design and scale up, where AI can accelerate process by conducting virtual DOEs, understanding complex relationships between the process and product quality attributes. And second, process monitoring and fault detection, where AI enables real-time monitoring and control, root cause analysis, and predictive detection. And third, advanced process control, where AI can implement real-time process control that automatically adjusts the process to maintain in the state of control. And last but not least, supply chain, where AI can predict drug shortages, predict trends, and optimize scheduling. So, all of those applications show how AI can make drug manufacturing more efficient, robust, and responsive.

Today I'd like to discuss two key regulatory challenges and considerations for drug manufacturing. First is the AI in the pharmaceutical quality system. So the challenges are real. Implementing AI in the pharmaceutical quality system requires integration with existing systems. But those systems were not built with AI in mind, so we might often run into compatibility issues with legacy systems. And there's a need for properly trained personnel who understand both pharmaceutical quality and AI technology. And the integration process is expected to be complex and lengthy.

So what are the considerations for addressing those challenges? First, use AI to augment human judgment, not replace it, and always have human in the loop and validate AI recommendations, especially for critical tasks. And also, the pharmaceutical quality system should provide the QA team all the information they need to make the right decision and ensure correct implementation through appropriate personnel training. The QA team needs to know how AI works, what their limitations are, and how to use it effectively.

And the second challenge is the AI model credibility assessment. So this is truly at the heart of how we evaluate and assess AI models. There's a perception of AI machine learning as black box models. You put your data in, you get your predictions out. So what's happening between is not always clear or explainable. So there's a need for publication of development and validation case studies. We want to see more real case scenarios of how AI models have been developed and validated successfully. And there are expectations for

software quality assurance and verification. So how to ensure implementing your software in your AI model is reliable and well maintained.

And there are considerations that can help. First, adopt a standard and guidance for evaluating AI systems, for example, using risk-based ASME V&V 40 standard and FDA AI drafted guidance. And then publish case studies on AI model development and validation. Sharing those can really provide practical insights and build trust across stakeholders. And also develop good AI machine learning practices, just like we have good manufacturing practice. We should have Good AI practice and develop trust in open source code through guidelines. We need to have clear guidelines for how to verify and trust that code.

So now I'd like to talk about how we address the second challenge using the FDA draft guidance. So the FDA draft guidance was issued in January 2025 and adopted key principles from ASME V&V 40 standard. It provides a seven-step risk-based credibility assessment framework. So, step one is to define the question of interest. This is to describe the specific question or concern being addressed by the model. And step two is to describe the specific role and scope of the AI model. And then we need to determine the AI model risk based on model influence and decision consequence. And now we need to develop a plan to establish AI credibility within the context of use. And then based on the model risk identified in step three, we need to execute the plan, carry out the model credibility assessment activities, document the results, and discuss any deviations from the plan. And then the last step, step seven, we need to gather all the evidence we have and determine whether the AI model is credible for the context of use.

Do I have time? Two minutes? One minute. Ok. This is a recent publication we just had, and just to show how to illustrate how the framework works in practice. So first we define the question of interest and the context of use. And then we determine the model risk. So for this model, the model influence is determined as medium because there are additional controls for monitoring API content. And the decision consequence is high because the blend uniformity is a critical quality attribute. And then non-conforming product will have a high impact on the product quality. And now we determine the model risk is medium. Next, we need to establish a model credibility assessment plan to establish the model credibility, including describing the model architecture, model inputs, and output. And now we have model training, and this is a very critical process to evaluate the AI model performance. Also, we have how the model developed, how the data is partitioned. So, in this case study, we have 70% partition to training and validation, testing for 15%. And those results just show that the AI model can provide very accurate representation of the system and nonlinear dynamics of the system. And then finally, we have to also tune the model parameters. And finally, we evaluate the AI model performance. And for the set point

tracking, the AI model showed the best performance against traditional control. And similarly, for the disturbance rejection, we also have the AI model having the best performance.

So, in conclusion, AI and machine learning in pharmaceutical manufacturing is rapidly evolving, so is our regulatory experience. So we should start with a risk-based approach. Always match your model assessment activities to your model risk and use standards and guidance and be clear about the context of use of the model and follow good AI practice for model development and validation. So get your data governance right and life cycle management matters. And thank you all my colleagues for contributing to this presentation. Thank you.

