
# COMP4920 AI Technology Case Study

## Using Deep-Learning Data Prediction/Simulation AI Technology in Drug Development

**Hayton Lam, Yue Yu, Yuyun Zhou and Azhad Haziq Kamarul Zaman**

# I Introduction[1]

Artificial Intelligence ('AI') is a technological phenomenon which has advanced productivity, efficiency and development since its inception. It can predict personality categories and political preferences (Bryson 2018, p. 15), and accelerate research by allowing researchers more flexibility, creativity and efficiency (Markowitz, Boyd and Blackburn, 2024, p. 8). This report focuses on the capabilities of deep-learning data prediction/simulation AI technologies, contending such technology should be deployed for use in drug development. Part II expounds the case study technology and highlights its scope, purpose, benefits and needs. Part III identifies important stakeholders and users. Part IV conducts a value-sensitive design investigation. Part V conducts a similar-technology investigation focusing on AlphaFold. Part VI compares the findings of Parts IV and V. Part VII concludes with this report's recommendation on deployment.

# II AI in Drug Development[2]

## A What is Deep Learning AI?

Part of the "brain-inspired" AI branch, deep learning AI ('DL-AI') focuses on simulating vectors of neural activity to represent concepts and draw relationships between concepts for enhanced generalisation (Bengio, Lecun and Hinton, 2018. p. 59). By allowing networks to form numerous layers (thus becoming 'deep') and exploiting layer compositionality permits layer combining, DL-AI are able to make more abstract features in deeper layers that permit greater and more accurate generalisation (Bengio, Lecun and Hinton, 2018. p. 59). A practical application of DL-AI is its capability to make accurate predictions by evaluating numerous factors to draw relationships that act as strong predictors for particular behaviours or trends. For example, DL-AI has been used to accurately predict clinical outcomes (Pettit et al 2021, pp. 737-8) and identify diseases in patients by analysing common factors and symptoms experienced by patients (Xie, Yu and Lv, 2021). DL-AI is also able to leverage the concepts and relationships it has identified to create simulations and evaluate outcomes. One example is using deep-learning to simulate protein-folding and determine the resulting structure (Noe, Fabritiis and Clementi, 2019). However, DL-AI

[1] This Part was written by the whole group.
[2] This Part was written by Yunyu Zhou, with contributions from Hayton Lam.

depends on varied and extensive datasets to facilitate accurate concept and relationship formalisation (Bengio, Lecun and Hinton, 2018. pp. 59-60).

## B Case Study: Usage in Drug Development

This report proposes leveraging the predictive/simulative capabilities of DL-AI in drug development to predict how different chemical compounds may work together and their effects on human bodies.

### *1. Purpose and Goal*

The purpose of incorporating DL-AI into drug development is to enhance efficiency throughout the process and reduce costs. Traditional non-AI methods demand astonishingly large amounts of time and money to develop a new, effective and marketable drug. On average, developing a new FDA-approved drug requires 10-15 years of research and development, costing over $2 billion USD, with 90% of products failing in development (DiMasi, Grabowski and Hansen 2016, p. 1). However, using AI models can significantly reduce these statistics. In 2024, generative AI developed ISM001-055, a promising drug candidate for idiopathic pulmonary fibrosis ('IPF') (a high-mortality aggressive interstitial lung disease) within 18 months for $2.6 million (Hutson 2024, p. 1494).

Using AI models also reduces development costs. Currently, the number of small drug-like molecules is estimated to be between $10^{23}$ and $10^{60}$. Traditional methods require significant effort to identify target molecules, requiring extensive research and considerable costs (Nowak 2023, p. 1). Researchers can utilise AI models to filter and focus on those most similar to known molecules with the desired properties, to minimise the number of costly experiments (Nowak 2023, p. 1-2).

## *2. Scope & Context*

While DL-AI can be implemented in many stages of drug development such as preclinical phases and post-market surveillance, this report primarily focuses on preclinical phases of drug development. This involves assisting in molecule identification, prediction of potential medicinal components (Vijayan et al 2022, p. 968) and predicting chemical interactions between humans and the drug candidate. The proposed technology will primarily be in the research/lab context and will allow for prediction of drug interaction with simulated humans to assist in refining the drug prior to clinical testing.

## *3. Needs*

A crucial reason for more efficient drug development is the inability for traditional processes to keep up with the rapid emergence of new diseases. New diseases are emerging at an unprecedented rate, with one new disease identified each year (O'Dowd 2007, p. 418). This is due to increased international travel, population growth and rising resistance to antimicrobial drugs (O'Dowd 2007, p. 418). Considering the extensive timeframe and cost for developing drugs for an unseen disease, DL-AI technologies are required to accelerate development and reduce costs to ensure medical research can match the pace of new diseases.

Another reason is that it can assist in shortening the clinical trial phase of drug development by pre-simulating human/animal reactions to the potential drug. Ordinarily, this phase is to test how the drug reacts in humans and animals to allow for further refinement and to reduce any undesirable side-effects. Research on organisms is restricted by strict ethical guidelines like the 3R principle (Replacement, Reduction, and Refinement) which emphasizes minimizing harm and using alternatives where possible. This is even stricter when transitioning to human research. While these guidelines are essential for ensuring ethical standards, they often slow research progress, with strict compliance sometimes halting projects that fail to meet these criteria. Although these rules are undeniably necessary, they pose significant challenges and delays for high-benefits studies.

DL-AI can meet these standards by modelling and simulating biological processes, potentially reducing the reliance on live subjects in early-stage testing (Rudroff 2024, p. 816). Researchers can perform detailed virtual studies before initiating actual organism testing. This satisfies ethical

requirements of minimising live organism use and accelerates research, allowing scientists to explore potential outcomes faster and more efficiently (Rudroff 2024, p. 816).

### *4. Benefits*

In general, DL-AI offers significant benefits in drug development by enhancing efficiency, reducing costs, and providing an ethical alternative to traditional animal and human testing. By bringing new drugs to the market more quickly and cost-effectively, DL-AI ultimately enhances accessibility and affordability for patients. DL-AI-driven drug development technology has the potential to make life-saving medications more widely accessible, particularly in underdeveloped regions with limited resources.

Furthermore, this technology can bring significant social impacts. Currently, over 7,000 rare diseases affect nearly 300 million people globally (The Lancet Global Health 2024, p. 341). This is exacerbated by the burden on those affected. In the United States, direct medical costs for individuals with rare diseases amount to approximately 64,000 USD annually (Yang et al 2022, p. 1). Medications for rare diseases are often ridiculously highly priced, primarily due to small patient populations and limited market for these drugs. Pharmaceutical companies thus raise prices to cover high costs of research and development. However, using DL-AI to streamline and reduce the cost of drug development could provide more affordable and accessible treatments for those affected by rare diseases, as well as individuals facing financial strain from purchasing necessary medications. With lower drug costs, it could ultimately help alleviate the financial burden on healthcare systems for society, especially in countries with universal healthcare.

