AI in Drug Discovery and Development
Prof. Rajnish Kumar
Dept. of Pharmaceutical Engineering and Technology
IIT-(BHU), Varanasi
Week-10
Lecture-47

Welcome to the course "AI in Drug Discovery and Development." In today's session, we will talk about AI in real-time monitoring and feedback. So by the end of this lecture, you will be able to explain the importance of real-time AI in critical care and therapeutic decisions. Distinguish between open-loop, closed-loop, and human-in-the-loop systems in clinical feedback; describe the core architecture and components of real-time AI systems. understand how AI enables real-time decision-making in healthcare, explore the ClinAIOps framework and its practical healthcare application, as well as identifying key challenges and safety concerns in deploying real-time AI systems. So, the question is, why do we need real-time monitoring? Because healthcare is one of the few domains where time sensitivity directly affects life-or-death outcomes, and the ability to monitor, analyze, and respond to patients' data in real time is not just a technological advantage; it is a clinical necessity.

So, the high-stakes nature of healthcare is derived from many critical decisions, such as adjusting a ventilator, administering epinephrine, or halting a drug trial, which depend on the rapid identification of clinical deterioration. Delayed recognition of symptoms or reactions can lead to severe complications or even fatalities as well. So, to avoid these complications, we can use real-time monitoring in healthcare management. So, the problems with traditional systems are that the manual checks are intermittent and reactive.

So, for example, checking vitals every 4 hours and then making clinical decisions often depends on retrospective data, leading to delays and a high dependency on human interventions that increase the risk of errors or delayed detection of complications. So, can we have a system which can monitor all these you know things in real time whether it is you know the vitals of a patient or it is the adverse drug events And then, based on that real-time monitoring of the events, the treatment can be altered, and that can be useful for conducting all those, especially those clinical trials, actually. So, the shift towards AI intelligent systems is, you know, with AI, where we can make the system dynamic, continuously learning from new data. We can make it responsive as well, where it adjusts the interventions based on live feedback, and proactive as well, where it can flag the risks before adverse events occur. So, AI doesn't just monitor; it learns and adapts to the emerging trends in the data as well.

For example, in sepsis management, every hour of delay in treatment increases mortality by almost eight percent, and AI-based real-time monitoring can predict onset four to six hours earlier using the vital signs and the lab markers. So the drivers of shifts, what are those factors that are, you know, inducing this shift? So these are the first things: the explosion of health data, including electronic health records, sensors, wearable devices, and applications. It could be the mobile apps or, you know, those health care wearable device apps. So, there has been a lot of technological development in this area, and that is making it possible to have a real-time monitoring system. And then there is a need for faster interventions in the ICUs, emergency rooms, and drug trials, as well as cost and resource efficiency by reducing hospital stays and manual efforts, as well as digital health transformation fueled by the cloud, 5G, and IoT.

So, all these technologies are making it possible to use real-time monitoring in clinical healthcare. So let us have a view of some of the, you know, key terms, like real-time monitoring, which is the continuous acquisition and analysis of data. This can be vital signs, biomarkers, behavioral patterns, or drug responses from multiple sources with minimal latency, often less than one second. and feedback loop, which is the automatic response mechanism that sends alerts, adjusts interventions, or modifies the treatment plan based on real-time insights. And the continuous therapeutic monitoring, which refers to the real-time and ongoing measurement of physiological or biochemical parameters that are critical for managing a patient's therapy.

as well as it enables dynamic adjustment of treatment based on live feedback ensuring personalized timely and precise medical interventions. So, latency is another key term that refers to the delay or time lag between the input or the event and the system's response or the output. In the context of real-time AI systems, it is defined as the time it takes for data to travel through the system, be processed by AI algorithms, and trigger a response or action. So, there can be multiple types of latencies, like data latency, which is the delay in capturing and transferring data from sensors to the processing units. It can be a processing latency, which is the time taken by an AI model or system to analyze the data and make a decision.

