Uncovering the Potential Impact of Large Language Models and Generative AI on Clinical Trial Efficiency to Excel in Drug Development

Ed Addison

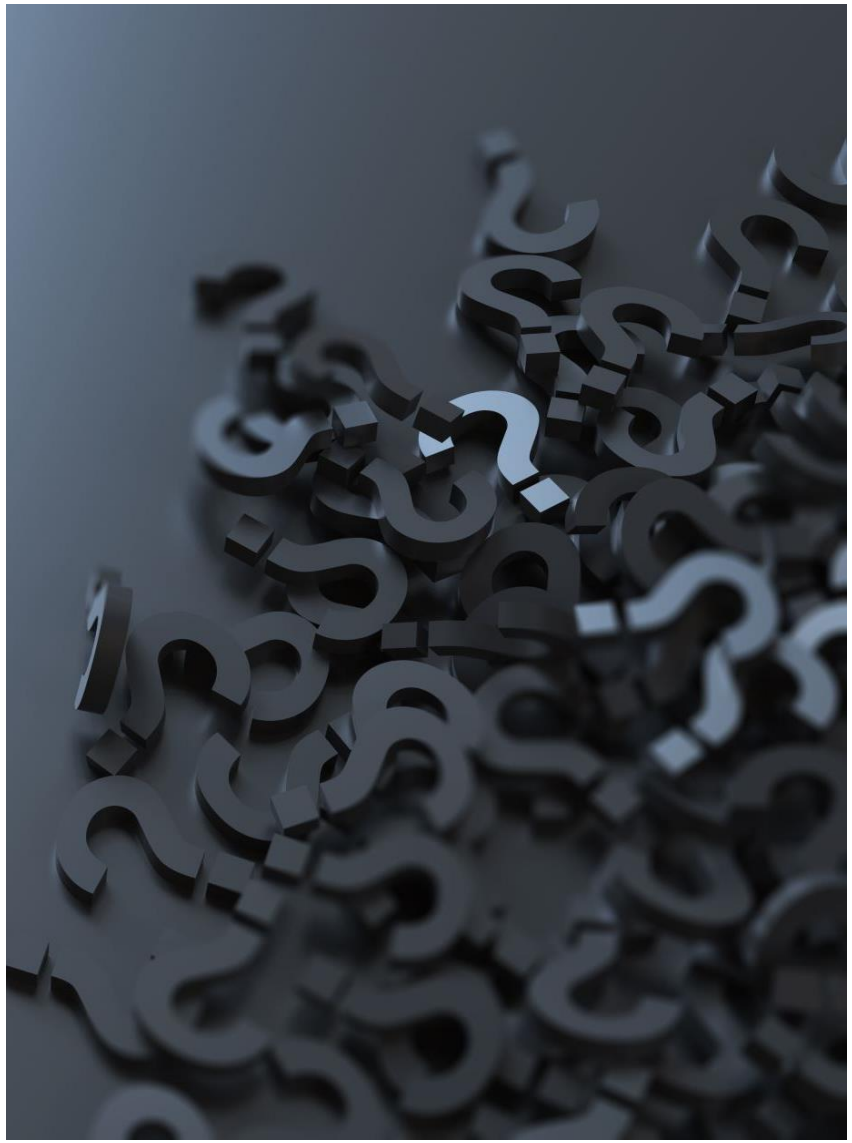
Chairman and CEO, Cloud Pharmaceuticals

Pharmaceutical Generative AI Market

- “GenAI will create $60-110 billion in value in the pharmaceutical industry” (1)
- “An accelerated drug discovery process ...will help cure more diseases quickly... ” (1)
- “The ability to generate insights from vast quantities of patient data will spark personalized treatments—and improved outcomes. ” (1)
- “Gen AI can create value across the entire business value chain” (1)

