Appendix 2

Expert Interview Questionnaire

1. What were your initial impressions of the overall functionality and usability of each AI-driven no-code application you tested?

2. Do you believe that the applications tested are capable of offering reliable and accurate results for drug discovery professionals? Please provide specific examples or scenarios to support your reasoning.

3. What are the main areas where these applications could be improved to better support drug discovery efforts, particularly in terms of performance metrics like speed, accuracy, and resource utilisation?

4. Do you have any additional comments or suggestions regarding the use and potential integration of AI-driven no-code applications in drug discovery that were not covered in the previous questions?

Scenario 3: Drug Design

(Used to test the functionality of apps such as Insilico Medicine's Pharma.AI, Atomwise, Schrödinger's LiveDesign, Exscientia, BenevolentAI, Cyclica)

Objective: Evaluate the application's effectiveness in designing new drug molecules optimised for specific properties such as binding affinity, solubility, bioavailability, and selectivity.

Description:

Target Protein Selection:

• HIV-1 Protease: A crucial enzyme in the life cycle of the HIV virus, making it a prime target for antiretroviral drugs.

• PD-L1 (Programmed Death-Ligand 1): A target in immuno-oncology therapies, where inhibiting the PD-1/PD-L1 interaction can enhance the immune response against tumors.

• ACE2 (Angiotensin-Converting Enzyme 2): A receptor for SARS-CoV-2, making it a target for designing antiviral drugs against COVID-19.

Design Objectives:

• *Binding Affinity:* The application should design molecules with strong binding affinity to the selected target protein.

• *Solubility:* The designed molecules should have high solubility in aqueous environments to ensure adequate bioavailability.

• *Bioavailability:* The molecules should be optimised for oral bioavailability, considering factors such as molecular weight, lipophilicity, and hydrogen bonding potential.

• *Selectivity:* The application should design molecules that are selective for the target protein, minimising off-target interactions that could lead to side effects.

Assessment Criteria:

Lipinski's Rule of Five: The designed molecules are to be assessed for drug-likeness based on Lipinski's Rule of Five, which includes criteria like:

• No more than 5 hydrogen bond donors (OH and NH groups).

• No more than 10 hydrogen bond acceptors (N and O atoms).

• A molecular weight under 500 Da.

• A partition coefficient (LogP) less than 5.

Additional Optimisation Goals:

Pharmacokinetics: The application should also consider pharmacokinetic properties such as half-life and metabolism.

Toxicity: The molecules should be designed to avoid known toxicophores and minimise predicted toxicity.

Evaluation Metrics:

Design Quality:

Definition: The degree to which the designed molecules meet the desired properties (binding affinity, solubility, bioavailability, selectivity).

Creativity:

Definition: The novelty of the molecular structures proposed by the application.

Optimisation Efficiency:

Definition: The time and computational resources required to generate optimised drug candidates.

Scenario 4: Lead Optimisation

(Used to test the functionality of apps such as Insilico Medicine's Pharma.AI, Atomwise, Schrödinger's LiveDesign, Exscientia, BenevolentAI, Cyclica)

Objective: Test the application's ability to refine lead compounds to enhance their efficacy, selectivity, and safety.

Description:

Selection of Target Protein and Lead Compounds:

Target Protein: VEGFR-2 (Vascular Endothelial Growth Factor Receptor-2)

Involved in angiogenesis, making it a target for cancer therapies, particularly in inhibiting tumor blood supply.

Lead Compound Example: Sunitinib

A multi-targeted receptor tyrosine kinase (RTK) inhibitor that has moderate activity against VEGFR-2 but requires optimisation to improve selectivity and reduce cardiotoxicity.

Target Protein: BACE1 (Beta-Secretase 1)

Involved in the production of amyloid-beta, making it a target for Alzheimer's disease.

Lead Compound Example: LY2886721

A BACE1 inhibitor that showed promise in early trials but was discontinued due to liver toxicity.

Target Protein: MMP-9 (Matrix Metalloproteinase-9)

Involved in tissue remodeling and implicated in cancer metastasis and chronic inflammatory diseases.

Lead Compound Example: Marimastat

A broad-spectrum MMP inhibitor that was discontinued due to musculoskeletal toxicity, requiring selective optimisation.