It can be the response latency, which is the time taken by the system to act on the decision, such as sending an alert or adjusting the doses. So, if you look at system architecture. So, at the bottom, it has the data collection layer, which pulls continuous data from the electronic health records, bedside monitors, wearable devices, applications, and lab systems. And it must support high-velocity, high-volume input because you can see that a lot of data is coming from all those, you know, sources, and then to handle that data we need a system that can support the high-velocity, high-volume input. And then the next layer is the data integration and pre-processing, where the real-time ETL processes unify

data from disparate formats and sources, standardize the timestamps, fill the missing values, and remove the noise.

And then the next layer is the analytical or ML layer, which hosts predictive models like early warning for sepsis and cardiotoxicity, and it must be optimized for speed, scalability, and interpretability. And the next layer is the decision layer, which applies the clinical rule or AI decisions, and it can trigger the alert, initiate a dose change, or pose a clinical trial arm. The next one is the feedback or user interaction layer, which consists of dashboards, SMS alerts, mobile notifications, or API responses to third-party tools. And then these are, you know, directly connected to either the user or, you know, the clinicians, actually. So, where they can see whatever decision is taken by the AI.

And then based on that, prioritization and explainability are critical in this case. And then on top of it, there is this learning loop that continuously improves model performance using the new outcome data. So, this is how the real-time system architecture is, you know, built. So if you look at the general workflow, as I said earlier, you have the AI engine, and then you have the data coming from different sources like imaging, genomics, electronic health records, or wearable devices; all of that data is fed into this AI engine. And then you have a clinical feedback loop which sometimes verifies whether the decisions taken by AI need to be checked by the clinicians or not, and then you have the safety alert system which starts, or which you know initiates, a safety alert on the basis of that a decision can be taken.

So, the real system designed principles for clinical validity; we have a look at it. So, there are multiple principles on the basis of that this real time monitoring system shall be you know designed. So, the first thing is scalability, as it should support the simultaneous monitoring of thousands of patients or devices. Especially in the clinical trial where you have thousands of participants, you should have a low latency threshold. It must deliver the inference results within seconds or even sub-seconds, actually.

And then it should be a failsafe design that falls back to clinician control in case of AI model uncertainty or failure, so that must be checked by the clinician, and that is how we can make it a failsafe design, actually. And then it should be explainable; actually, the outputs must be transparent, especially in high-risk decisions. And then it should comply with the regulatory bodies; it must align with HIPAA, GDPR, FDA, and other data and software safety regulations to ensure it can be used without any regulatory hurdles. So let us talk about some of the data sources for real-time monitoring. So the data can come from real-time clinical and biological signals.

So it is usually high-frequency data from sensors tracking, for example, the vitals such as

heart rate, blood pressure, SpO2, respiratory rate, and EEG, which are captured via monitors and wearables; these signals offer early insights into physiological distress such as arrhythmia or respiratory decline. And then you have the imaging, genomics, and laboratory data, so AI enables real-time interpretation of CT, MRI, and ultrasound for rapid diagnosis. And then genomic profiles, particularly pharmacogenomics, inform personalized dosing, and live lab feeds trigger alerts for critical shifts in biomarkers like troponin or creatinine. And then the data is coming from EHRs as well. So, EHRs are a major source of, you know, data for real-time monitoring.

So, the integrated EHR system provides structured data and unstructured nodes for contextual analysis as well. The real-time correlation with live data helps detect trends, adjust thresholds, and personalize alerts based on the patient's history. And then you have mobile and wearable technologies like smartwatches, CGMs, digital inhalers, and pill trackers, which support continuous at-home monitoring; these devices generate behavioral and physiological data used for dynamic dosing, adherence tracking, and lifestyle interventions. And then you have streaming data and clinical trials as well as participants in decentralized trials. They contribute real-time data via apps, sensors, and ePROs, and continuous input enables early detection of adverse events, adaptive randomization, and improved safety monitoring.

Okay then, coming to the next part, which is the sensors and IoT devices in real-time AI systems. So, sensors are usually the hardware components that detect changes in the environment. For example, it could be the temperature, pressure, motion, or biological signals, and then you have the IoT, which is a network of interconnected physical devices embedded with sensors, software, and communication technologies to collect and exchange data in real time. So, the role of, you know, sensors with the role of sensors in real-time AI monitoring systems. So, it acts as a data source layer; sensors are the frontline generating real-time data streams, and then they are, you know, capable of handling the high frequency and high volume.