# III Users & Stakeholders[3]

## A Users

Researchers are the primary user for DL-AI in drug development research. Those involved in drug development tend to be experienced researchers with high income and have wide knowledge about their relevant medical specialization and drug chemical composition. Introducing DL-AI will change how they conduct their research, adapting their work to train AI models and validate results. This will introduce greater emphasis on ensuring data is kept secure and private, and following ethical guidelines for AI usage. They may also need to upskill in data management and collection, as DL-AI is reliant on curated datasets, and require validation to ensure results are correct. Hence, they are both a user and stakeholder as it not only affects the nature of their job, but as they are the primary users of this technology in drug development.

Consider the following persona:

> Clara Chen is a 50-year-old senior researcher at BigPharma, a pharmaceutical company. She is creative, hardworking and patient, having developed these attributes from decades of research experience. She is knowledgeable in chemical drug compositions and is familiar with biomedical technologies used to develop drugs.
>
> Her goal with using DL-AI is to more accurately identify suitable drug candidates and hasten research and development. She also aims to refine the drug via simulation testing to reduce the time and money spent on live clinical trials.
>
> When interacting with DL-AI, she will need to have high computing skills and be familiar with operating AI models. She will also need to be diligent in validating outputs and ensuring that data privacy is maintained during queries and development.

[3] This Part was written by Azhad Haziq Kamarul Zaman, with contributions from Hayton Lam.

## B Stakeholders

### *1. Patients*

Patients are a primary stakeholder as they are the ones who will benefit (or be harmed) by AI-developed medication. They are a varied population with a spectrum of attributes. However, DL-AI deployment will significantly affect patients with rare diseases or currently ineffective medication. While well-off patients may be less affected, those with lower incomes will benefit from reduced drug costs and expedited development times which will lower their total medical costs. It is unlikely that they will understand much of the deployed technology other than what is available online.

Consider the following persona:

> Osas Loba is a 57-year-old retired veteran with low income. He has basic understanding in AI and medicine. He currently suffers from high-cholesterol which makes him dependent on medication with severe side-effects such as headaches and dizziness. It has started to lose potency.
>
> He hopes that AI-developed medication can resolve his disease with minimal side-effects. He understands that he may need to volunteer data for research, and wants his data kept secure and private.

### *2. Pharmaceutical Companies*

Pharmaceutical companies are an entity stakeholder as large corporations. They are primarily the ones employing experienced researchers to assist with developing new drugs, with the aim of developing successful drugs that can be sold and return profit. By using DL-AI, they can reduce costs, accelerate development and increase profit to allow for better research. They may also be able to tap into new markets such as personalised drugs or more willingly develop drugs for rare diseases. However, the use of DL-AI may impact their reputation should a drug cause significant side-effect, and thus they want to ensure DL-AI errors are minimised.

Consider the following persona:

> BigPharma is a well-known pharmaceutical company which produces and develops drugs, generating billions in revenue each year. They employ many experienced researchers and want to develop a new drug for high-cholesterol patients with minimal side-effects.
>
> They want DL-AI technologies to be easily implemented and used by their researchers. They also require the AI-model to be reliable, easily validated and easy to modify for their proprietary purposes. This might be done by partnering with another IT firm to develop in-house models.

3. *International Medical Organisations (e.g. World Health Organisations)*

International medical organisations act as regulators and assist in fostering public trust in new medical developments, drugs and technology. They are generally made up of experienced and well-known figures in the health/medicine world and provide guidance to pharmaceutical companies, practitioners and researchers on ethical practices. They also advocate for fair access to medical care. Introducing DL-AI into drug development will affect these organisations as they will need to determine whether to support such a technology, and if so, ensure that it is receptive to the public. They will also need to evaluate what are ethical/best practices for using this technology.

Consider the following persona:

> WHO is an international medical organisation which regularly provides ethical guidelines and publishes resources to inform the public about recent medical developments and foster trust. They will be responsible for making DL-AI developed drugs palatable and conducting their own research on what ethical/best practices should be adopted. They likewise will call-out companies not adhering to ethical usages.
>
> They expect that DL-AI technologies will be regulated and have proper safeguards to uphold public trust. They expect to have access to the technology to conduct their own studies. They hope that such technology will be used ethically, and to increase access to medicine and accelerate medical development.

*4. Doctors & Pharmacists*

Doctors and pharmacists are indirect stakeholders as they will be the ones prescribing/supplying DL-AI developed drugs to patients. While they vary in private attributes, they are generally very knowledgeable about the drugs they prescribe, and whether they would be suitable for patients. If DL-AI drugs have serious actual/potential side-effects, they will need to warn patients, and/or determine whether it is appropriate to prescribe said drugs. They will also need to inform patients, as they have a right not to use DL-AI developed drugs.

Consider the following persona:

> Dr Natalie Strange is a young GP at a local clinic. She sees various patients with varying health conditions, from general sickness to chronic illnesses. As a compassionate and patient person, she needs to explain the effects of drugs she prescribes and tries to accommodate to her patients' idiosyncrasies.
>
> While she has little knowledge of AI, she hopes that DL-AI will develop more effective drugs for her patients with minimal side-effects. She will want to know whether the drugs developed are reliable, as she doesn't want to prescribe harmful drugs to her patients. She knows some of her patients are distrustful, and so hopes the drugs developed are clearly labelled as 'AI developed' so that she can distinguish and tell her patients this information.

# IV Value-Sensitive Design & Investigation[4]

This Part considers a value-sensitive design of DL-AI in drug development. It examines the value alignment in deploying such a technology based on Microsoft's six ethical principles, considering the perspectives of researchers, patients and pharmaceutical companies ('companies').

## A Fairness

Fairness refers to the absence of bias in scientific research, and the equitable distribution of risks/benefits of research and preventing discrimination (Department of Health, Education and Welfare, 1979) based on personal characteristics (Ballard, Chappell and Kennedy 2019, p. 425).

### *1. Researchers*

DL-AI training relies on diverse and large datasets. Researchers must ensure the integrity of the dataset, ensuring data properly represents patients with differing private attributes such as race, colour, sex, ethnicity, etc (Bernal and Mazo 2022, p. 14). Datasets containing historical data may introduce biases against certain groups, even if unintended. Sometimes, even when protected variables like gender, race and religion are excluded from the dataset on purpose, unanticipated proxies for these variables could still be used to reconstruct biases, leading to "bias by proxy" that is difficult to detect and avoid (Tsamandos et al 2021 p. 221).

### *2. Patients*

For patients, accessibility of research results is a paramount consideration. While using DL-AI may lower drug development costs, it may raise the cost due to needing to develop relevant technologies and collect accurately labelled datasets. Another concern is that the convenience of DL-AI may create personalised/privatised drug development (Levaggi and Levaggi, 2024). Either consequence leads to less accessible drugs, resulting in unfair distributions of healthcare. Furthermore, as AI models require large volumes of data, research may skew towards more

[4] This Part was written by Yue Yu, with contributions from Hayton Lam.

common diseases with more available data, placing patients suffering from rare disease in a disadvantaged position.

