Deep Learning Advanced: Agentic Generative AI for Precision Medicine and Health

Date: March 23, 2026

Module Objectives

- Understand the current bottlenecks in modern healthcare and how large-scale biomedical data combined with AI can accelerate precision medicine.
- Describe the use of human-collaborative multi-agent systems for privacy-preserving drug screening.
- Explain the application of generative AI models (diffusion, RL, VAEs) in biomedicine, including for TCR engineering, protein complex design and molecule optimization.
- Understand the future trends and challenges in precision medicine and health, including digital twins and AI-driven robotic surgery.
- Frame why medical decision-making is a natural “agentic” setting involving many steps, many sources and high cost of errors.
- Introduce representative agentic and generative method families: RAG, tool-using LLM agents, RLHF, diffusion/latent-variable models.
- Situate these systems within health data standards, regulation and evaluation (FHIR, FDA CDS/GMLP, CONSORT-AI/SPIRIT-AI).
- Understand how precision medicine problems are fundamentally workflow problems requiring chaining of many imperfect steps.

Key Concepts & Definitions

- Precision Medicine: An innovative medical approach that analyzes individual variability in genes, environment and lifestyle to tailor medical care, rather than using a one-size-fits-all strategy.
- De Novo Drug Design: A computational approach to drug discovery that generates entirely new, distinct molecular structures from scratch. On average, designing one new drug takes over 10 years and costs more than $2 billion, with a success rate under 10%.
- T-Cell Receptors (TCRs): Heterodimeric protein complexes found on the surface of T-cells responsible for recognizing and binding to specific antigens, thereby triggering an adaptive immune response.
- TCR Engineering: The process of computationally optimizing TCR sequences to enhance their binding strength and specificity for target antigens, as demonstrated for cancer immunotherapy applications.
- Reinforcement Learning (RL): A machine learning paradigm where an AI agent learns to make a sequence of decisions by interacting with an environment to maximize a cumulative reward.
- Diffusion Models: Generative models that diffuse training data with random noise, then learn to reverse the diffusion process to output new, high-quality data samples.
- Human-Centric Multi-Agents: An AI architecture consisting of multiple specialized LLM agents that work collaboratively with human experts, enabling interactive, privacy-preserving insight discovery.

- Digital Cell / Digital Patient: Highly detailed, data-driven virtual simulations of biological entities, combining predictive AI, generative AI and physical engines (cell world models).
- Agentic System: A system that iteratively plans → acts (tool calls) → observes → updates plan to accomplish a goal, rather than emitting a single-pass answer.
- RAG (Retrieval-Augmented Generation): A method that retrieves relevant documents from an external corpus and conditions generation on those retrieved snippets to improve factuality and provenance.
- RLHF: Reinforcement learning from human feedback, where preference judgments train a reward model, then RL optimizes the policy to better match desired behavior.
- PPO (Proximal Policy Optimization): A widely used RL optimizer that stabilizes learning via a clipped surrogate objective, commonly applied in RLHF and sequence optimization.
- Foundation Model: A large pre-trained model (often transformer-based) that can be adapted to many tasks via prompting, fine-tuning or tool-augmented inference.
- VAE (Variational Autoencoder): A latent-variable generative model trained with an evidence lower bound (ELBO); widely used for representation learning and integration in high-dimensional biomedical settings.
- DP-SGD: A differential privacy training variant that clips per-example gradients and adds noise, reducing memorization risk at some cost in utility.
- Tool Calling: A pattern where the model triggers external functions (search, calculators, simulators, lab automation) and integrates outputs back into generation, improving factuality and controllability.
- Domain Feedback / Gatekeeper: A structured evaluation layer (rules, predictors, constraint checks, human review) that filters or scores candidate outputs, reducing hallucinations and enforcing scientific/clinical constraints.
- Clinical Decision Support (CDS): Software that provides recommendations or information to clinicians or patients; some CDS functions fall under FDA medical device oversight.
- FHIR: A standard for exchanging electronic healthcare information using 'resources' and profiles, supporting interoperability needed for scalable CDS and analytics.
- LLM Itself is Not Enough: When prompted to generate a cartoon antigen-antibody complex image, a general-purpose LLM produces a cartoon drawing rather than a scientifically accurate 3D structure, illustrating that domain-specific tools and agents are essential in biomedicine.