These real-time AI monitoring systems need structured, clean, and timestamped data from the sensors, which enables automation and allows AI systems to react dynamically, for example, by stopping machinery, adjusting insulin levels, or alerting the clinicians. So, you can see here, for example, in this figure, which shows how sensors and IoT devices continuously collect data. So, you have all these, you know, the sensors and IoT devices. So, they collect the data that goes to, you know, the cloud infrastructure, and then it is fed into the machine learning models, especially the streaming models, and then that is used for making decisions. So, the AI models analyze the data in real time to generate actionable insights enabling automated feedbacks and decision making across various applications.

So, there are different kinds of sensors that are commonly used, such as those for measuring body and environmental temperature. It can be used for monitoring the patient's health and industrial safety. Then there are, you know, motion sensors that measure the acceleration and orientation. Sometimes they are used for, you know, fault detection, and then there are, you know, pressure sensors which are measuring the barometric fluid and mechanical, you know, properties. So, these are used mainly in the ventilators; for example, then you have the optical sensors, which are used for measuring the light and also for capturing   the   images.

So, these are used generally for diagnostics, like, you know, the SpO2 sensor. And then you have biometric sensors that measure heart rate, SpO2, and ECG, mainly used in medical variables. And then you have the chemical or biosensors that are measuring the pH, glucose, or toxin levels in the blood or, you know, the body fluid. So, these are mainly used for glucose monitoring, for example. So, there are some challenges with using those sensors,   and   some   of   those   challenges   are   sensor   accuracy   and   drift.

So, the AI predictions are only as good as the incoming data because the sensors are not getting the correct data. So, the data coming from the sensors is not reliable. So, then the you know the AI system will not be able to give the correct outcome. And then power efficiency is important because these are especially wearable sensors or remote sensors; they should be power efficient because recharging them repeatedly and replacing their batteries can create significant trouble. And then security and privacy, protecting live data from   being   intercepted   or   misused,   is   also   one   of   the   concerns,   actually.

Then we come to the types of feedback loops. There are basically two kinds of loops: one is the open loop system, and the other is the closed loop system. So in the open loop setup the AI system monitor data and provide recommendations without executing direct actions and the decision remains with the human user actually. So this model supports clinical autonomy and is often preferred in high-risk settings like the ICU and drug trials. And then you have a closed-loop system that enables autonomous feedback, where AI is not only detecting issues but also directly initiating actions such as adjusting a ventilator setting or insulin   dose.

So, these systems they reduce response time, but they require high reliability and safety validation and they cannot be used in emergency situations actually. So, they can be used for routine, you know, tasks. And then you have the human-in-the-loop feedback, which integrates human oversight into the real-time system. So, HITL systems are designed to keep a clinician, researcher or operator involved in the decision process. While AI may detect a risk, prioritize an alert, or recommend an intervention, the human-in-the-loop validates,   modifies,   and   approves   the   action.

So, this approach is particularly essential in domains where safety, ethics, and contextual understanding cannot be fully automated, such as emergency medicine, diagnostic, or patient-specific dosing decisions. You can see here, like this is the, you know, for example, patient, which is getting where the data is coming from, and then AI is making a decision. And then a human is reviewing the decision, and then an action is taken, verified by the human or the clinician, and then that is being applied to the patient. So, it is also helpful in balancing speed and safety. So, where AI enables high speed analysis, but clinical decisions often require the nuance.

So, HITL ah models allow real time systems to operate quickly without sacrificing the safety. For instance, in a closed-loop insulin delivery system, the AI may suggest a dose adjustment, but in HITL configuration, a clinician reviews the recommendation before it is applied to the patient. So, this not only mitigate a risk, but also builds confidence in the systems reliability. And then HIDL acts as a learning mechanism as well. So, human-in-the-loop feedback is not static, as clinicians interact with AI systems by approving, correcting, or rejecting suggestions.