*3. Companies*

Likewise, fairness for companies involves ensuring that research is not optimised on profit-potential, but on equitable access to medicine. Companies must ensure that rare diseases are still researched, potentially through a system similar to mandatory 'pro-bono' requirements some systems of law adopt (Maguire, Shearer and Field 2014, p. 1176). This would mean companies may be required to conduct at least one drug research related to rare or disadvantaged diseases, which may alleviate equity concerns.

## B Privacy and Security

Considering DL-AI requires large volumes of data, its application in drug development means collection of sensitive medical data from patients. Hence, data privacy and security refer to protecting data and maintaining privacy rights (Ballard, Chappell and Kennedy 2019, p. 425).

*1. Researchers*

As users of DL-AI, researchers must protect the privacy and security of patient data. Simply removing all sensitive information from data collected is infeasible as it would significantly reduce the level of detail within data, making potentially useful data unavailable and increasing unreliability of the DL-AI model. A balance must be struck between privacy, security, availability and informativeness of user data (Bernal and Mazo 2022, p. 12). Researchers must also ensure data security is preserved by anonymising data, conducting data privacy impact assessments (Bernal and Mazo 2022, p. 13). Data encryption and limiting data access to necessary personnel will ensure sensitive patient data will not be easily exposed. '

*2. Patients*

From a patient's perspective, privacy and security is vital as it is their data which may be exposed. While not their responsibility to balance privacy and availability, they are able to control how their data is used or accessed. Patients are generally unaware of what type of information is stored about them and its usage (Tsamados et al 2022, p. 223). Hence, obtaining informed consent per the Belmont Report is critical, showing respect to their autonomy as research participants volunteering their personal data (Department of Health, Education and Welfare 1979, p. 10).

*3. Companies*

Privacy and security are also important for companies as respecting the privacy of user information is essential for building good reputation and rapport amongst patients. If privacy is managed poorly, it is unlikely patients would willingly volunteer data, resulting in insufficient datasets for research. Furthermore, all research data/results are the company's intellectual property meaning that the company will prioritise data security to ensure all relevant information will not be leaked. Maintaining security may involve device encryption, proper personnel training to encourage them to act accordingly with moral principles and avoiding common cybersecurity mistakes and traps (Bernal and Mazo 2022, p. 13).

## C Reliability and Safety

Reliability refers to the consistency and safety of deployed AI systems in various conditions (Ballard, Chappell and Kennedy 2019, p. 425). In drug-development contexts, this is particularly important as unreliability may cause serious health/financial consequences.

*1. Researchers*

Traditionally, researchers face the issue of variable reliability when evaluating the safety and reliability of developed drugs. DL-AI presents a solution capable of overcoming such challenges

in traditional methods by analysis and simulation (Abbas et al 2024, p. 2). However, results must still be validated by researchers through aspects such as reproducibility and replicability. Reproducibility refers to whether a model yields consistent results across various studies using their own data, while replicability tests whether a model yields consistent results using identical data, computational methods and conditions of analysis (US Food and Drug Administration 2023, p. 18).

*2. Patients*

Patients are powerlessly dependent on companies and researchers to ensure DL-AI developed drugs are safe and reliable. Failure to do so may result in serious health/financial consequences. Arguably, patients are most concerned that this ethical principle is met by other stakeholders.

*3. Companies*

Reliability and safety of DL-AI models is vital to companies as it is essentially in building good reputation and rapport with patients who are the target market group. Furthermore, developed drugs undergo approval testing from relevant regulatory institutions including the FDA and EMA. If rejected, companies lose development time and increase costs. Hence, they must ensure the model is safe and reliable.

## D Transparency

Transparency refers to an AI's capacity to describe its decision-making process to relevant stakeholders, sometimes termed as 'interpretability' (Ballard, Chappell and Kennedy 2019, p. 425).

*1. Researchers*

Researchers require an AI model which can be assessed, validated and evaluated to facilitate iterative development processes. These require a transparent AI model, yet most AI models (including DL-AI) are black-box, meaning the internal mechanisms of the AI are unknown. In drug development contexts, drug discovery is considered a black-box optimisation problem where researchers search for small molecules which optimise specific properties (Quereshi et al

2023, p. 15). Without an interpretable/transparent DL-AI model, determining result accuracy or understanding a measurement's precision to properly exploit data would be hindered (Shockley 2016, p. 296). A possible solution is to ensure Explainable DL-AI models are used. This would set out the decision-making process, assisting researchers in assessing and validating the AI's result, or trace steps to correct potential errors (Malizia and Paterno 2023, p. 21).

*2. Patients*

For patients, transparency arises from the notion of 'trusting' the medication. Unlike the scientific community where transparency is required to validate results and determine exploitable data points, patients' perception of transparency is of whether the developed drug and relevant parties are trustworthy. Acceptance of AI products is built upon transparency (Zhang et al 2020, p. 103), and understanding how a product is created such that it meets patient standards. XAI may not be an appropriate solution, as professional explanations may have little to no explanatory patients as non-technical audiences (Malizia and Paterno 2023, p. 21). Instead, making the development process publicly accessible may foster greater patient trust/confidence, even if they may not fully understand the content, as they appreciate that researchers and companies are acting in good faith by being transparent.

*3. Companies*

Like patients, companies are motivated to make DL-AI models transparent to foster patient trust and to promote their drugs. Additionally, companies must ensure the development process is transparent enough to comply with regulatory requirements from relevant agencies such as the FDA and EMA who are increasingly wary of AI drug development. For example, the European Union's ('EU') *EU AI Act 2024* places requirements on AI applications in medicine to be sufficiently transparent to enable users to interpret system output and use it appropriately (Vokinger and Gasser 2021, p. 4).[5] This may conflict with companies' own interests as the development process and data may be considered their intellectual property. During COVID-19, WHO established a technology share pool for companies to foster vaccine development (World Health Organisation 2020, p. 3), which no vaccine-producing company has signed up to (Bernal

[5] Though Vokinger and Gasser refer to the *European Union Artificial Intelligence Act* as a proposed regulation, it has since been passed on 13 March 2024.

and Mazo 2022, p. 14). While regulatory pressure may motivate companies to be more transparent, companies will need to strike a balance between transparency (for patients) and opacity (for its own interests).

## E Inclusion

Inclusiveness focuses on the design and data collection phase in AI. This is different from the fairness principle, which focuses on how AI should treat all equally (Ballard, Chappell and Kennedy 2019, p. 425). An inclusive AI system should be designed to prevent exclusion and benefit everyone.

### *1. Researchers*

For researchers, an inclusive dataset ensures scientific accuracy and data integrity. Without a diverse dataset, DL-AI models will produce results relevant only to groups represented by the data, thus misrepresenting/omitting other populations and resulting in unfair results. Researchers therefore must prioritise the diversity and inclusiveness of training data collected, ensuring it encapsulates diverse patient populations and accounts for demographic, socioeconomic and cultural factors (Tilala et al 2024, p. 3). For rare diseases, this poses an additional challenge beyond the existing difficulty in collecting data for training. Ensuring each group is evenly represented in such a small sample size is difficult, exacerbated by the possibility that some underrepresented groups may be less available to supply data for research. Hence, researchers must also ensure inclusivity by using a variety of data augmentation techniques, and actively seeking and incorporating underrepresented groups to ensure diversity (Tilala et al 2024 p. 3).