Moderator (Meng Hu): Thank you, Dr. Tien. Let me introduce the last speaker for this session, Miss Torrey Ward. Miss Torrey Ward joined the FDA as an investigator with the Office of Inspection and Investigation, conducting complex pharmaceutical establishment inspections. With that, let's welcome Ms. Ward.

Presentation 9: Advancing Pharmaceutical Quality System Assessment to Support ICH Q12 Implementation: Leveraging Evidence-Based AI Tools

Torrey Ward *(Consumer Safety Officer, DQI-II/OQS/OPQ/CDER/FDA)*

Good afternoon. Thank you for having me. My name is Torrey Ward, and I'm with the Office of Quality Surveillance. Today, I'll be walking you through how FDA is advancing its assessment of pharmaceutical quality systems, or PQS, to support implementation of ICH Q12, and how we're using evidence-based AI tools to do that more efficiently.

At FDA, we believe that everyone deserves confidence in their next dose of medicine. Pharmaceutical quality assures the availability, safety, and efficacy of every dose.

Here's our roadmap for the next 10 minutes. We'll start with a quick overview of ICH Q12 and why regulatory flexibility matters for generic manufacturers. Then we'll talk about how FDA is using AI-assisted tools to evaluate submissions more efficiently. We'll cover the PQS assessment framework, share some data on outcomes, and close with key takeaways for industry.

ICH Q12 creates a structured, predictable framework for managing post-approval manufacturing changes. It applies broadly. Innovators, generics, biosimilars, and qualifying drug-device combination products all fall under its scope. If your establishment has a

strong pharmaceutical quality system, you can manage certain manufacturing changes with less regulatory oversight. That means fewer submissions and faster implementation, all while maintaining product quality.

Let's briefly walk through some of the key concepts of ICH Q12. Established conditions, or ECs, are the specific parameters, attributes, and controls that are defined as necessary to assure product quality. If an EC changes, a regulatory submission is required. The Product Life Cycle Management Document, or PLCM, is where ECs and associated reporting categories are documented. The Post-Approval Change Management Protocol, or PACMP, is a pre-approved plan to manage future changes. And underpinning all of this is an effective pharmaceutical quality system. A robust PQS supported by appropriate scientific justification for proposed ECs and reporting categories are required to enable regulatory flexibility.

This decision tree shows how manufacturing process parameters are classified and what that means for reporting. Applicants have the option to propose their own ECs and regulatory flexibilities. That includes identifying specific elements to be designated as ECs, making changes to an EC, and proposing alternative reporting categories for changes to those ECs. Applicants can also determine what is not an EC, and those elements would be managed internally through the establishment's PQS.

Here's what regulatory flexibility could look like in practice. Take batch size scale-up for a drug product. Normally a prior approval supplement, but with regulatory flexibility that could be reduced to a CBE-30, allowing scale-up to be implemented faster. Or consider filling line equipment changes. Also typically a prior approval supplement, but with flexibility, those changes can be implemented just 30 days after notification. The regulatory flexibility demonstrated by these examples is enabled by an effective pharmaceutical quality system and supported by scientific justification.

FDA's evaluation of an establishment's PQS leverages a coordinated evidence-based review process, one that is increasingly supported by AI-assisted tools that improve consistency and efficiency. Now, let's discuss those AI-assisted tools that FDA is using to support evaluations. First is our analytics-driven supplement evaluation system, or ASE, which uses AI to efficiently prioritize incoming submissions. Next are a set of AI prompts developed specifically for PQS assessment. These extract and synthesize information available to the Office of Quality Surveillance, producing outputs that help assessors work more consistently across the board. And third is post-approval monitoring. Once the flexibility is approved, we don't consider that a closed loop. We use data analytics to continuously monitor establishment performance, tracking inspection outcomes and post-

market quality signals over time. Together, these tools support a more efficient, evidence-based review process, while keeping human assessors in the decision seat throughout.

So here I will walk you through our coordinated review process and utilization of these AI tools. When a submission comes in with proposed established conditions, ASE, or analytics-driven supplement evaluation, uses AI to triage and prioritize the submission for review. Applicants should flag proposed ECs clearly in the cover letter, with relevant FEIs listed in the PLCM. From there, the Office of Program and Regulatory Operations assembles an assessment team from the EC Coordination Committee, the Q12 Assessment Implementation Team, and an OQS PQS assessor. OQS then evaluates and finalizes the PQS determination, leveraging AI prompts to extract and synthesize relevant data, all prior to the goal date. A key point is that no mandatory inspection is required. We assess PQS effectiveness using existing FDA and international partner inspections, as well as other quality intelligence. And once an establishment is approved, the work doesn't stop there. Our post-approval inspection monitoring tool continuously evaluates inspection outcomes and post-market quality signals. We also provide targeted support for upcoming inspections through site dossiers, pre-inspection briefings, and SME support.