Main Content / Topics

1. Precision Medicine and the Case for Transformation

Modern healthcare faces three structural global pressures that demand transformation:

- A growing elderly population driven by increasing lifespans will pressure existing medical infrastructure.
- Widespread low-cost world travel facilitates rapid spread of disease, increasing the risk of frequent global pandemics.
- A medical knowledge explosion requires an ever-increasing number of experts to manage and deliver treatments.

The traditional de novo drug discovery pipeline is extremely slow and expensive. On average, it takes over 10 years and ~$2 billion to bring one new drug to market, spanning target validation, compound screening, lead optimization, pre-clinical testing, Phase I–III clinical trials and FDA approval. The overall success rate remains under 10%, with more than 10,000 initial candidates narrowing to roughly 2 by Phase III.

To overcome these bottlenecks, precision medicine leverages large-scale biomedical data, such as personal genomics, physiological data and medical records, combined with artificial intelligence. AI agents are increasingly used to accelerate every phase of drug development. A 2025 FDA announcement to phase out animal testing requirements for monoclonal antibodies and other drugs further signals a shift toward AI-driven, digital pre-clinical and clinical workflows.

Precision medicine pushes clinical decision-making toward high-dimensional personalization (genomic + clinical + environment) and evidence dynamism (guidelines and drug labels evolve). Traditional “single-model” automation struggles because medicine is not one prediction, it’s a workflow. You gather patient facts, consult evidence, compute risk scores, check contraindications, reconcile uncertainty and document reasoning.

2. Human-Collaborative Multi-Agents for Drug Screening

NEC developed a human-collaborative multi-agent system for drug screening that addresses two key technical problems:

- Reading thousands of literature papers is time-consuming.
- Sending private company data to public LLMs poses data-leaking and privacy risks.

The NEC system uses three specialized agents:

- Agent 1: A public Guidance LLM that processes public documents and generates expert guidance.
- Agent 2: A local retrieval agent that searches private documents and extracts relevant reference chunks without exposing data externally.
- Agent 3: A local QA LLM that answers questions using the retrieved private data, keeping sensitive information secure.

The system provides interactive visualization of large-scale compound libraries, allowing researchers to filter candidates, ask natural language questions (biodegradability, flexibility, molecular weight), and receive answers with confidence scores. Visual encodings map multiple properties simultaneously. For example, fill color for biodegradability, stroke for curdlan-based composition and position axes for molecular weight and density.

This architecture exemplifies the principle that 'model capability' and 'data access' should be decoupled in regulated environments. The best planner is not necessarily allowed to see protected health information. The strong public model designs the workflow; local smaller models execute it near sensitive data.

The lecture also emphasized the broader vision of transitioning from standalone LLMs to Structured Workflow AI: domain expertise is captured into reusable, auditable workflows, letting AI scale beyond individual query-response interactions into complex multi-step pipelines.

3. LLM-Guided Drug Design with Gate Keepers

A key example of agentic drug design presented in the lecture is ChatDrug. The system addresses drug editing for small molecules, peptides and proteins via a conversational LLM interface augmented by a Retrieval and Domain Feedback (ReDF) module.

The workflow proceeds iteratively across multiple conversation rounds:

- The user inputs a drug molecule (SMILES) and a task prompt.
- The LLM proposes an edit (e.g., adding a tertiary amine group to improve water solubility).
- The ReDF module retrieves relevant chemical literature from a database and applies domain feedback to evaluate the proposed edit.
- If the edit is incorrect, the system flags it and the LLM proposes a revised candidate in the next round.

ChatDrug achieves best performance across 39 drug-editing tasks. The key design principle is that the LLM serves as a reasoning and interface layer, while correctness is enforced by domain gatekeepers, a canonical demonstration that LLM alone is insufficient for biomedical applications.

4. Deep Reinforcement Learning for TCR Engineering (TCRPPO)

T-cell receptors recognize specific peptide-MHC complexes on cell surfaces during viral infections or tumor development. TCRs consist of alpha and beta chains; the complementarity-determining region (CDR) is the key region responsible for antigen recognition. Major challenges include poor generalizability to rare or novel antigens and a very large sequence search space.

Policy learning follows the PPO (Proximal Policy Optimization) framework with a clipped surrogate objective, a value network for advantage estimation via Generalized Advantage Estimation (GAE), and an entropy bonus to encourage exploration. A complementary approach uses a disentangled Wasserstein autoencoder to separate TCR representations into a sequential embedding (structural characteristics) and a functional embedding (peptide-binding behavior), which enables targeted modifications.