### *2. Patients*

For patients, inclusivity is important to avoid potential biases in DL-AI models, which may result in drugs with disproportionate effects. Hence, patients are motivated to ensure their own group is not underrepresented in terms of gender, culture and other relevant personal backgrounds. This may thus reduce data volunteering hesitancy. However, for patients from disadvantaged groups like low-income regions, patients with rare diseases or with limited reach to the company developing the drug it becomes difficult to ensure they are fairly represented. Where there are small sample sizes (like rare diseases), patients may also lack confidence that DL-AI models

have enough data representative of their group.

*3. Companies*

For companies, inclusivity may be mandated by regulatory bodies such as the FDA and EMA, who advise that AI medical products must identify and mitigate biases in data (US Food and Drug Administration 2024, p. 3). When using DL-AI for medical products, a major root of bias arises from a dataset not being inclusive enough, resulting in skewed results towards certain groups. Thus, if the developed drug is to obtain approval from regulatory bodies, companies must ensure inclusivity occurs in DL-AI. Furthermore, inclusivity is financially beneficial as the more diverse the patients are, the more consumers they have purchasing their product. Hence, fostering an inclusive brand is vital, even from a purely financial perspective.

## F Accountability

Accountability refers to the idea that the people who design and deploy AI systems must be accountable for how their systems operate (Ballard, Chappell and Kennedy 2019, p. 425). This is crucial in drug development where identification of liability in accidental scenarios is necessary.

*1. Researchers*

Researchers will require precise definitions of responsibilities, otherwise they will be hesitant to participate in research with indeterminate liability. Despite the efficacy and efficiency of DL-AI, it is not infallible (Tilala et al 2024 p. 5). Technical bottlenecks such as lack of transparency and explainability undermine DL-AI's scrutability (Tsamados et al 2022, p. 224). These could make researchers accountable if they do not fully comprehend DL-AI's decision-making process, or where unexpected results arise from drug usages in untested scenarios. If erroneous results are produced during the molecule prediction stage, reconducting research on new molecules would increase costs. Likewise, mispredictions or simulation errors in testing interactions with organisms may cause serious medical malpractice if uncaught before live clinical testing occurs.

*2. Patients*

Patients are likewise concerned about DL-AI accountability. Unlike doctors, technologists are not accountable by law for their actions (Naik et al 2022, p. 4). Where medical malpractice occurs, it may be difficult to ascertain which party is responsible for compensation and liability. DL-AI's black-box nature gives leeway for insurance companies, researchers and companies leeway to 'kick around responsibility'. Even if patients can initiate legal proceedings, patients are immensely disadvantaged when facing companies with specialised legal departments and massive budgets. They may engage in costly stalling tactics such as 'trolley-loading' (see *Thomas v SMP (International Pty Ltd*) to inflate costs and force patients to back down.[6]

*3. Companies*

Likewise, companies highly value accountability as they need to manage regulatory compliance with ethical responsibility, market trust and liability risks. Interestingly, studies have found humans trust humans with AI assistance more than humans or AIs by themselves (Bilos and Budimir 2023, p. 39). Yet most patients are yet to trust DL-AI enough to interfere with medicines, as consequences seriously affect one's livelihood. Without clear definitions of accountability, companies' reputations would be at great risk if sued, and any attempt to distribute responsibility is yet to be defined.

[6] In *Thomas v SMP (International Pty Ltd)* [2010] NSWSC 822, SMP filed 500-paged affidavits which were largely irrelevant material. They also continually objected to each of the 3000 paragraphs, which hindered the administration of justice and created prohibitively expensive legal costs.

# V Ethical Concerns in Similar Technologies[7]

Utilising artificial intelligence technologies like deep learning is not new to the medical and bioscience landscape. AI has been used to assist in recommending treatment plans to cancer patients (Smith 2021, pp. 535-6), analysing radiology images (Tiang 2020, p. 3), and simulating protein folding (Nussinov et al, 2022). This Part examines AlphaFold, a revolutionary deep-learning AI model designed to simulate protein folding (Nussinov et al 2022, p. 6732) as a similar technology. Its contribution to medical research is unparalleled, allowing for accurate structure predictions of biomolecular interactions and vastly expanding the scope and complexity of experiments (Abramson et al 2024, p. 1; Varadi et al 2022, pp. 442-3). These capabilities assist in identifying drug candidates and simulating drug-to-drug interactions.

This Part will consider how AlphaFold is affected by two ethical areas of concern: accountability and data privacy/security.

## A Accountability

Accountability is particularly important for patients as false-negatives/positives and faulty data simulations have potential for significant harm in medical settings such as exacerbated injuries, severe and long-lasting side effects or even death. We examine two aspects: AlphaFold as an accountable AI data prediction technology and attribution of liability for AlphaFold errors. We then discuss these aspects' impact on the technology.

### *1. Accountable AI*

AlphaFold has been observed to produce spurious predictions due to noisy data (Abramson et al 2024, p. 8) and suffers from hallucinations where it predicts plausible but non-existent molecular structures and has varying rates of success dependent on the particular structure it is predicting (Desai et al 2024, p. 6). This is exacerbated by the restricted access to AlphaFold 3's codebase which has negative implications of proper accountability in scientific research (Desai et al 2024,

[7] This Part was written by Hayton Lam.

p. 6). Furthermore, Tan and Zhang (2023, p. 2166) highlight that AlphaFold is a black-box system where the question of 'why' particular results are produced is not easily known. It is difficult for researchers to fine-tune or debug AlphaFold, nor can it be utilised for screening tools or develop optimised tools (Trager, Tuting and Kastritis 2024, p. 1555). Rudroff (2024, pp. 813-14) further suggests that such black-box models cause difficulties in validating results and causes over-reliance. This ultimately undermines research and public accountability in research contexts where AlphaFold is being used (Tan and Zhang 2023, p. 2173).

Tan and Zhang (2023, p. 2173) suggest that AlphaFold's next step is 'ExplainableFold', an explainable AI version which explains its reasoning/steps in folding protein structures. This allows for accountability from researchers as they comprehend the 'why' of results and have confidence in applying prediction results into drug development. In the meantime, accounting for AlphaFold's mistakes/hallucinations requires heightened diligence from researchers to ensure drugs developed using such AI technology do not rely on incorrect information. When something goes wrong, it may be difficult to determine the proper cause. This leads to the second accountability aspect: attribution of liability.

### *2. Attribution of Liability*

When a developed drug has unexpected side-effects/consequences due to errors or misprediction, who is accountable? The researchers for lack of diligence, the AI designer for designing a flawed prediction model or the pharmaceutical company for improper supervision? There is currently no unified school of thought, and no such test case involving attribution of liability caused by AI exists in Australia (Nolan 2022, p. 68). We consider two possible schools of thought below.