So now that we've reviewed the coordinated application review process, let's discuss the role of the PQS assessment. So why does the PQS assessment matter so much in this process? Establishments must demonstrate an effective PQS and sound scientific justification, as both are required before FDA can make decisions on ECs and reporting categories. At its core, the PQS assessment objectively answers a critical question: Can this establishment effectively manage changes throughout the product life cycle without extensive regulatory oversight? To do this consistently and rigorously, we rely on the ICH Q10 PQS model, a globally recognized framework grounded in ISO quality concepts, ensuring that our assessments are aligned with international standards.

ICH Q10 describes a pharmaceutical quality system that spans the entire product lifecycle, from early development through product discontinuation. At its foundation are two enablers, knowledge management and quality risk management, which power the four PQS elements: process monitoring, CAPA, change management, and management review. When these layers work together under strong management leadership, they don't just satisfy GMP requirements, they strengthen them, creating a proactive and continuously improving quality system.

To assess PQS effectiveness, we leverage information sources that span inspection data, post-market reports, and regulatory intelligence, some of which are uniquely available to our Office of Quality Surveillance. We review inspection data, which can include inspection reports, FDA 483 observations, and firm responses, and supporting documentation such

as site master files and product quality reviews. This gives us insight into establishment operations and compliance history. We also review post-market reports. Field alert reports, BPDRs, MedWatch reports, and recalls reveal how effectively a firm identifies and addresses quality issues. And other regulatory intelligence includes drawing on information from our MRA partners, 704(a)(4) responses, informants, and regulatory actions like warning letters and regulatory meetings. Our comprehensive assessment combines qualitative evaluation of how mature and effective the establishment's quality culture is with quantitative analysis of hard data, such as deviation rates and complaint trends.

And so what are we actually looking for when we evaluate PQS effectiveness? The next slide will break that down. So an effective PQS isn't just about having systems in place, it's about demonstrating that those systems actually work. And here's what that looks like in practice. An effective PQS reliably delivers products with the right quality attributes, maintains a state of control through monitoring and control systems, identifies and acts on opportunities for continual improvement, and applies knowledge management and quality risk management across the product life cycle. A PQS may be found ineffective if there's insufficient data showing that CAPAs are effective, particularly if repeat deviations are not being adequately addressed across the product life cycle.

So let's review assessment demographics. The program started with a pilot beginning in 2019, and to date, OQS has completed 111 PQS assessments across 12 applications and 33 supplements. Most are BLAs with 17 NDAs received to date. But as you can see, we have not received any ANDA submissions. We encourage generic manufacturers to take advantage of this pathway, and the data we've gathered so far is promising. FDA's experience shows that regulatory flexibility under the Q12 framework is both accessible and low risk. With 95% of the establishments that requested flexibility found to have an effective PQS, this transparent, objective framework offers high reward for manufacturers who have invested in robust quality systems.

In short, for generic manufacturers, ICH Q12 means less regulatory burden and faster time to market when changes are needed, while actively encouraging continuous improvement, process innovation, and investment in quality management maturity. And speaking of which, my colleague, Dr. Eric Twum, will be presenting Quality Management Maturity as a framework for driving operational performance and business excellence tomorrow in Session 4, so be sure to catch that.

I'd like to take a quick moment just to acknowledge the contributions of Dr. McGuire, director of OQS, the ECCC, and the entire OQS-PQS assessment team. With that, I will end my presentation. Thank you all for your time today.

Meng Hu, **PhD**: All right. Thank you, all the speakers. So, according to the agenda, we will have a 10-minute break. Please note, we are not done yet. We will come back in 10 minutes, and Dr. Qing Liu will kick off the panel discussion. Thank you. See you soon.