So, they provide valuable input for model refinement. So, this adaptive feedback allows AI to learn from expert decisions, improve accuracy, and reduce reliance on manual validation over time. So, coming to the core AI technologies that are being used in real-time monitoring systems. The first one is NLP, or natural language processing, because it is helpful in the real-time extraction of clinical data from patient reports, outcomes, or EHR nodes. For example, it converts patient-reported outcomes, either text or audio, into structured feedback.

And then the time series models, like LSTM or GRU and transformers, are used for modeling patient vitals and trends. And then anomaly detection techniques like isolation forests, autoencoders, and one-class SVMs are used to detect deviations from normal patterns, such as unusual ECG waveforms. Reinforcement learning involves adaptive decision-making in ventilator settings or drug dosing. Classification models that can be used to detect risk classes such as high versus low bleeding risk. And the multi-modal AI fusion can also be used, which combines input from video, text, and structured data for holistic insights.

So these are some of the real-time deployment strategies, like edge AI, which is a kind of lightweight model running on local devices with minimal latency, and an example of that could be smart insulin pens or wearable ECG analyzers. And then you have the cloud-based AI, which is centralized computation on large data sets ideal for large hospitals. An example could be the predictive analytics engines in the ICUs. And then you have the

hybrid systems that combine pre-processing with cloud-based inference and storage. So this could be like on-premises pre-processing and cloud AI analytics for scalability as well as privacy.

So an example of that kind of system could be the mobile health apps with real-time alerts synced to the hospital dashboards. So then we have this ClinAI Ops, which is a framework for AI-integrated clinical feedback. So clinical AI operations are a new framework that links patients, clinicians, and AI developers through continuous therapeutic monitoring and feedback loops, transforming how AI operates within real-time clinical workflows. So, traditional care models rely on delayed clinical visits to adjust therapies, and ClinAI Ops proposes tighter real-time loops powered by variable devices, AI analytics, and cloud-connected systems. So, it enables therapy adjustment based on live patient signals without always needing in-person visits.

So, the patients, clinicians, and AI developers are continuously connected, creating an ecosystem of adaptive care. So, let us have a look at the feedback loops of ClinnAI Ops. So these ClinAI Ops connect stakeholders: the patients, clinicians, and AI developers via three feedback loops. The patient AI loop, which is where AI interprets real-time health signals like glucose levels and heart rhythms, suggests therapeutic changes. And the patient may self-adjust doses within a clinically approved range.

And then you have the patient-clinician loop, where the clinician receives AI-generated summaries and focuses on personalized coaching, non-monitored symptoms, and goals. So it enables dynamic scheduling of widgets based on need and routine, and then you have the clinician AI developer loop. Where the clinician feedback on AI suggestions helps improve model accuracy, safety, and utility. So, it can be like a rejected or modified recommendation that provides real-world ground truth for the model retraining. So, this is one of the applications of ClinAI Ops for the management of blood pressure.

So, in the patient AI loop, the patient wears a passive continuous blood pressure monitor and reports antihypertensive administration. So, the AI generates recommendation for the antihypertensive dose titration. And then in the clinician AI look, the clinician sets and updates the AI's limits for the titration of the antihypertensive dose. And the AI alerts of severe hypertension or hypotension, promoting follow-up or emergency medical services. And then you have the patient-clinician group where the patient discusses the AI-generated summary of their blood pressure trend and the effectiveness of the therapy, and the clinician checks for adverse events and identifies patient-specific modifiers such as diet or exercise.

So, this is the timeline for the application of clinAI ops to CTM. So, the patient, for example, shares challenges and health goals, and the clinician prescribes an AI-powered

sensor with an initial therapy plan, and the AI developers supply a low-power sensor with their own device AI for inference. And then the patient modifies therapy within the clinician's bounds, okay, because the clinician has suggested that within this range, you can modify the therapy and change the dose, but only within that range. And then it modifies therapy after the clinician's approval, and then the clinician is alerted to an out-of-bounds update to the therapy. So, the clinician approves the recommendation, and then the developer updates the algorithm based on the alert and the updates to the therapy and reinforces the algorithm's recommendation using the clinician's decision.