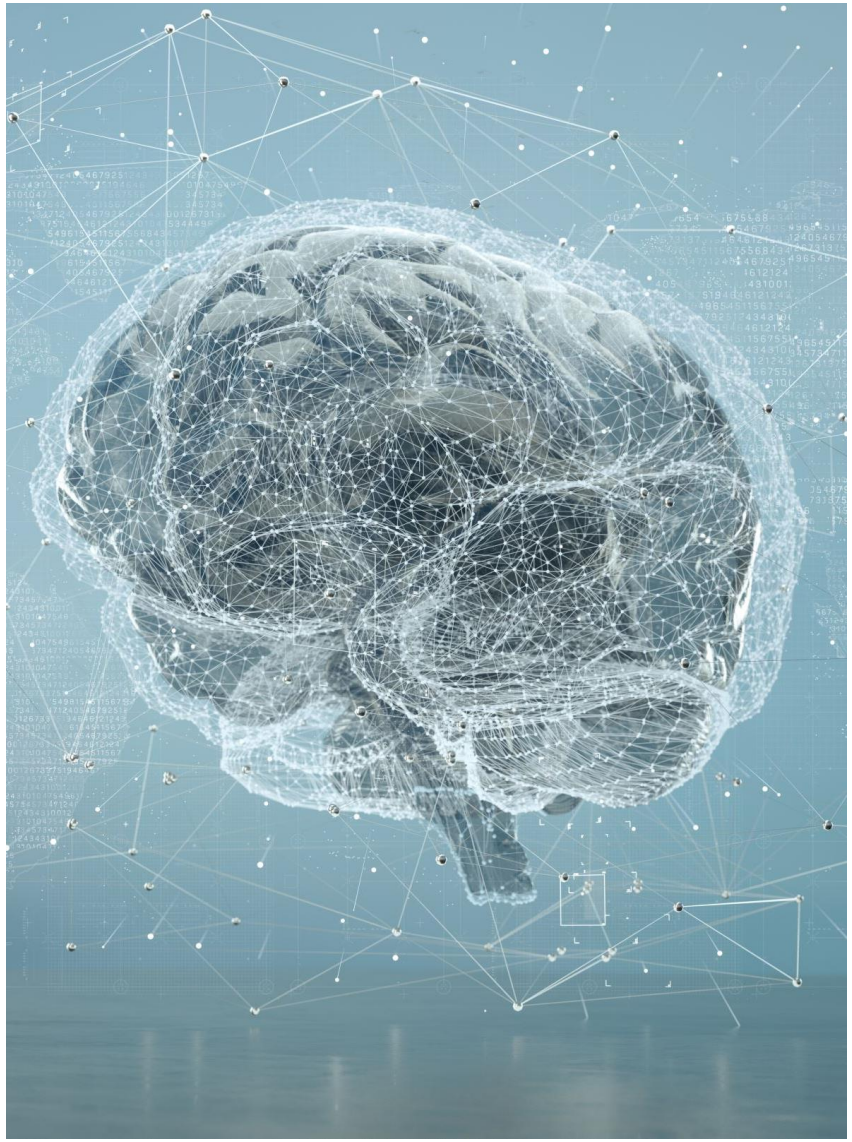
---

(1) McKinsey, January 12, 2024



The Pharmaceutical Industry AI Problem

- It takes 10 years and $2.6B to develop a drug
- Precision Medicine is still elusive
- Artificial Intelligence for drug discovery has deen proven for point solutions only
- Generative AI promises to solve this, but costs >$B to train models, usually on uncurated and dirty data



The Generative AI Solution

- Generative AI models cover all phases of drug discovery and development, and will speed drug development, reduce cost, provide better drugs, and generate more revenue due to longer patent life
- Models are trained ”orders of magnitude” faster using a fast data center, unique “convergence algorithms” and by leveraging existing GenAI models

Enormous
Untapped
Applications!

Here are 10 of them....

Generative AI for Clinical Trials

1 Protocol Development Assistant

LLMs can assist in drafting and refining clinical trial protocols by generating text based on existing protocols, scientific literature, and regulatory guidelines. They can help ensure that protocols are comprehensive, clear, and consistent with relevant standards and regulation

2 Clinical Trial Design Optimization

- LLMs can analyze vast amounts of textual data, including scientific literature, patient records, and clinical trial protocols, to identify relevant information and patterns.
- This analysis can help researchers optimize the design of clinical trials by identifying suitable inclusion and exclusion criteria, selecting appropriate endpoints, and designing patient recruitment strategies.