Panel Discussion

Qing Liu, PhD (Moderator): Hello! Welcome back. OK, we are going to start our afternoon panel discussion. We, in addition to all the speakers, we have five additional panelists joining us. The first one is Doctor Joga Gobburu. Doctor Gobburu is a professor at the University of Maryland, Baltimore in the School of Pharmacy and Medicine, and the co-founder of Pumas-AI Inc. and Vivpro Corporation. The second panelist is Doctor Sharmista Chatterjee. Doctor Chatterjee is the Division Director in the Division of Pharmaceutical Manufacturing Assessment III within the Office of Pharmaceutical Manufacturing Assessment. The third panelist is Doctor Mihir Jaiswal. Dr. Jaiswal is a senior FDA leader specializing in AI, advanced analytics, and enterprise digital transformation within the Center of Drug Evaluation and Research. The fourth panelist is Doctor Zhen Zhang. Doctor Zhang is a master pharmacologist in the Division of Bioequivalence I within FDA's Office of Generic Drugs. The fifth panelist is Doctor Robert Lionberger. Doctor Lionberger is the Director of the Office of Research and Standards in the Office of Generic Drugs. He leads the implementation of GDUFA science and the research commitments. So welcome all the panelists and our speakers. So because we only have half an hour for the panel discussion, we will start with a question for our panelists first, and then we will take any questions from our audience. So the first question is for Arti and Jeremy, because you mentioned this in your presentations. And also, for Diana, because you work with IID. So the first question is, how can AI be used to deal with inconsistent IID naming conventions? Because as we all know, right, the IID has records for the same excipients, but with multiple different names, and also sometimes a record with a name that is not commonly used. So when anyone searches the IID database, and then they cannot find a match, but it's actually there. So the question is, how can we utilize AI to solve this problem?

Arti Patel, MS (Cosette Pharma): Yes. So already IID database has a UNII number listed in the database. So if we go through with the UNII number, it will reduce chances of evaluation of the same excipient for the multiple times. And that is easily acceptable from the available database itself.

Qing Liu, PhD: Thank you Arti.

Jeremy Evans, PhD (USP): Yeah, I think Arti kind of jumped on one of the key capabilities that I had in mind in terms of that link to the Global Substance Registry System. And so, of course, the UNII, you know, if that is duplicated, then that's a clear signal, but also, you know, having that link allows you to leverage the rich data within the GSRS, such as the, you know, the vast synonym database. You know, you could leverage fuzzy matching or AI synonym generation as a way to sort of pick some of those up. It might not be perfect, but I think it'll really help to kind of bring down the number of duplicates.

Qing Liu, PhD: Thank you, Jeremy.

Diana Solana-Sodeinde, PharmD (FDA): Can you hear me? Okay, great. Yeah, thank you very much. That's a great question. So exactly what you said, we call it the Unicode, but again, I learned something today. It's the UNII, right? Which stands for the Unique Ingredient Identifier, and it is listed on the IID system when you do a search for the specific excipient. So that's one way to, you know, use a unique identifier, right, when trying to eliminate the gaps with unstructured text. And then also the CAS number as well. The chemistry, the Chemical Abstracts Service number is also an identifier to use for the excipients. So we hope for more AI consideration, right, and tools that can help identify those errors or those naming convention mishaps. So thank you.

Robert Lionberger, PhD (FDA): So I have one comment on that. So if you look, I think the IID does the excipient nomenclature well using the standard dose, the standard identifier. But I think if you look at our inactive ingredient database, it has routes and dosage forms that are sort of, let's just call them random. They're not found in the orange book. And they come from the data source that was used to produce that. So I'd like to hear from the people outside FDA, like, is there a value in normalizing and cleaning up the dosage form and route information in the IID, in the same way that the unique identifier is now used for the excipient.

Arti Patel, MS (Cosette Pharma): Yes, definitely. Because before when we were going on the IID database for the same excipient, we were looking for the multiple entries. And for the dosage form also like minor, minor changes, but it was coming with the multiple entries. So when we came with going with the route of administration, as per the SPL database, structured database, it helped a lot during the evaluation, so that's helpful.

Joga Gobburu, PhD (University of Maryland): I wanted to make a more, you know, a broader comment, listening to all these talks, that if you think about this from an MBA perspective, that these are all use cases to prove the power of AI as well as to gain confidence in what it can do. And to a point where we want to industrialize as much as possible the integration of AI into the regulatory practices. From my own experience

working with AI-based software and such, one of the key bottlenecks, in my opinion, is that the standards we use in terms of how we write reports, how we format our data sets, and simple things, you know, formatting tables or figures or the statistical data sets, et cetera, they're all geared towards humans using and interpreting that data. But they're not designed for machines. And hence, I think one of the biggest areas for all of us to work together like CDISC is going to be in the area of standardizing the input. Inputs, broadly speaking, documents, data, and what have you, such that they are machine ready, AI ready, that to me is where the next, you know, 5, 10 years is going to be a major focus. And for that, exactly, we can follow. I mean, FDA has always done that in a very transparent and collective manner. Same things we have done for CDISC and so on. We can embrace the same kind of path to come up with standardization for AI-ready data sets and documents.

