Article:
He S Yang, Fei Wang, Matthew B Greenblatt, Sharon X Huang, and Yi Zhang.
AI Chatbots in Clinical Laboratory Medicine: Foundations and Trends
Clin Chem 2023; 69(11): 1238–46. https://doi.org/10.1093/clinchem/hvad106

Guest: Dr. He Sarina Yang from the Department of Pathology and Laboratory Medicine at New York Presbyterian Hospital/Weill Cornell Medicine.

Bob Barrett: This is a podcast from *Clinical Chemistry*, a production of the Association for Diagnostics & Laboratory Medicine. I'm Bob Barrett. With the increasing complexity of clinical laboratory testing, patients and healthcare providers regularly have questions about test selection and result interpretation. Clinical laboratorians field these questions from medical personnel, but generally are not available to interact in the same way with the general public. One possible solution is artificial intelligence chatbots, available 24/7 with the capacity to provide medical information in real time in response to user inquiries.

While chatbots are widespread in customer support settings and have been explored in other clinical areas, their application to laboratory medicine is in its infancy. In order to provide a valuable service in the laboratory medicine space, the information provided must be accurate, consistent, and reflective of the subtle differences that make each patient unique.

Is Dr. Siri ready for primetime, or should we treat chatbot-provided medical information with caution? A new mini-review appearing in the November 2023 issue of *Clinical Chemistry* discusses approaches to build AI chatbots, proposes possible applications in laboratory medicine, and highlights challenges that must be overcome prior to successful implementation.

In this podcast, we're pleased to welcome the lead author of that mini-review. Dr. Sarina Yang is an associate professor in the Department of Pathology and Laboratory Medicine at New York Presbyterian Hospital Weill Cornell Medicine. She is a member of the IFCC Working Group on Artificial Intelligence and Genomic Diagnostics.

Dr. Yang, to the best of my knowledge, anyway, this is probably the first review published in *Clinical Chemistry* about AI chatbots. Can you first give us a sense of what you cover in this review?

Sarina Yang: Sure. First, I'd like to thank *Clinical Chemistry* Editorial Board for inviting us to write this review and also picking our paper for the podcast. Our review paper provides an overview of the history of AI chatbots, and two major approaches to develop chatbots and their respective advantages and limitations. We further discuss opportunities and challenges of using chatbots in healthcare with a focus on the field of laboratory medicine, as well as future trends and research directions to improve their clinical utility.

Bob Barrett: Can you describe how chatbots' underlying architecture affects their ability to execute their essential functions?

Sarina Yang: Yes. An effective chatbot must have some essential capabilities. First, it needs to accurately comprehend the user's queries in plain natural language. Next, it should be capable of generating relevant, logical, coherent, and accurate responses that align with users' request. These functions require the chatbot to have a robust understanding of relevant domain knowledge, as well as comprehension of a broader factual context, including common sense knowledge. Additionally, it should also be able to interpret the ongoing conversation and provide appropriate feedback. Furthermore, it should determine the right time and method for executing actions. Lastly, the chatbot must be interactive, facilitating an engaging conversational experience for the users. Therefore, developing a sophisticated AI chatbot is a challenging task.

There are two main approaches to chatbot development, a modularized approach or an end-to-end approach. Basically, a modularized approach integrates a few functional modules together. Each module performs a specific task and then the entire system functions by combining results from all these modules. For example, the natural language understanding module attempts to understand the meaning of sentences and transform them into structured format. Then, a dialogue management module receives inputs and direct the flow of conversation by determining system next action or response. It maintains and updates the memory of the conversational context, including the user's preference. Then it passes the output to the natural language generation module, which generates a response. Overall, the modularized approach allow developers to build flexible, robust and scalable systems, as individual modules can be updated or replaced as needed. However, the variability and ambiguity of human languages pose a significant challenge.

Human can express the same idea in many ways, using different words, phrases, and structures, and structurally similar sentences can have very different meanings. Also, during a conversation, the meaning of the sentence is often influenced by contest and preceding statements. Without a

grasp of the discourse, the chatbot may provide irrelevant or confusing responses.