Results: From over 6.4 billion mutated TCR candidates, two optimized TCRs targeting the melanoma antigen MART-1 were selected for wet-lab validation. One engineered TCR demonstrated over 1,000 times greater sensitivity than the original template. Incorporating functional cellular data from response assays (DCP component) further improved results.

TCRPPO is a canonical example of 'agentic generation' where the generator is a policy, not merely a sampler: the system chooses where and how to mutate based on learned advantage

estimates. The PPO clipping objective is doing critical systems work here, it reduces instability and reward hacking risk when the reward is partly model-based and thus imperfect.

5. Diffusion-Based Generative Models for Molecular and Protein Design

PPDiff: Addresses protein-protein complex design by diffusing in a hybrid sequence-structure space. Given a target protein's amino acid sequence and alpha-carbon backbone structure, PPDiff co-designs the binder protein's sequence and structure using a hybrid discrete-continuous diffusion process. It was pretrained on PPBench (734,032 complexes) and fine-tuned for mini-binder and antigen-antibody complex design. PPDiff achieved the highest success rate and novelty compared to baselines.

MolDiffAE: A disentangled autoencoding equivariant diffusion model for controlled 3D molecule generation. It encodes molecules into separate semantic and structural embeddings, enabling controlled manipulation of properties like QED, LogP, SAS and halogen content. The model supports both unconditional generation and source-conditioned property manipulation, so it achieves efficient multi-objective optimization.

Modof: Performs 2D fragment-based molecule optimization. It modifies molecules by predicting a single site of disconnection, then removing or adding fragments at that site. A pipeline of multiple Modof models (Modof-pipe) enables iterative optimization, progressively improving drug-likeness (QED) and target activity (DRD2) scores.

Controllable Fragment-based 3D Molecule Generation: Learns disentangled equivariant representations separating property and structure latent spaces using SE(3)-equivariant graph neural networks, leading to explicitly controllable 3D fragment-based molecule generation.

Diffusion models generate by reversing noise. Training often reduces to predicting the added noise ε at random time t. In medicine, surveys and systematic reviews report rapid growth of diffusion applications (generation, translation, reconstruction, denoising) while emphasizing significant remaining barriers to reliable clinical deployment.

PPDiff illustrates 'agentic generation' in a different sense than tool calls: the key is a structured generative prior (diffusion + equivariance) that embeds physics/geometry constraints into the sampling process. The work on MolDiffAE and the AAAI 2025 paper both show the power of disentangled latent spaces. Separating what a molecule does functionally from what it looks like structurally allows much more precise, controllable design.

6. Key Mathematical Frameworks

VAE ELBO: The VAE objective maximizes the Evidence Lower Bound:

$$\log p_\theta(x) \geq \mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x \mid z)] - \mathrm{KL}\left(q_\phi(z \mid x) \parallel p(z)\right).$$

DDPM Denoising Loss: Diffusion models are trained to predict the noise ε added at timestep t. This is equivalent to multiscale denoising score matching.

$$\mathcal{L}_{\text{diff}}(\theta) = \mathbb{E}_{t,x_0,\varepsilon}\left[\|\varepsilon - \varepsilon_\theta(x_t, t)\|^2\right].$$

PPO Clipped Objective: Clipping prevents overly large policy updates and is critical for stable TCR optimization.

$$\mathcal{L}^{\text{PPO}}(\theta) = \mathbb{E}\left[\min\left(\rho_t(\theta)\hat{A}_t,\ \text{clip}(\rho_t(\theta), 1-\epsilon, 1+\epsilon)\hat{A}_t\right)\right], \quad \rho_t(\theta) = \frac{\pi_\theta(a_t \mid s_t)}{\pi_{\theta_{\text{old}}}(a_t \mid s_t)}.$$

RAG Marginalization: RAG treats retrieved documents as latent variables, leading to grounded and updateable generation without retraining the generator.

$$p_\theta(y \mid x) = \sum_{z \in \mathcal{N}_k(x)} p_\eta(z \mid x)\, p_\theta(y \mid x, z)$$

TCRPPO Final Reward:

$$\mathcal{R}(c_T, p) = p_{\text{r}}(c_T, p) + \alpha \min(0, s_{\text{tcr}}(c_T) - \sigma_c)$$

Here, p_r is the predicted binding probability from ERGO/AVIB, s_tcr is the validity score (reconstruction + density), and σ_c is the validity threshold. The penalty term checks that only valid TCR sequences are rewarded for high binding affinity.