3 Patient Recruiting and Retention

- LLMs can analyze patient data, electronic health records (EHRs), social media, and online forums to identify potential candidates for clinical trials and predict patient eligibility and willingness to participate.
- They can also generate personalized communications to engage and retain participants throughout the trial.

4 Regulatory Compliance and Documentation

- LLMs can assist in automating regulatory compliance tasks, such as preparing regulatory submissions, compiling safety reports, and generating documentation for Institutional Review Board (IRB) approvals.
- By generating structured text based on predefined templates and guidelines, LLMs can streamline the regulatory review process and ensure compliance with relevant regulations.

5 Real World Evidence Generation

- LLMs can analyze real-world data sources, such as patient registries, claims databases, and wearable devices, to generate real-world evidence (RWE) for supporting regulatory submissions, label expansions, and post-market surveillance.
- By integrating structured and unstructured data, LLMs can provide insights into treatment effectiveness, safety outcomes, and patient preferences in diverse patient populations.

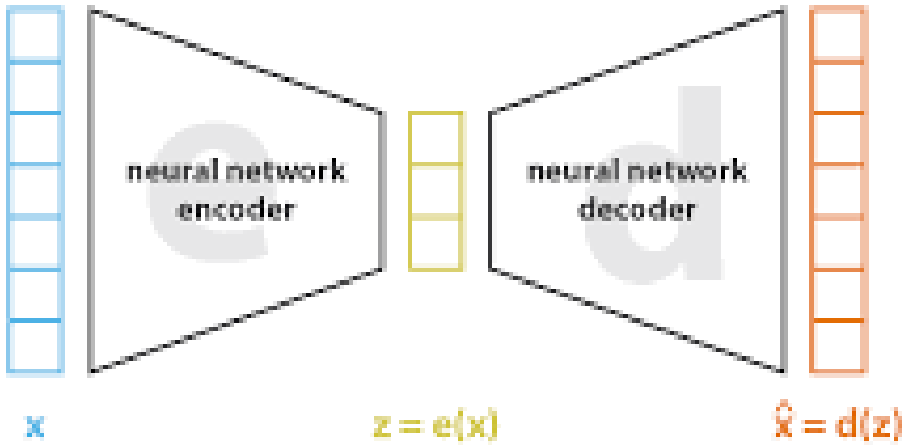
6 Drug Repurposing and Biomarker Design

- LLMs can analyze scientific literature, clinical trial data, and molecular databases to identify potential drug candidates for repurposing and biomarkers for patient stratification.
- By extracting and synthesizing information from diverse sources, LLMs can accelerate the identification of novel therapeutic targets and treatment strategies.

7 Analysis of Clinical Trial Data

- Generating synthetic data to augment limited datasets, improving model performance.
- Imputing missing values in the dataset, enhancing data completeness for analysis.
- Predicting clinical outcomes based on patient characteristics and treatment interventions, guiding decision-making in trial design and patient management.