On the other hand, the end-to-end approach involves developing a single integrated machine learning model, such as a large language model that has all necessary functions ranging from language understanding to decision making algorithms. The underlying large language models are trained on billions of words derived from books, articles, and other internet-based content. So large language model can predict the response to a given input from scratch without relying on handcrafted features or domain specific knowledge. The end-to-end system focuses on processing long windows of text that provide sufficient context to handle complex conversations and understand the nuances of human communications. For example, the very popular Chat GPT utilize an end-to-end approach.

Bob Barrett: So what are some examples of AI chatbots that have already been implemented in different healthcare settings?

Sarina Yang: AI chatbots have been used in an expanding range of healthcare domains. For example, the COVID-19 pandemic has accelerated the adoption and deployment of chatbots, which were designed to disseminate health information and knowledge about symptoms, precautionary measures, and medical treatment. The WHO Technology Program developed a chatbot to combat COVID-19 via WhatsApp and Facebook Messenger. In addition, an AI chatbot can answer common inquiries related to a particular disease or health condition, such as smoking cessation, hypertension, obesity, diabetes, or overcoming unhealthy habits. Furthermore, health chatbots such as Ada Health have been used as self-triage and personal risk assessment tools.

Moreover, chatbots offer a convenient and accessible way to provide remote patient support. For example, chatbots can be programmed to remind patients to take their medications, monitor vital signs, and track symptoms. These functions may be particularly beneficial for patients with chronic conditions such as cancer. By providing continuous support and monitoring, chatbots can improve patient outcome and quality of life while also reducing the burden on healthcare providers.

Bob Barrett: Well, finally, Dr. Yang, how might AI chatbots be used in laboratory medicine and what challenges are we likely to encounter when doing that?

Sarina Yang: In the field of laboratory medicine, chatbot holds exciting potential to improve medical education, provide timely responses to routine questions regarding laboratory tests, as well as facilitate the interpretation of laboratory results. For

example, many clinicians are perplexed by the complexity of laboratory tests and struggle to find detailed testing instructions. However, clinicians often lack the time to thoroughly search literature or request a laboratory consultation. As a result, clinicians may need an accurate, understandable, rapid, and interactive approach to obtain information about laboratory testing and interpret laboratory results. Chatbot that can provide curbside consultation could support clinicians who require immediate answer to those routine questions. Furthermore, many patients use online resources to search their laboratory test results and understand their medical conditions. Consequently, they are likely to seek chatbot's assistance in interpreting laboratory data or understanding how to utilize clinical laboratory services. In addition, chatbots can also greatly benefit laboratorians by saving their time spent answering routine questions and handling regular phone calls. However, the use of chatbots in clinical laboratories is still in its infancy stage and fraught with risks that must be considered. The current chatbots do not display adequate accuracy for clinical deployment and cannot reliably interpret complex laboratory data. There are several reasons for their limitations, which include a lack of domain knowledge training and expert annotation.

Furthermore, there is a lack of robust evaluation and validation of chatbots in laboratory medicine, making it difficult to determine their accuracy and reliability. Concerns have been brought up that chatbots may generate responses that appear convincing but contain inaccuracies or fabricated information. Moreover, large language models such as Chat GPT are not able to reliably interpret laboratory results in clinical context, since these often require specialized knowledge to understand.

Overall, at the current stage, users should be very vigilant of existing chatbot's limitations, such as misinformation, inconsistency, and a lack of human-like reasoning abilities. To be effectively used in laboratory medicine, chatbots must undergo extensive training on rigorously validated medical knowledge and be thoroughly evaluated against standard clinical practice. Well, I think laboratory medicine is a constantly progressing field with noble tests and techniques being developed and incorporated into clinical laboratories at an astonishing rate. As such, AI chatbots should be continuously updated to stay current with the latest knowledge in laboratory medicine. Considering these factors, if the limitation of AI chatbots can be addressed, they are poised to become a powerful tool for clinicians and laboratorians to embrace in their future clinical and laboratory practice.

Bob Barrett: That was Dr. Sarina Yang from New York Presbyterian Hospital/Weill Cornell Medicine in New York City. She and her colleagues published a mini review describing the use of AI chatbots in laboratory medicine in the November 2023 issue of *Clinical Chemistry*, and she's been our guest in this podcast on that topic. I'm Bob Barrett. Thanks for listening.