And then finally, it is being reviewed in the patient monitoring summary to evaluate the progress towards the goal, and then if the clinician can reject out-of-bounds suggestions and call in the patient for a visit. And then the developer reinforces against the algorithm's recommendation using the clinician's decision. And then finally, the clinician reviews the patient monitoring summary to suggest an updated resume. And then the developer updated the monitoring algorithm on the basis of the patient outcomes. So, this is how you know they work in tandem with each other, leading to a common goal.

So, this is, you know, another real-world application of ClinAI Ops. Okay, and then let's talk about the stakeholders' perspective in the operation of the Clin AIOps feedback ecosystem. So if we talk about the clinician's perspective, the focus shifts from repetitive data entry to interpreting AI summaries and providing higher-order care. And then clinical time is spent on setting goals, reviewing exceptions, and making strategic decisions. And it also provides feedback to the developers, which refines AI tools to better support workflows and reduce false positives.

So, if you look at the patient's perspective, the patient engages in self-care with AI-assisted guidance like dose titration or behavior nudging. So, it has a reduced dependence on routine clinical visits, more flexible and personalized care, and it encourages autonomy and adherence while still being clinically supported. And then, looking at the AI developer's perspective, it has access to real-world feedback, which improves model generalizability and safety. It gets direct insights from clinician overrides that help reduce hallucination and error propagation, and it also enables the dynamic model updates that remain clinically grounded. So bringing ClinAI ops into clinical practice so we can use, you know, the HR and device integration where the AI system must integrate with existing clinical systems like EIC, Cerner, and device ecosystems.

Standardized APIs, HL7/FHIR protocols, and cloud-based data platforms enable real-time data exchange, allowing adaptive workflows where clinicians are alerted based on exceptions rather than volume. And then feedback loops adapt to patient needs and clinical preferences; some decisions are automated, while others are being reviewed, ensuring

scalability and continuous learning. where the feedback loops are designed to scale across patient populations and adaptive learning ensures models improve over time without needing full redevelopment And the real-world readiness, where the early deployment should focus on conditions with clear measurable parameters like glucose or BP, and co-design with clinicians ensures adoption and reduces resistance. Okay, in the end, coming to the challenges and safety considerations, there are multiple challenges and safety considerations as well. So, the safety risks and mitigations, such as real-time AI decisions, may lead to patient harm if the model misfires, and that risk is mitigated through constraints like dose limits, alert thresholds, HITL checkpoints, and continuous validation.

And then there is feedback fatigue and alert overload. So the overly frequent feedback can overwhelm the users, and the ClinAI OPS encourages goal-driven feedback, where updates are sent only when deviations from the patient-defined targets are detected. So, another challenge is the data quality and integration, where sensor noise, missing values, and misaligned timestamps can impair accuracy. So, as a solution, you know that the system must include a pre-processing layer, error detection, and the fallback logic. And then there are challenges with the regulatory and ethical frameworks, like the dynamic model challenge challenging static validation paradigms, so the ongoing collaborations with the regulators are essential to define safe and updated AI practices and maintain transparency.

Okay, coming to the summary. So, these AI enabled real time monitoring allows continuous assessment of patient data for timely and personalized care. So, techniques such as time series analysis, anomaly detection and reinforcement learning support proactive clinical decision making and the frameworks like Clin AI Ops which enables dynamic feedback loops between patients, clinicians and AI systems. These systems enhance responsiveness, safety, and precision in therapeutic interventions. Despite deployment challenges, real-time AI marks a significant shift toward adaptive and intelligent healthcare delivery. So, now we have seen how we can use those real-time monitoring systems, especially in clinical trials, and how they can help.

So, in the end, I have an open question for you. So, if AI can continuously monitor and adjust treatments in real time, should clinical decisions ever be fully automated, or should humans always remain in the loop? So, just ponder over it, and I have suggested some publications that you can go through if you want to learn more about this topic. And with that, thank you.