Sandip Tiwari, PhD (BASF): thank you, I think I just have a comment. I think you asked a very interesting question, you know, whether the route information is essential. I think from a scientific perspective, I think route information is helpful, but I'm not sure if the dosage form information is helpful. You know, as long as the route is considered, I think it would be much better if the dosage form related information is either eliminated or abbreviated, you know, because that creates a lot of confusion for formulators.

Joga Gobburu, PhD (University of Maryland): I want to push back on that a little bit. The reason is that if you think IID is in isolation, then, sure, you can build a system that only works for IID, but if you think it's a global, you know, it's a much larger system where there is something upstream, something downstream, and everything is interconnected, then we would think differently and these do become important, these metadata or whatever you want to call them.

Qing Liu, PhD: Okay, so I think, well, we heard that for the IID database, there's some kind of suggestion regarding like merging GSRS and the IID is a suggestion I heard, and also maybe regarding the database itself, right? So there is a need to kind of streamline the format and make it easier for human and also for AI to read, right? So good suggestion. Thank you. Okay, so that's the first question. I wonder if there are any in-person questions from our audience. You may walk to the microphone. Okay, sounds good. So, Meng, do you want to… Okay. All right. So let's move on to the second question. I think this question is for Jenny. So what tools and data would be needed to develop the IID and MDD AI system?

Jenny Wang, PhD (FDA): Thank you for the question. Because currently the tool I presented is mainly focused on the MDD part, so I will just only answer the MDD part. For our current MDD tool, we used the ELSA, like the AI, the large language model AI system in FDA, so that's kind of the brain for that MDD tool. For agentic AI tools, we also have the capability to integrate like other engineering ways to better control the data quality and we control the

intermediate outcomes between the communications of like each agent using like a programming language, for example, like Python. Yeah, thank you.

Qing Liu, PhD: Thank you Jenny, Diana I wonder if you want to share what you think regarding the IID part.

Diana Solana-Sodeinde, PharmD (FDA): Yes, the only thing I just want to add is we are exploring the AI tools that have been presented to us by different offices in OBI. And like Jenny said, I know there's more to come, right? So right now it's the AI IngenTech tool that has been used. And then there are more applications that are being considered in the pipeline. So more to come in the future.

Qing Liu, PhD: Thank you again

Robert Lionberger, PhD (FDA): A related question to this. So as we've been working on this and other projects, many of the things that we're doing internally, they need access to what's the label of the drug. And so I'd like to hear from the people outside FDA on what do you think the current state of access to FDA approved labels is? What could be done to make that better? Are there any gaps? Are there any ways that that data could be organized better to make it more useful to people who are building AI tools or applications or for drug developers or people planning to submit ANDAs? And, you know, certainly you can also comment to the docket if you think, if you're listening and you think you have comments on this question, but I'm interested to hear people's thoughts on access to the labels that you need to do different things related to AI tools.

Arti Patel, MS (Cosette Pharma): Yes, one thing which I see immediately, that's the maximum daily dose. Because MDD is something like can be available on the SPL structure. And right now we are doing multiple CCs just to get the clarification on that one. Even we are working on the multiple AI tools to get the clarification. But whatever we do, it needs to be cleared from FDA, so those understanding if we get it into the PI or in the label, it will, it will be, I think I will say it's a blessing for the generic industries to get that clarification.

Sandip Tiwari, PhD (BASF): I think another question, you know, which is something really a little basic, is that many times when you look at labels, you know, it may have ingredients listed there, but those ingredients might be listed, some of them as trade names, some of them as generic names, so there could be uniformity there that would help. In addition to the manufacturing location, you know, if it becomes easier to identify.

Diana Solana-Sodeinde, PharmD (FDA): I have a follow-up question real quick in alignment to what Dr. Lionberger was saying. So what labeling sources do you use right now, just off the top?

Arti Patel, MS (Cosette Pharma): So, as guideline says, we have to refer Drugs@FDA site to get the approved labels, but sometimes we also need to go on DailyMed to find the all details with the excipients and everything. So, these are the two sources which we use.