In AI medical decision-making, Chung and Zink (2018, pp. 77-8) contend that doctors are accountable and ought to double-check AI outputs as the human intermediaries between patients and AI. Adapting them to AlphaFold, this would mean researchers are primarily accountable for lack of diligence. Yet such a heavy attribution of 'blame' or accountability might create a

cautious attitude towards using AI-assisted simulation/prediction technology, as simple mistakes may result in millions in damages.[8]

Meanwhile Frank (2019, pp. 293-4) suggests that the AI designers (or company) ought to be liable for negligent design and training of the AI technology. This attribution of liability is a classic legal model;[9] a manufacturer must ensure their product is designed properly and does not harm users. Design flaws such as poor training from limited datasets and lack of clinical trials are all accountable actions which trace back to the manufacturer.

Regardless, the uncertainty in liability is problematic as patients and other affected stakeholders will not have certainty as to who is liable for compensation. This risks creating a 'legal limbo' situation where each related party 'kicks responsibility around', leading to uncompensated injured parties.

### *3. Impacts of Accountability*

Ultimately, uncertainty as to who should be accountable may lead to hesitation in adopting AI-assisted research techniques by researchers. As Tetlock (1991, p. 470) suggests, where ambiguity exists, people tend to avoid the ambiguity. Applied here, an impact of ambiguous/uncertain accountability is that researchers and pharmaceutical companies are less likely to adopt AI-tools in an abundance of caution. Furthermore, legislative and regulatory frameworks are slow to address concerns about accountability and often lack necessary expertise (Paterson 2024, pp. 24-6). Given that lack of attributable accountability undermines public acceptance of AI technology (Zhang et al 2020, p. 103), ambiguity in accountability diminishes the possibility of practical deployment of any AI technologies, especially in a sensitive, high-stakes field like medicine. Hence, AlphaFold's effectiveness is dependent on its actions being accountable, and for attributions of liability to be transparent and defined to reduce ambiguity.

[8] See for example the current class action legal proceedings against AstraZeneca for rare side-effects caused by COVID-19 vaccines with estimated damages of £120,000 per plaintiff (Dyer 2023, p. 1).

[9] See the seminal case *Donoghue v Stevenson* [1932] AC 562; All ER Rep 1 (snail in glass bottle case), which established manufacturer's strict liability for their products.

## B Data Privacy and Security

AlphaFold can be used for more individualised medicines to develop specific use-case drugs but requires sensitive information such as genomic and medical data (Desai et al 2024, p. 5). Even in general use-cases, Vora et al (2023, p. 1916-18) highlights large, accurate, and diverse datasets are important for deep-learning AI technologies like AlphaFold to simulate drug-to-drug interactions and predict drug candidates to diseases. We analyse three aspects: cloud-computing, data anonymisation and data security practices before addressing their impacts.

### *1. Cloud Computing*

AlphaFold currently functions as a cloud database, where experiments can be queued and retrieved from a server (Varadi et al 2024, pp. 369-70). If the cloud is insecure, or if results are publicly available, then data provided can be easily compromised. AlphaFold does not have privacy-enhancing techniques to protect sensitive data in query sequences (Unal, Pfeifer and Akgun 2024, p. 2). This leaves genomic and biological data vulnerable to bad-actor interception of query sequences and attacks on the cloud database to mine sensitive information. Contrasting with a traditional local-hosted model, cloud-based computing introduces a new dimension of attacks which need to be shielded against.

### *2. Synthetic Data & Anonymisation*

AI simulation/prediction technologies bring about a new type of data: synthetic data (Gal and Lynskey 2024, p. 1105). For example, the protein-folding outputs produced by AlphaFold are synthetic data that can be used to further refine queries or guide research directions. Gal and Lynskey (2024, p. 1125) note that synthetic data risks re-identifying original personal data either by the model itself or being reverse engineered by an attacker. Likewise, AlphaFold tends to reuse data from previous queries (Gadiya et al 2023, p. 3). This introduces a secondary angle whereby patient data can be indirectly compromised; discovery via independent queries when AlphaFold reuses results/query data from previous queries.

While data anonymisation is traditionally suggested to protect data privacy, introducing noise or bucket variables to obscure data (Breugel et al 2024, p. 9), Agrawal, Misra and Padmanabhan

(2020, p. 21) highlight that genomic data is uniquely counter-anonymisation; genomic data are unique identifiers of individuals which make data masking unviable. Suggested solutions include structuring laws to prevent interlinking of genomic data with other datasets, introducing harsh penalties to disincentive re-identification and rigid consent-revocation options (Agrawal, Misra and Padmanabhan 2020, p. 21), however these are reliant on specific jurisdictions and are unlikely to provide data privacy individuals expect. Though AlphaFold has no specific data anonymisation practices, Roy et al (2023, pp. 10-12) suggest integrating 'differential privacy' into medical research AI models like AlphaFold, which integrates noise into each stage of the AI training, query and prediction to mitigate genome-level anonymity issues. This may mitigate issues with anonymisation that currently exists with AlphaFold.

*3. Data Security Practices*

Humans are the weakest point in cybersecurity (Leach 2003, p. 692). Thus, researchers using AlphaFold and similar AI models in data simulation utilise data encryption and limited personnel access permissions to protect data at rest (Roy et al 2023, p. 7). In transit, AI models like AlphaFold use HTTPs protocols and digital signatures to prevent interception and modification from bad-actors, which can cause malicious mispredictions (Roy et al 2023, p. 8). By maintaining strict data security practices and control, the risk of data compromise due to humans is limited, maintaining data privacy and security.

*4. Impacts*

Users are less likely to volunteer data or otherwise use services where they are unsure about their data privacy (Schaewitz, Winter and Kramer 2021, p. 769). If proper data privacy is not achieved, patients are likely less inclined to volunteer their data for medical research which may undermine any optimisation or enhancements brought about by using AI for drug development. This would undermine AlphaFold's prediction effectiveness as its dataset becomes less diverse and smaller. Furthermore, improper data privacy and security attract bad-actors into attacking researcher databases and the cloud-hosted models directly. This is evident in the Optus and Medibank data breaches, where attackers exploited poor data privacy and security policies to mine millions of users' data.

# VI Comparison to Conceptual Investigation[10]

When comparing Part III and IV, a lot of the conceptual investigation findings on accountability and data privacy were complementary. However, in many areas, similar technologies highlighted new dimensions in existing concerns that were overlooked by our conceptual investigation. These include privacy implications from current technologies being reliant on cloud-computing and the effect of black-box AIs on ascertaining accountability. However, a major finding is that the AlphaFold investigation did not consider companies as an important stakeholder, hence many issues and concerns related to companies raised in the conceptual investigation were not addressed in the former. The below table summarises and highlights the key differences and similarities between the findings and conceptual investigation.

| Ethical Concerns | Specific Issues | Conceptual Investigation | AlphaFold Investigation | Conclusion |
|---|---|---|---|---|
| **Accountability** | Accountability of DL-AI Model/System | Considered the risk of error, and the issues with black-box AI models.<br><br>Highlighted that it is difficult to have a transparent 'deep-learning' model. | Complementary findings, finding that AlphaFold was not transparent with how it achieved its results. Also highlighted a trend of overreliance and reduced research accountability associated with black-box AI models. | Both investigations highlighted issues with black-box AI models and the difficulty in making them accountable.<br><br>AlphaFold expanded on the practical aspect of this, demonstrating the diligence required by researchers to validate results, and the possibility of creating 'ExplainableFold |