8 Variational Autoencoder for Data Analysis

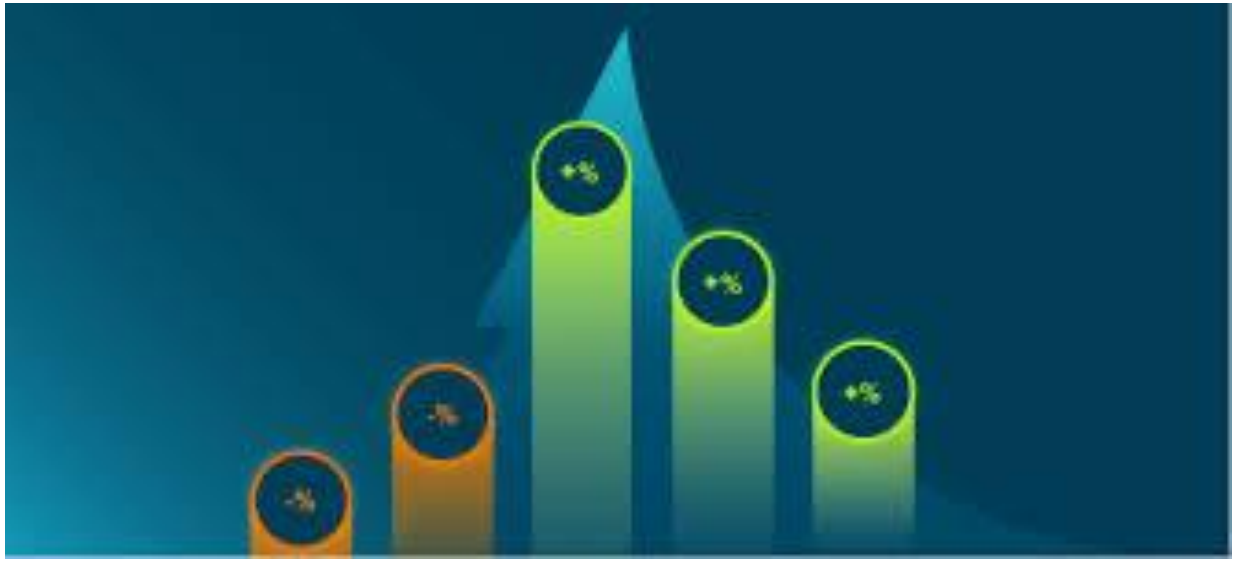


$$\text{loss} = \|x - \hat{x}\|^2 = \|x - d(z)\|^2 = \|x - d(e(x))\|^2$$

- VAEs can learn low-dimensional representations of complex medical data, such as patient electronic health records (EHRs) or genomic sequences.
- These learned representations can facilitate data analysis, visualization, and patient stratification in clinical trials.

9 Deep Reinforcement Learning Can Optimize Trial Performance

- DRL techniques can optimize patient treatment strategies
- DRL can optimize trial protocols by learning adaptive trial designs
- DRL can also optimize patient enrollment