Optimisation Objectives:

Potency Improvement:

Increase the binding affinity of the lead compounds to their respective target proteins. For example, optimising Sunitinib to achieve stronger inhibition of VEGFR-2.

Selectivity Enhancement:

Reduce off-target interactions that could lead to side effects. For instance, modifying the structure of LY2886721 to reduce its interaction with liver enzymes, thus lowering hepatotoxicity.

Safety Improvement:

Optimise pharmacokinetic properties, including absorption, distribution, metabolism, and excretion (ADME), to enhance the safety profile. For Marimastat, structural changes could be explored to reduce its impact on musculoskeletal tissues.

Lead Compound Optimisation Process:

Structure-Based Drug Design:

The application will use the 3D structures of the lead compounds bound to their target proteins to identify regions of the molecule that can be modified to improve binding affinity and selectivity.

Computational Screening:

The application will generate and virtually screen multiple analogues of the lead compound, predicting their binding affinities, selectivity profiles, and ADME characteristics.

Iterative Optimisation:

The application will iteratively refine the lead compounds, using feedback from each round of screening to enhance efficacy and reduce potential toxicity.

Evaluation Metrics:

Improvement in Potency:

Definition: The increase in binding affinity or activity compared to the original lead compounds.

Selectivity:

Definition: The ability to reduce off-target interactions while maintaining or improving efficacy.

Safety Profiles:

Definition: Predicted improvements in pharmacokinetic properties and reduced toxicity.

Scenario 5: Toxicity Prediction

(Used to test the functionality of apps such as Insilico Medicine's Pharma.AI, Schrödinger's LiveDesign, Exscientia, Cyclica)

Objective: Assess the application's capability to predict potential toxic effects of drug candidates, focusing on common toxicity endpoints such as hepatotoxicity, cardiotoxicity, and genotoxicity.

Description:

Selection of Drug Candidates:

Candidate 1: Acetaminophen (Paracetamol) Derivatives

Potential Toxicity: Hepatotoxicity

Background: Acetaminophen is widely used as an analgesic and antipyretic, but overdoses can cause severe liver damage. The application will be tasked with predicting hepatotoxicity in novel acetaminophen derivatives.

Candidate 2: Doxorubicin Analogues

Potential Toxicity: Cardiotoxicity

Background: Doxorubicin, an anthracycline antibiotic used in chemotherapy, is known for its effectiveness but also for its dose-dependent cardiotoxicity. The application will predict cardiotoxic effects in new analogues.

Candidate 3: Nitrosourea-Based Compounds

Potential Toxicity: Genotoxicity

Background: Nitrosoureas are alkylating agents used in cancer treatment, but they can cause DNA damage leading to mutagenesis and genotoxicity. The application will assess the genotoxic potential of new nitrosourea derivatives.

Toxicity Endpoints:

• *Hepatotoxicity:* Predict the likelihood of liver enzyme elevation (e.g., ALT, AST) and liver damage, which could lead to conditions like drug-induced liver injury (DILI).

• *Cardiotoxicity:* Predict the potential for QT interval prolongation, cardiomyopathy, or arrhythmias, which could lead to heart failure or sudden cardiac death.

• *Genotoxicity:* Predict the potential for DNA damage, mutagenesis, and chromosomal aberrations, which could result in carcinogenicity or teratogenic effects.

Toxicity Prediction Process:

Data Input:

The application is to be provided with the chemical structures and relevant physicochemical properties of the drug candidates.

Prediction Models:

The application will use in silico models to predict toxicological endpoints based on known toxicity data, structure-activity relationships (SAR), and machine learning algorithms trained on large toxicity datasets (e.g., Tox21, ToxCast).

Output:

The application will generate a toxicity profile for each drug candidate, indicating the likelihood of various toxic effects, such as hepatotoxicity, cardiotoxicity, and genotoxicity.

Evaluation Metrics:

Accuracy:

Definition: The percentage of correctly predicted toxic and non-toxic compounds compared to experimental data.

Definition:

Sensitivity: The ability to correctly identify compounds that are toxic.

Specificity: The ability to correctly identify compounds that are non-toxic.

Predictive Power:

Definition: The application's ability to predict specific types of toxicity with high confidence.