[10] This Part was jointly written by Hayton Lam and Yue Yu.

| | | | | |
|---|---|---|---|---|
| | | | | ’ which would assist in accountable and transparent research. |
| | Attribution of Liability | Found that there was a real risk of indeterminate responsibility that each party (researcher, pharmaceutical company, insurance company, AI designer) blames each other.<br><br>Also highlighted that indeterminate liability may result in hesitancy from patients on accepting AI technologies in drug development, and from researchers in using such technologies. | Had similar findings, but highlighted the practical complications of attributing liability, especially when the AI model is not necessarily accountable. It highlighted conflicting views of accountability, alongside an uncertain legal landscape. | Both investigations agreed on core issues of indeterminate liability, but the AlphaFold investigation provided additional dimensions, and grounded the speculative reasoning behind indeterminate liability in the conceptual investigation by providing background information as to the source of indeterminacy. |
| **Data Privacy & Security** | Cloud Computing | Not contemplated | Highlighted the need for privacy-enhancing techniques internal to AI models, and the vulnerabilities associated with cloud-hosted AI models. | AlphaFold highlights new possible angles of attack and key features of this vulnerability. |

| | | | | |
|---|---|---|---|---|
| | Synthetic & Anonymised Data | Discussed the need for data anonymisation to protect private data, and the importance of informed consent with large datasets. | Demonstrated the practical difficulties of anonymising genomic data required for similar technologies and introduced an issue of synthetic data protection. | AlphaFold highlighted overlooked practical considerations of measures used to protect data such as anonymisation unique to the conceptual investigation's AI use-case. |
| | Security Practices | The importance of good personnel data security practices and keeping data secure. | In full agreement with conceptual investigation. | No substantial difference |

# VII Recommendation[11]

We **recommend** the use of 'deep-learning' AI in drug development in simulating the effects of potential drugs on humans and other test subjects. The societal benefits of deploying such technology are immense. It will advance medical research progress significantly, allowing researchers to identify trends and new drug candidates more efficiently while shortening the simulation period. It also allows for cheaper drug development and for drugs to be developed to cure currently incurable diseases, allowing for cheaper and more accessible medicines.

However, our recommendation is subject to the following changes which need to be made to any 'deep-learning' AI technology:

1. Proper accountability
2. Explainable & Privacy-Integrated AI
3. Localised Model or Secure Cloud Data Control

These changes are derived from the conceptual and AlphaFold investigations.

*1. Proper Accountability*

As highlighted in Parts III and IV, accountability attribution must be clearly defined in order for such technology to be accepted by the public, and by researchers. This may take the form of legislation being passed or developing a transparent accountability policy which all involved parties can access and understand. This will reduce the ambiguity of accountability and thus reduce hesitation in adopting AI. Furthermore, it will allow potential patients/users of AI developed drugs to understand who is responsible/liable to compensate when things go wrong, giving reassurance. An alternative scheme is to adopt the indemnification scheme undertaken by the Australian Government when COVID-19 vaccines were released. This involved the government compensating all parties affected by serious side-effects from vaccines developed. This equally would assuage the accountability ambiguity.

[11] This part was written by Hayton Lam with contributions from the whole group.

### *2. Explainable & Privacy-Integrated AI*

The AI technologies must also be an explainable and differential-privacy integrated AI. This is to firstly make the AI accountable and assist researchers in efficient and accountable drug development research. As explored in Part IV, 'ExplainableFold' produces transparent and accountable outcomes for AlphaFold-like technologies, and thus this approach should be embraced. Furthermore, differential-privacy integration will ensure that the practical difficulties of anonymising genomic and medical data essential to drug development AIs are overcome, maintaining user privacy and security. This will bolster confidence in users, making them more likely to volunteer their data for research.

### *3. Localised Models or Secure Cloud Data Control*

Finally, to mitigate vulnerabilities to cloud-hosted AI models, we recommend researchers use localised versions of AI models. Sharing similar data privacy and security concerns as lawyers, adopting their approach of a custom, local-based model would ensure that sensitive research data is not exposed to external parties.[12] If that is not feasible, then secure data protocols and practices including digital authenticated query signatures and data encryption must be used to protect data transmission between the cloud-hosted model and the researcher. Furthermore, standard security practices must be employed to protect against bad-actor attacks against the cloud model to secure sensitive data.

[12] See MinterEllison's implementation of a local-stored version of ChatGPT for its lawyers: https://www.techrepublic.com/article/australian-legal-profession-adapting-ai/.

# VIII Reference List

Abbas, MKG, Rassam, A, Karamshahi, F, Abunora, R, and Abouseada, M 2024, 'The Role of AI in Drug Discovery', *ChemBioChem*, vol. 25, no. 14, accessed 2 November 2024, <https://doi.org/10.1002/cbic.202300816>.

Abramson, J, Adler, J, Dunger, J, Evans, R, Green, T, Pritzel, A, Ronneberger, O, Willmore, L, Ballard, AJ, Bambrick, J, Bodenstein, SW, Evans, DA, Hung, C-C, O'Neill, M, Reiman, D, Tunyasuvunakool, K, Wu, Z, Žemgulytė, A, Arvaniti, E, Beattie, C, Bertolli, O, Bridgland, A, Cherepanov, A, Congreve, M, Cowen-Rivers, AI, Cowie, A, Figurnov, M, Fuchs, FB, Gladman, F, Rishub, J, Khan, YA, Low, CMR, Perlin, K, Potapenko, A, Savy, P, Singh, S, Stecula, A, Thillaisundaram, A, Tong, C, Yakneen, S, Zhong, ED, Zielinski, Michal, Zidek, A, Bapst, V, Kohli, P, Jaderberg, M, Hassabis, D, and Jumper JM 2024, 'Accurate Structure Prediction of Biomolecular Interactions with AlphaFold 3', *Nature*, vol. 630, pp. 493–500: DOI:10.1038/s41586-024-07487-w

Agrawal, H, Misra, P and Padmanabhan A 2020, 'Balancing Privacy with Technology Use: Lessons from Covid-19', Digital Pathways Paper Series, University of Oxford.

Ballard, S, Chappell, KM, and Kennedy, K 2019, 'Judgment Call The Game: Using Value Sensitive Design and Design Fiction to Surface Ethical Concerns Related to Technology', *Proceedings of the 2019 on Designing Interactive Systems Conference*, pp. 421-455, DOI: 10.1145/3322276.3323697.

Bengio, Y, Lecun, Y and Hinton, G 2021, 'Deep Learning for AI', *Communications of the ACM*, vol. 64, no. 7, pp. 58–65, DOI:10.1145/3448250.

Bernal J, and Mazo, C 2022, 'Transparency of Artificial Intelligence in Healthcare: Insights from Professionals in Computing and Healthcare Worldwide', *Applied Sciences*, vol. 12, no. 20, accessed 2 November 2024 <https://www.mdpi.com/2076-3417/12/20/10228>.