7. Digital Cell, Digital Patient and Future Directions

The concept of a generative digital cell combines predictive AI, generative AI and physical engines (cell world models) to simulate cellular behavior. This draws on foundational work in computational biology such as identifying regulatory modules from gene expression data and predicting gene expression from sequence. Both pharmaceutical and IT companies are increasing investment in digital twins of cells and patient models.

The healthcare digital twins market is projected to grow from approximately $903 million in 2024 to over $9 billion by 2034, representing a CAGR of about 25.9%. Other future trends include:

- Connected health (market growing at 20.3% CAGR through 2030), where providers will use AI to summarize records, triage patient inquiries and augment doctor-patient interactions.
- AI-driven robotic surgery (market valued at $6.4 billion in 2022, expanding at 18.9% CAGR through 2030).

The lecture outlined a hierarchy of validation challenges for AI-designed therapeutics, progressing from binding affinity (molecular level) → cellular response → clinical response (efficacy and safety) → sub-population personalized response. Key open challenges include

fueling scientific discovery with accumulated data, increasing availability and quality of biomedical data while protecting privacy, and efficiently validating AI-generated candidates through AI agents operating self-driving labs.

The lecture's 'digital cell / patient digital twin' vision should be interpreted as an aspirational research program rather than a solved capability. Reviews in systems biology show that post-transcriptional regulation and protein degradation substantially shape protein abundance, and this limits approaches that treat transcriptomics as the whole story.

A compelling framing from the lecture: the ultimate vision is summarized as 'Of the Human, By the Human, For the Human', suggesting that all AI agents and digital systems should remain fundamentally oriented toward human benefit.

Discussion / Comments

Generative AI and multi-agent systems are driving a major shift in precision medicine, addressing the high costs and slow pace of traditional drug discovery. This lecture demonstrates how deep reinforcement learning and equivariant diffusion models are revolutionizing T-cell receptor engineering and 3D molecule generation. Meanwhile, human-centric multi-agent systems provide a secure bridge between human expertise and massive biomedical datasets.

A central lecture insight is that health AI's bottleneck is rarely model capacity alone; it is the end-to-end socio-technical pipeline: data standardization, privacy, evaluation and the ability to incorporate domain tools and validate in the real world. The literature cautions that clinical evidence for LLM systems is still maturing, with many evaluations being retrospective or simulated, an institutional barrier motivating the emphasis on gatekeepers, human oversight and rigorous reporting standards.

A key open question that emerges from this lecture is: as agentic AI systems become more capable, where exactly should the human remain in the loop, and what does 'human oversight' actually mean when a system can act faster and access more information than any individual clinician? This connects to broader ethical questions about accountability, liability and the risk of automation bias in high-stakes medical decisions. Additionally, the lecture's roadmap from binding affinity → cellular response → clinical response → sub-population personalized response highlights that the technical challenges are compounded at each level of biological complexity. This suggests that computational and experimental approaches must remain tightly coupled rather than treating in silico validation as a substitute for wet-lab work.

Suggested Readings

1. Bran et al. ChemCrow: augmenting large-language models with chemistry tools. Preprint at https://arxiv.org/abs/2304.05376 (2023).
   This paper shows LLMs acting as autonomous agents using external chemical tools to plan syntheses and help human scientists.
2. Klebanoff et al. T cell receptor therapeutics: immunological targeting of the intracellular cancer proteome. Nat Rev Drug Discov. 2023 Dec;22(12):996-1017. doi:

10.1038/s41573-023-00809-z.
This provides some biological background to understand the de novo design of T-Cell Receptors and why increasing their affinity (the goal of the TCRPPO model in the lecture) is so important for targeted cancer treatments.

3. Shubham Vatsal et al. Agentic AI in Healthcare and Medicine: A Seven-Dimensional Taxonomy for Empirical Evaluation of LLM-Based Agents. Preprint at https://arxiv.org/abs/2602.04813 (2026).
This review paper categorizes exactly how multi-agent architectures are being used for clinical tasks, treatment planning and research workflows, to bridge the lecture's AI concepts with its healthcare application goals.