10 Clinical Trial Outcome Prediction

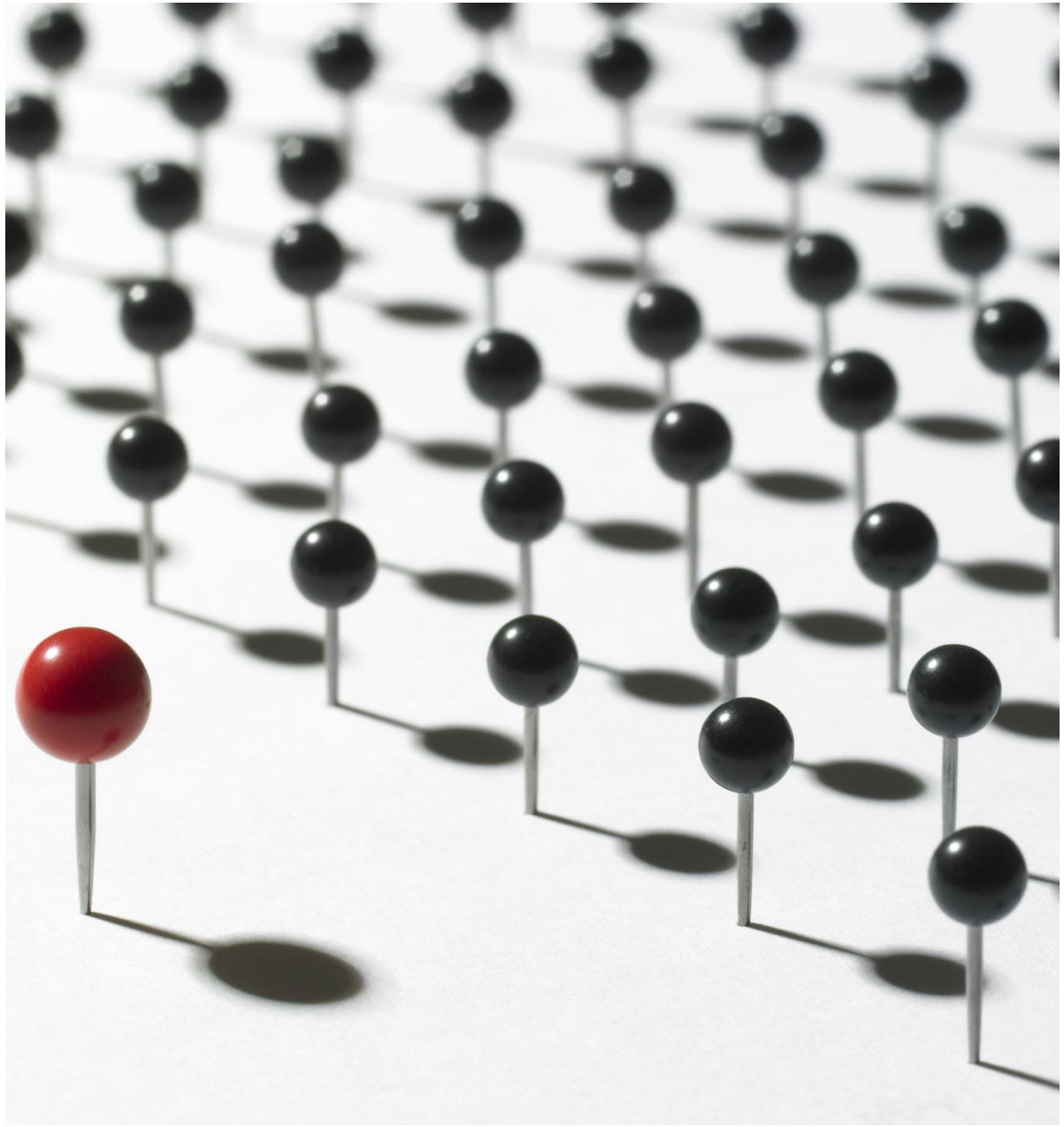
- LLMs can analyze historical clinical trial data and patient records to predict clinical trial outcomes, such as treatment response, disease progression, and survival outcomes.
- By identifying predictive features and patterns in the data, LLMs can help optimize patient selection, treatment allocation, and trial design to improve trial success rates.

Cloud's Generative AI Models in Development...

- Molecular Optimization
- Data Extration from EMR
- Scientific Knowledge Extraction
- Large Molecule Design
- Indication Selection
- Trial Optimization
- Clinical Trial Enrollment
- Trial Performance Co-Pilot
- Smart Data Management
- Major Submission Content Writer

Proprietary Value

- Vertical AI Models for Drug Discovery and Development
- Curated Data Sets known to be both TRUE and VALUABLE (a drawback of many public GenAIs)
- Novel ”Convergence Algorithms” to rapidly train models based on ”exploratory data analysis” and data compression
- Comprehensive end-to-end models covering the entire business value chain in drug development – we are NOT a point solution!



Drug Repurposing

The top 3 GenAI models above will enable Cloud Pharmaceuticals to generate singular and polypharmacy drug repurposings of over 3,000 approved drugs agains over 22,000 indications very rapidly!

Data Center for Pharma Generative AI

- Employs Nvidia Blackwell GPUs and makes use of NIM and BioNeMo
- Built incrementally in a capital efficient manner
- Used to train multiple GenAI models that are later deployed in the cloud

This is the actual site managed by Borelais Data Center where our training data center will be located in Blonduos

Data Center Location: Blonduos, Iceland

- Leverages clean geothermal power
- Run at night when power is cheap
- Managed by Borealis Data Center
- Cooling is cheap in Iceland
- Pay only for what you use

Contact

- **Ed Addison:**
- ed.addison@cloudpharmaceuticals
- 910-398-1200

- **Cloud Pharmaceuticals:**
- www.cloudpharmaceuticals.com