Bilos, A, and Budimir, B 2023, 'How Much Do We Trust in AI? Exploring the Impact of Artificial Intelligence on User Trust Levels', paper presented at the 6th International Scientific Conference on Digital Economy DIEC 2023, Tuzla, BiH, August 2023.

Breugel, BV, Liu, T, Oglic, D, and Schaar, MVD 2024, 'Synthetic data in biomedicine via generative artificial intelligence', *Nature Reviews Bioengineering*, accessed 8 November 2024, <https://www.nature.com/articles/s44222-024-00245-7>.

Bryson, JJ 2019, 'The Past Decade and Future of AI's Impact on Society', *Towards a New Enlightenment? A Transcendent Decade*, vol. 11 accessed 18 November 2024, <https://purehost.bath.ac.uk/ws/portalfiles/portal/190042349/Bryson_BBVA_Sep2018.pdf>.

Chung, J and Zink, A 2018, 'Hey Watson - Can I Sue You for Malpractice? Examining the Liability of Artificial Intelligence in Medicine', *Asia Pacific Journal of Health Law & Ethics*, vol. 11, no. 2, pp. 51-80.

Department of Health, Education and Welfare 1979, *Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*, United States of America.

Desai D, Kantliwala, SV, Vybhavi, J, Ravi, R, Patel, H, and Patel J 2024, 'Review of AlphaFold 3: Transformative Advances in Drug Design and Therapeutics', *Cureus*, vol. 16, no. 7, accessed 6 November 2024, <https://pmc.ncbi.nlm.nih.gov/articles/PMC11292590/pdf/cureus-0016-00000063646.pdf>.

DiMasi, JA, Grabowski, HG, and Hansen, RW 2016, 'Innovation in the Pharmaceutical Industry: New Estimates of R&D Costs', *Journal of Health Economics*, vol. 47, pp. 20–33.

Dyer, C 2023, 'Patients Launch Legal Action Against AstraZeneca Over its Covid-19 Vaccine', *BMJ*, vol. 380, accessed 8 November 2024, <https://www.bmj.com/content/380/bmj.p725>.

Frank, X 2019, 'Is Watson for Oncology per se Unreasonably Dangerous?: Making A Case for How to Prove Products Liability Based on a Flawed Artificial Intelligence Design', *American Journal of Law & Medicine*, vol. 45, no. 2-3, DOI:10.1177/0098858819871109.

Gadiya, Y, oannidis, V, Henderson, D, Gribbon, P, Rocca-Serra, P, Satagopam, V, Sanson, SA, and Gu, W 2023, 'FAIR Data Management: What Does it Mean for Drug Discovery?', *Frontiers in Drug Discovery*, accessed 8 November 2024, <https://www.frontiersin.org/journals/drug-discovery/articles/10.3389/fddsv.2023.1226727/full>.

Gal, MS, and Lynskey, O 2024, 'Synthetic Data: Legal Implications of the Data-Generation Revolution', *Iowa Law Review*, vol. 109, no. 3, pp. 1087-1156.

Hutson, M 2023, 'AI for Drug Discovery is Booming, But Who Owns the Patents?', *Nature Biotechnology*, vol. 41, pp. 1491-1503.

Leach, J 2003, 'Improving User Security Behaviour', *Computers & Security*, vol. 22, no. 8, pp. 685-692, DOI:10.1016/S0167-4048(03)00007-5.

Levaggi, L and Levaggi, R 2024, 'Competition, Value-based Prices and Incentives to Research Personalised Drugs', *Socio-Economic Planning Sciences*, vol. 93, accessed 2 November 2024 <https://www.sciencedirect.com/science/article/pii/S0038012124000697>.

Maguire, R, Shearer, G, and Field, R 2014, 'Reconsidering Pro Bono: A Comparative Analysis of Protocols in Australia, The United States, The United Kingdom and Singapore', *University of New South Wales Law Journal*, vol. 37, no. 3, pp. 1164-1197.

Malizia, A, and Paterno, F 2023, 'Why Is the Current XAI Not Meeting the Expectations?', *Communications of the ACM*, vol. 66, no. 12, pp. 20-23, DOI: 10.1145/3588313.

Markowitz, DM, Boyd, RL and Blackburn, K 2024, 'From Silicon to Solutions: AI's Impending Impact on Research and Discovery', *Frontiers in Social Psychology*, vol. 2 accessed 18 November 2024, <https://www.frontiersin.org/articles/10.3389/frsps.2024.1392128/full>.

Naik, N, Hameed, BMZ, Shetty, DJ, Swain, D, Shah, M, Paul, R, Aggarwal, K, Ibrahim, S, Patil, V, Smriti, K, Shetty, S, Rai, BP, Chlosta, P, Somani, BK 2022, 'Legal and Ethical Consideration in Artificial Intelligence in Healthcare: Who Takes Responsibility?', *Frontiers in Surgery*, vol. 9, accessed 2 November 2024, <https://www.frontiersin.org/journals/surgery/articles/10.3389/fsurg.2022.862322/full>.

Nowak, D, Bachorz, RA, and Hoffmann, M 2023, 'Neural Networks in the Design of Molecules with Affinity to Selected Protein Domains', *International Journal of Molecular Sciences*, vol. 24, no. 2, pp. 1762-1778, DOI:10.3390/ijms24021762.

Nussinov, R, Zhang, M, Liu, Y, and Jang, H 2022, 'AlphaFold, Artificial Intelligence (AI), and Allostery', *Journal of Physical Chemistry B*, vol. 126, pp. 6372-6383, DOI:10.1021/acs.jpcb.2c04346.

O'Dowd, A 2007, 'Infectious Diseases are Spreading More Rapidly Than Ever Before, WHO Warns', *BMJ,* accessed 5 November 2024 <https://doi.org/10.1136/bmj.39318.516968.DB>.

Paterson, JM 2024, 'Regulating Generative AI in Australia: Challenges of Regulatory Design and Regulator Capacity' in P Hacker, P, S Hammer, A Engel, & B Mittelstadt (eds), *Handbook on Generative AI*, Oxford University Press (forthcoming), pp. 1–26.

Pettit, RW, Fullem, R, Cheng, C, and Amos CI 2021, 'Artificial Intelligence, Machine Learning, and Deep Learning for Clinical Outcome Prediction', *Emerging Topics in Life Sciences*, vol. 5, pp. 729-745, DOI:10.1042/ETLS20210246.

Qureshi, R, Irfan, M, Gondal, TM, Khan, S, Wu, J, Hadi, MU, Heymach, J, Le, X, Yan, H, and Alam, T 2023, 'AI in Drug Discovery and Its Clinical Relevance', *Heliyon*, vol. 9, no. 7, accessed 2 November 2024, <https://pmc.ncbi.nlm.nih.gov/articles/PMC10302550/>.

Roy, P, Chandrasekaran, J, Lanus, E, Freeman, L, and Werner, J 2023, 'A Survey of Data Security: Practices from Cybersecurity and Challenges of Machine Learning', research paper, Cornell University, accessed 8 November 2024, <https://arxiv.org/abs/2310.04513>.