Diana Solana-Sodeinde, PharmD (FDA): Thank you very much.

Sandip Tiwari, PhD (BASF): And also complementary prescribing information many times.

Qing Liu, PhD: Arti, from what you said, I was thinking about a system like you input your plan, like development of drug, like just the labeling of the RLD and you add your generic drug like formulation and then you get an answer right just there. That would be ideal, I guess. So let's move on to the next question. I think it was mentioned by many of our speakers, so it's an open question to all of you. So this is as AI shifts teams from drafting to validating, what specific human in the loop safeguards are required to ensure the accuracy and integrity of automated regulatory outputs? So it is about human in the loop.

Joga Gobburu, PhD (University of Maryland): Can you give me an example of what is a regulatory output?

Qing Liu, PhD: I think it's a regulatory decision, yeah

Joga Gobburu, PhD (University of Maryland): I have strong views on this one. So, if we think about the regulatory review process, there is some effort in terms of, if you make it into three steps. So first is assuming that submission is filed. OK, so first the reviewer is going to master the application, you know, perform all kinds of research and understand what was done and gain an overall view about the entire drug development evidence. Then, the second step is going to be the authoring that part. You can say maybe fifty-fifty is describing, summarizing what was submitted; that part is definitely AI can help generate. If there is a particular format, it can make things easier so that you don't have to go to 17 different places and get that information. But the third part, that has to be human, because that's interpretation. That's using judgment. I mean, it's like saying that the bioequivalence limits are 0.79 to 1.25. Do you reject it or not? It's not a machine decision. The opposite. So that part I don't think, I would urge that we preserve that to the professional judgment, SMEs. But the other two parts, you can make them more efficient, no questions.

Sharmista Chatterjee, PhD (FDA): So definitely agree with what you just said, Joga. So what we are seeing now, at least within our office and Office of Pharmaceutical Quality, and we can speak for our office, Office of Pharmaceutical Manufacturing Assessment in OPMA. So we are using ELSA, some of the AI tools that we have on site, for like data extraction. So we can pull out the information from the submissions, but the critical review piece to make the determination whether it's adequate or not, that has to be done by the assessor. And we are also using AI, we have some like, we have a structured way for like

prompt generations. When we send out like information requests, we have to follow the four-part harmony. So the assessor comes up with an IR, we can use ELSA, we can use the built prompts to make sure that it aligns with the four-part harmony. So it is being used to facilitate assessment, but the final decision is on the human. And I want to add, because I'm in manufacturing side, I want to say another plug for manufacturing, that in the manufacturing arena, AI can be used to optimize or to facilitate the process. The final decision is on the human. There was a famous, infamous warning letter, which was issued by the agency about two months ago, where the company had just used ChatGPT to come up with specifications. And there was no human oversight. So it really strengthens the point that having adequate human oversight is really important.

Arti Patel, MS (Cosette Pharma): I would just like to add up here that definitely oversight is definitely needed. One more thing, what the AI is expecting, need to have citations for those extractions. Because if something is coming which is not there and coming with the hallucination, then that need to be addressed. So whatever numbers are coming, whatever the sources are there, that need to be addressed and it should be come up as a score or whatever we establish in the system. But there is no room of hallucination when we talk about the AI in pharma industry. Because it is dealing with the safety, quality, and efficacy, where direct impact is on the patients and on the human being, so that should be also considered.

Jeremy Evans, PhD (USP): Yeah, I think the one thing that I would add as well, I think one of the places where there's a lot of room for work in, I would say across the entire AI industry, as well as within the drug industry, is in how we work with the models and the user experience in terms of how the models present their retrieved information to the user, because ultimately I think these are tools to augment a human and scientific judgment. So I think as much as can be done to put the critical information front and center and make it as easily understandable as possible, I think that's where we make some real wins. I think right now, you know, we see kind of the wall of text that comes down and it can be even difficult to interpret what comes out of these models. So I think having that sort of UX type work to sort of improve that experience and make it more digestible and understandable, I think is a huge win.

