PharmaSUG China 2025 - Paper 151X-B8B6H6F6C9

Interpretation of FDA Guidance for AI use in Drug development and Biologic products and a Real-World case discussion

Qiao Deng, Boehringer Ingelheim (China) Investment Co., Ltd;

ABSTRACT

AI has been playing an increasingly important role in the development of Drug and Biological products. To make the use of AI reliable and normative, FDA released a draft guidance which provides a risk-based credibility assessment framework. This topic is intended to interpret the Guidance and discuss a Real-World case of dermatological decision-making through AI-powered skin analysis which fits the framework.

INTRODUCTION

FDA released a draft guidance to provide recommendations for using AI to support regulatory decision-making regarding the safety, effectiveness, or quality of drugs. It introduces a risk-based credibility assessment framework to evaluate the credibility of AI models for specific contexts of use (COU). The key considerations for AI use in the guidance include a risk-based credibility assessment framework, the life cycle maintenance, and engagement with FDA.

The guidance is currently in draft stage, so we want to interpret it based on a Real-World case of dermatological decision-making through AI-powered skin analysis and discuss the potential challenges of AI use in Drug development and Biologic products. In this case, AI is used to develop the CE certified Software as a Medical Device (SaMD) which acts as automatic calculator of severity assessment for dermatological conditions. The challenges in the development are avoiding bias and integration.

FDA KEY CONSIDERATIONS FOR AI USE

AI use in the drug product lifecycle is growing, with applications such as reducing animal-based studies, predictive modeling for clinical pharmacokinetics, and analyzing real-world data or digital health technologies, etc. The involvement of AI use could also bring challenges, including dataset variability, bias, and reliability, methodological transparency, performance changes over time. To ensure AI models used in regulatory decision-making are credible and fit for purpose, FDA emphasizes a risk-based approach for sponsors and other interested parties and encourage early engagement with FDA to align on expectations and address challenges.

The detailed steps in the risk-based credibility assessment framework can be summarized as the following:

1. Define the question of Interest: clearly articulate the specific question or decision the AI model addresses.
2. Define the Context of Use (COU): specify the role and scope of the AI model, including whether other evidence will complement the model's output.
3. Assess Model Risk: evaluate model influence, decision consequence and use a risk matrix to classify model risk.
4. Develop a credibility assessment plan: the plan includes model description, data description, model training, and model evaluation.
5. Execute the Plan: implement the credibility assessment plan, ensuring alignment with FDA expectations.
6. Document results: prepare a credibility assessment report summarizing results, deviations, and evidence supporting the AI model's credibility.
7. Determine adequacy: evaluate whether the AI model is fit for its intended COU based on the documented evidence.

Two examples of question of interest, and detailed clarification for each step could be find in the draft guidance. However, there is not a case thoroughly discussed in the guidance based on the risk-based assessment framework. Therefore, next a Real-World case would be introduced to help reviewers better understand the workflow.

REAL-WORLD CASE DISCUSSION

Dermatological imaging is extensively for diagnosis and severity assessment, but could have huge variability caused by lightning, distance to the lesion, focus, etc. Such variability might in the end make an image clinically useless. Therefore, an AI-based tool was developed to assess dermatological image quality and ensure the image utility in remote situation or clinical trials.

This case described above happened prior to the release of FDA draft guidance. Nevertheless, the workflow and overall consideration can be fit for the risk-based assessment framework. Therefore, to begin with, the case would be introduced step by step according to the framework.

RISK-BASED ASSESSMENT FRAMEWORK

The detailed steps in terms of risk-based assessment framework for the case to be discussed in this article can be summarized below.

1. Define the question of Interest

Generalized Pustular Psoriasis (GPP) is a rare, chronic, life-long inflammatory disease with an unpredictable and diverse nature. Unfortunately, patients often experience delays before being diagnosed with GPP caused by misdiagnosis, lack of insurance, visiting multiple healthcare professionals to obtain diagnosis. Therefore, can we enhance dermatological decision-making through AI-powered skin analysis?

2. Define the Context of Use (COU)

The CE certified Software as a Medical Device (SaMD) is used to support professional differential diagnosis which perform skin severity assessment for GPP and identifying various dermatologic diseases. It is targeted to help healthcare professionals with dermatologic knowledge but not expert in GPP diagnosis, and public who already present dermatologic symptoms but lack specialist access or clinical awareness. The AI model is deployed in a dermatologic platform as a progressive web app. And outputs would identify more appropriate patients for company study drug and generate data and insights to drive data-based decision making. These outputs are available in a dermatologic platform as a progressive web app.

3. Assess Model Risk

The role of this AI model is to inspire and support the differential diagnosis of dermatologic symptoms for healthcare professionals or provide speedy and credible remote dermatologic suggestions for public. The consequence brought by this AI model is 1) identifying appropriate patients to receive company study drug, 2) generating data and insights to drive data-based decision-making. Therefore, model consequence risk is high. However, it not the sole determinant for identifying dermatologic symptoms. Instead, its outputs should be reviewed and validated by healthcare professionals and supplemented by traditional methods. Therefore, model influence risk is medium.

To conclude, even though the model is not the sole determinant, misdiagnosis of a severe disease like GPP would result in an incorrect decision in treatment. Therefore, the overall model risk is still high.

4. Develop a credibility assessment plan

The AI model applies a Convolutional Neural Network (CNN) to identify dermatologic symptoms. Input data for this model come from thousands of images for three visual signs of pustulation, erythema, and scaling, labeled by four expert dermatologists. The whole database is stratified for training (90%) and testing datasets (10%). The developed medical device is tested in the live environment by 15 experienced dermatologists to compare the assessment of 100 images between 15 skin pathologies Specialist alone

and Specialist with Medical Device.

5. Execute the Plan

The credibility assessment plan was executed as expected and got acknowledgement from national health systems in Spain and the US.

6. Document results

The automatic dermatologic severity assessment tool is considered to preventing unnecessary referrals and speed-up the urgent cases.

7. Determine adequacy

Doctors with the help of the automatic dermatologic severity assessment tool improve diagnosis accuracy by 2 folds compared with the traditional work. And the results are consistent across a list of most probable pathologies.

DISCUSSION

AI use has been playing an increasingly important role in the development of Drug and Biological products. Even though it significantly enhances the research efficiency, we should be careful about the risks caused by bias and difficulty in result interpretation, especially in regulatory submissions to health authorities.

The draft guidance of AI use released by FDA provides relatively comprehensive consideration in AI use which helps sponsors to get a clear view of the question of interest in a specific COU. It is also encouraged to have early engagement with FDA prior to and during the execution of plan to fully align on the AI model expectations and challenge resolution. Documentation is also important, which justifies the AI model development, evaluation, and maintenance processes, tailoring credibility activities to the model's risk and COU.

REFERENCES

Hernández Montilla I, Mac Carthy T, Aguilar A, Medela A. Apr 2023. “Dermatology Image Quality Assessment (DIQA): Artificial intelligence to ensure the clinical utility of images for remote consultations and clinical trials.” *J Am Acad Dermatol*, 88(4):927-928.

Food and Drug Administration Center for Drug Evaluation and Research. “Considerations for the Use of Artificial Intelligence to Support Regulatory Decision Making for Drug and Biological Products Guidance for Industry and Other Interested Parties.” Draft Guidance released January 2025. https://www.fda.gov/media/184830/download

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Qiao Deng
qiao.deng@boehringer-ingelheim.com