Rudroff, T 2024, 'Artificial Intelligence as a Replacement for Animal Experiments in Neurology: Potential, Progress, and Challenges', *Neurology International*, vol. 16, no. 4, pp. 805-820, DOI:10.3390/neurolint16040060

Schaewitz, L, Winter, S, and Kramer, NC, 'The Influence of privacy control options on the evaluation and user acceptance of mobile applications for volunteers in crisis situations', *Behaviour & Information Technology*, vol. 40, no. 8, pp. 759-775, DOI:10.1080/0144929X.2020.1723703.

Shockley, KR 2016, 'Quantitative High-Throughput Screening Data Analytics: Challenges and Recent Advances', *Drug Discovery Today*, vol. 20, no. 3, pp. 296-300, DOI: 10.1016/j.drudis.2014.10.005.

Smith, H 2021, 'Clinical AI: Opacity, Accountability, Responsibility and Liability', *AI & Society*, vol. 36, pp. 535-545, DOI:10.1007/s00146-020-01019-6.

Tan, J, and Zhang, Y 2023, 'ExplainableFold: Understanding AlphaFold Prediction with Explainable AI', *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 2166-2176, DOI: 10.1145/3580305.3599337.

Tang, X 2020, 'The Role of Artificial Intelligence in Medical Imaging Research', *BJR*, vol. 2, no. 1, accessed 6 November 2024, <https://academic.oup.com/bjro/article/2/1/20190031/7240329?login=false>.

Tetlock, PE 1991, 'An Alternative Metaphor in the Study of Judgment and Choice: People as Politicians', *Theory & Psychology*, vol. 1, no. 4, pp. 451-475, DOI:10.1177/0959354391014004.

The Lancet Global Health 2024, 'The Landscape for Rare Diseases in 2024', vol. 12, no. 3, accessed 5 November 2024 <https://www.thelancet.com/journals/langlo/article/PIIS2214-109X(24)00056-1/fulltext>.

Tilala, MH, Chenchala, PK, Choppadandi, A, Kaur, J, Naguri, S, aoji, R, and Devaguptapu, B 2024, 'Ethical Considerations in the Use of Artificial Intelligence and Machine Learning in Health Care: A Comprehensive Review', *Cureus*, vol. 16, no. 6, accessed 2 November 2024, <https://pmc.ncbi.nlm.nih.gov/articles/PMC11249277/>.

Trager, TK, Tuting, C, and Kastritis, PL 2024, 'The Human Touch: Utilizing AlphaFold 3 to Analyze Structures of Endogenous Metabolons', *Structure*, vol. 32, accessed 6 November 2024, <https://www.cell.com/structure/pdfExtended/S0969-2126(24)00335-6>.

Tsamados, A, Aggarwal, N, Cowls, J, Morley, J, Roberts, H, Taddeo, M, and Floridi, L 2022, 'The Ethics of Algorithms: Key Problems and Solutions', *AI & Society*, vol. 37, pp. 215–230, DOI:10.1007/s00146-021-01154-8.

Unal, AB, Pfeifer, N, and Akgun, M 2024, 'A Privacy-preserving Approach for Cloud-based Protein Fold Recognition', *Patterns*, vol. 5, accessed 8 November 2024, <https://www.cell.com/patterns/pdfExtended/S2666-3899(24)00156-9>.

United States Food and Drug Administration 2023, *Using Artificial Intelligence & Machine Learning in the Development of Drug & Biological Products*, United States of America.

Varadi, M, Anyango, S, Deshpande, M, Nair, S, Natassia, C, Yordanova, G, Yuan, D, Stroe, O, Wood, G, Laydon, A, Zidek, A, Green, T, Tunyasuvunakool, K, Petersen, S, Jumper, J, Clancy, E, Green, R, Vora, A, Lutfi, M, Figurnov, M, Cowie, A, Hobbs, N, Kohli, P, Kleywegt, G, Birney, E, Hassabis, D, and Velankar, S 2022, 'AlphaFold Protein Structure Database: Massively Expanding the Structural Coverage of Protein-sequence Space with High-accuracy Models', *Nucleic Acids Research*, vol. 50, no. D1, pp. D439-D444, DOI:10.1093/nar/gkab1061.

Varadi, M, Bertoni, D, Magana, P, Paramval, U, Pidruchna, I, Radhakrishnan, M, Tsenkov, M, Nair, S, Mirdita, M, Yeo, J, Kovalevskiy, O, Tunyasuvunakool, K, Laydon, A, Zidek, A, Tomlinson, H, Hariharan, D, Abrahamson, J, Green, T, Jumper, J, Birney, E, Steinegger, M, Hassabis, D, and Velankar, S 2024, 'AlphaFold Protein Structure Database in 2024: Providing Structure Coverage for over 214 Million Protein Sequences', *Nucleic Acids Research*, vol. 52, no. 1D, pp. D368-375, DOI:10.1093/nar/gkad1011.

Vijayan, RSK, Kihlberg, J, Cross, JB, and Poongavanam, V 2022, 'Enhancing Preclinical Drug Discovery with Artificial Intelligence', *Drug Discovery Today*, vol. 27, no. 4, pp. 967-984, DOI:10.1016/j.drudis.2021.11.023.

Vokinger, KN, and Gasser, U 2021, 'Regulating AI in Medicine in the United States and Europe', *Nature Machine Intelligence*, vol. 3, pp. 738-739, DOI:10.1038/s42256-021-00386-z

Vora, LK, Gholap, AD, Jetha, K, Thakur, RRS, Solanki, HK, and Chavda, VP 2023, 'Artificial Intelligence in Pharmaceutical Technology and Drug Delivery Design', *Pharmaceutics*, vol. 15, no. 7, accessed 8 November 2024, <https://www.mdpi.com/1999-4923/15/7/1916>.

World Health Organization 2020, *Operationalising the COVID-19 Technology Access Pool (C-TAP)*, The United Nations, accessed 2 November 2024, <https://cdn.who.int/media/docs/default-source/essential-medicines/intellectual-property/who-covid-19-tech-access-tool-c-tap.pdf?sfvrsn=1695cf9_36&download=true>.

Xie, S, Yu, Z, and Lv, Z 2021, 'Multi-Disease Prediction Based on Deep Learning: A Survey', *Computer Modeling in Engineering & Sciences*, vol. 128, no. 2, pp. 489-522, DOI:10.32604/cmes.2021.016728.

Yang, G, Cintina, I, Pariser, A, Oehrlein, E, Sullivan, J, and Kennedy, A 2022, 'The National Economic Burden of Rare Disease in the United States in 2019', *Orphanet Journal of Rare Diseases*, vol. 17, no. 1, accessed 5 November 2024 <https://ojrd.biomedcentral.com/articles/10.1186/s13023-022-02299-5>.

Zhang, Y, Liao QV and Bellamy, RKE 2020 'Effect of Confidence and Explanation on Accuracy and Trust Calibration in AI-Assisted Decision Making' , paper presented at the Conference on Fairness, Accountability and Transparency, Association for Computing Machinery, Barcelona, 27th-30th January.