Mihir Jaiswal, PhD (FDA): Yeah, that's a very good point, Dr. Evans. And what Sharmista was mentioning about how we are using some of the prompts to extract the data or organize the information, that's our focus as well, because we just don't want to just get some text or data. It needs to be in certain structures, certain format that can be directly used. And to that point, the way we approach AI within OPQ, especially in terms of testing validation, is not necessarily testing or validating the use of the AI, but the overall impact on

the purpose itself. Because whether we are using AI or not, if we are getting the data out of submission, it just needs to be correct and correctly represented. So that's how we view it. It's not necessarily the prompt is 90% accurate, 80% complete or something like that. But ultimately, are we able to accomplish the task on hand with greater efficiency and able to get ROI around it? And, and that's how we define human in the loop, because ultimately human in the loop is not just validating the output coming out of AI, but if the use of AI is for high-risk tasks where the impact of something going wrong is going to be quite devastating, then we need to make sure that human in the loop is incorporated in such a way where it can change the behavior of AI itself, because ultimately the AI output or AI use is not important. It's the accomplishing of that task is important. So that's how we approach testing and validation from the task or workflow perspective, not just the use of AI.

Sandip Tiwari, PhD (BASF): Just a quick comment, Mihir, to what you just mentioned about prompts. It is also a general observation that besides prompt, even if you use two different tools, they might give you directionally different conclusions. You know, so I think that is also an area that needs to be analyzed and reviewed.

Zhen Zhang, PhD (FDA): Yeah, I want to make a comment around the similarities. So, yes, so when you want when you validate the outcome that generative AI output, so we often ask it to generate AI to give us a source data source, so we can validate. But one caveat is that, so what if generative AI missed something? So, without reading through everything, so we don't know what is missing. Yeah, that's one comment, and then it comes to Doctor Joga's comment regarding the format. So, I agree generative AI can help us summarize submissions. So, actually, we do actively test this idea, and so we upload study reports and ask generative AI to give us a summary table, and we see some hallucinations. Yeah, but it comes to the question whether we can ask applicants to provide this information, because applicants are more familiar with the study reports, right? They know everything. And so if applicants can provide all the evaluation criteria, evaluation tables that we use, and those evaluation tables actually you can find from public available reviews. And if applicants can provide that information, then it will be much easier for reviewers to check and to evaluate. So reviewers can focus on the most highly valuable areas to facilitate the overall review process.

Joga Gobburu, PhD (University of Maryland): I want you to be careful about that. The reason is, if you say this is my standard table in my review or for evaluation, and you only ask that be submitted, then if you think long term, what are the other regulatory needs in the next 5, 10 years? You may have to really go back a step or two to make sure that the source data for that table is what is machine readable for you, because this is only an

instance of that original source, so in the long run, if we only focus on this particular template, we might be losing the big picture.

Zhen Zhang, PhD (FDA): Yeah, I totally agree. This is not only submit summary tables, it's along with all the raw data and study reports. And also in the summary tables, also add the link. Yeah, so everything, raw data, chromatograph.

Audience Member 3: I have a question to ask to the FDA actually. See, my question is when we talk about the AI tools, when the AI tool is actually collecting the data of what you all have, and coming up with some standard guidance or anything. The question I have is, if you look at the standard review process, you know, five years ago, the review process was totally different. Today is different. Same product five years ago, was approved with the limited data, where today the same product will not be approved with the limited data because of requirement changing and everything. So my question is when AI is collecting the data and evaluating, assessing and coming up with a certain standard, what is the criteria the agency is allowing? What should be the critical parameter that this product should meet finally to get approval, let's say next five years? So because I lived in this and I know standard requirements have changed a lot for complex, even complex generic products. Three years ago, different, and today is different. So I'm just trying to understand, before we go to the next level of setting up guidance and going, what is the actual criteria? How is this AI tool coming up with evaluating the variances that existed before and now and in future it may be changing. What is the predictability?

Robert Lionberger, PhD (FDA): All right, well, again, as you heard from our discussion, right, the ultimate responsibility on the standards, that's what ought to rely on human judgment. And, you know, whether there's AI or not AI, things change over time. Products that may have been acceptable yesterday don't meet what we currently think about that. I don't think that's, you know, that's part of the current judgment that you get from FDA. I do want to loop back to the discussion on the human in the loop and say something controversial. Like, I don't think we should aspire for humans to be in the loop. We should aspire to be thinking about how we validate processes and establishments without. So give the example of, I want to extract the stability data from the application. Do we then want the reviewer to check that extraction every time? That's putting the human inside the loop. I want to put the human at the responsibility point. And I think of it just to use the OPQ language, you know, quality by design instead of quality by testing, right? How do you design really robust systems to do these things so that humans don't have to live inside the loop at the low level and they can be focused on the higher level? So I think that's an important, you know, research frontier as to how do you establish really robust systems that use AI, given that sometimes large language models are not quite deterministic

elements, right? And how do you account for that in your system as a whole? I think that's a very important point. But I also want us to aspire to tools that don't require humans to check every aspect of it. I think one of the things that AI allows us to do is to do things that from product quality and review aspects that we should be doing, but we just weren't able to do before because we just didn't have enough humans to do them. And I think if we think about that, then it opens some new frontiers and new ways of thinking about how we look at the data that's in applications. You know, I think from FDA's perspective now, we're trying to say, how do we help AI help our reviewers do the things that they're already doing better? But there's a world beyond that when it's, what should we be doing if we weren't limited by the capabilities of our reviewers alone? And there may be different ways to look at product quality or evaluations for that. So I encourage people to think about that as we focus on what the, you know, what the future uses of AI will be in, you know, in the evaluations of pharmaceutical development and our assessments of pharmaceutical quality and applications.

Sharmista Chatterjee, PhD (FDA): To the point that you mentioned that we really don't want every time if a model were to extract the information, you wouldn't want necessarily each time to go back and check the submission to see if it's extracted correctly. That should happen when you're doing the initial, what we call about the model credibility assessment. So when we're going through the phase of model credibility assessment, that's the phase when we would test to make sure that whatever the model, say, if it's used to extract the information from the application, if it's doing that correctly or not. But once the model's credibility has been established and it's in use, then we wouldn't have to keep doing that.

Joga Gobburu, PhD (University of Maryland): I mean, the assumption here, going back to all these points, is when you are using any tool, whatever it is, even SAS or R, you still have to go through that phase of making sure that it is accurate. Only then you will use it. So in that manner, there are several examples where it helps during the regulatory review process totally. Take a simple example of a meeting with a sponsor. Do we have access to the meeting minutes of a similar topic that some other team has gone through with another sponsor? At least several years ago, there was no system like that. I doubt if there is one now. But you could, that's where AI can leverage and then say, if there is a question from the sponsor, can it tell you at least the top five responses of what happened, you know, in most recently? That makes it actually improve the agency's consistency in terms of what type of advice that is given to the sponsors. That's an example. Second, from a new drug perspective, if you want to evaluate what is the benefit-risk profile of another drug for the same indication, it's not easy to go back and compare the adverse event rates or the efficacy rates and so on, when you really have the original data, but you cannot be just

confined to a PDF label. So those aspects that the first step in my three-step process is going to be where FDA would get the most benefit. Number two is, you said, what if AI doesn't give me the whole list? That is not AI's job. It will never give you the full list. You can ask it, give me all the list of drugs approved for a particular indication. It can never do that. That's an engineering solution. So you have to think about a hybrid way of making sure that there is a way you enable both search, which is like the Drugs@FDA, which definitely could use a facelift, then plus AI, that is going to be more definitive.

Jeremy Evans, PhD (USP): I think you, in your comments, I think you brought up a key kind of definitional concept in terms of what is meant by human in the loop. And I think what you're describing there, and it's quite accurate, is separating the retrieval from the judgment. Again, we're not, we don't want to hand over the judgment to the AI, but I think it is reasonable to expect that models can extract and summarize information that is given to it reasonably well. And even the advancements in the last year have kind of shown how far they're coming. And coupling that with strategies such as retrieving from multiple models and comparing, I think there are ways to get to a point where that is a valid, testable and validated process to say, okay, we're going to pull this information from the source, we're going to collect it, we're going to put it together in a way that is consumable. And then I think we're still in a place where the judgment piece we need to reserve for a human, but the parts that come before the judgment, I think, are automatable and then can be validated.

Qing Liu, PhD: Okay, anyone else want to share? We are running overtime. OK, so I guess we won't have any further questions, and this concludes our afternoon session. Doctor Lionberger will give our closing remarks.

Robert Lionberger, PhD (FDA): Thank you very much. I will be very brief because we're coming back tomorrow. But if you're not coming back tomorrow, on the tables, there's a questionnaire. So we ask that you please fill that out and drop it in the box if you're not going to be back here tomorrow. But again, appreciate everyone who contributed to our morning session on nitrosamines, very deep understanding of a lot of the issues related to that, and as well as our great discussion that we've had on the AI sections this afternoon. And with that, we'll see many of you here tomorrow, and for the audience online, please come back tomorrow for more of this same type of content. Thank you all very much.