GPX4 Drug Development Pipeline

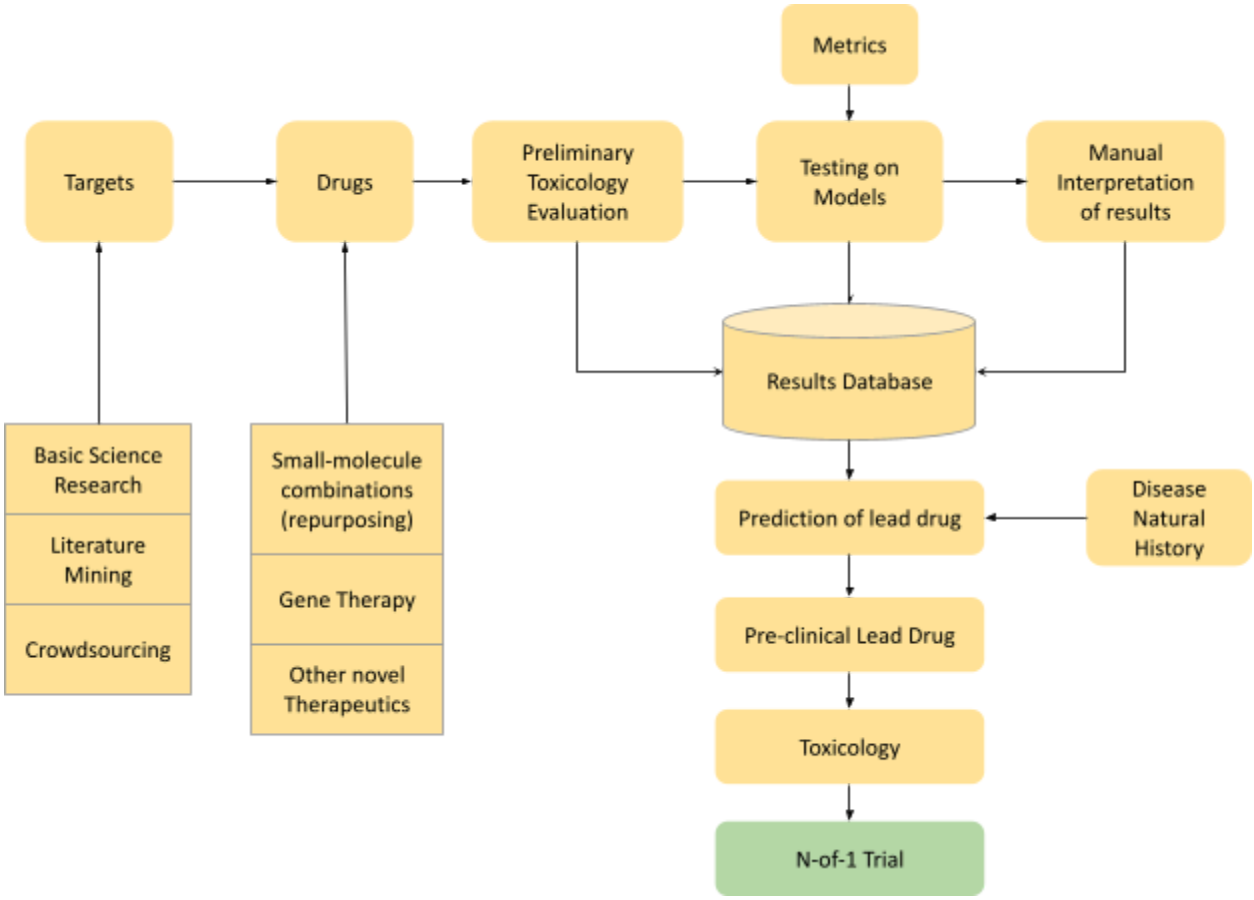
We need to intervene with a drug in the next 6-12 month to create meaningful improvements in the kid's quality of life. But drug development is expensive, time consuming and has a low probability of success (DiMasi et al). The average time from starting a pipeline to dosing the first human patient with a new drug is 4-7yrs for a new small molecule drug, 2-4yrs for a new gene therapy drug, and 1-2yrs for a new anti-sense oligonucleotide (ASO) drug. To meet our aggressive timelines, we apply the following principles to simplify the problem and also potentially increase the probability of success:

- **Incremental over Big-Bang**: Instead of investing several years creating one highly effective drug, we want to quickly discover a drug with reasonable efficacy, dose our patients, and buy us more time as we iterate to improve the drug's efficacy.
- **Treatment over Intellectual Property:** Drug developers go after novel technologies or molecules that could be patented. Given the goal of finding a treatment, we are completely okay with repurposing existing drugs, using naturally occurring substances, utilizing existing patented technology, or using generics. This opens up the door to new possibilities not available to commercial drug developers.
- **N-of-1**: Clinical trials contribute a significant portion to the cost and duration of drug development. In a rare disease with just four patients, clinical trials are not practical. Therefore we are adopting the N-of-1 drug development model where the risk of dosing patients with an experimental drug is acceptable in comparison to the damage caused by the disease during the wait for a perfect trial.
- **Reduce Time over Money**: Given two choices of activities, we prefer to execute the faster, more expensive activity over the cheaper one. It does not necessarily mean our pipeline is very expensive to operate. In fact, if we draft a plan with a billion dollar budget, we are just as likely to fail as a plan with a million dollars. On the contrary, activities that are quick to execute tend to be small and cheap. By choosing to reduce time over money, we not only find treatment faster but also potentially cheaper.
- **Early Failures over Early Success**: Drug development pipelines follow a multi-step "waterfall" approach of starting with a library of tens of thousands of molecules and identifying top performers by testing on models of increasing biological complexity (single cell -> tissues -> worm -> fish -> mouse). Companies spend several months designing the perfect molecule and a pipeline upfront because failures can cost thousands if not millions of dollars. In a novel rare disease, we don't know enough about the disease to design the perfect pipeline upfront. Instead of trying to prevent failures, we assume failure is inevitable in all our activities. We will design the pipeline to fail fast, maximize the learning from failures, and fail often to generate enough insights to eventually lead to a successful drug. Successful drug is one that provides a meaningful improvement to the patient's quality of life.

Pre-clinical Pipeline Design

Only goal of the pipeline is to identify drugs that provides meaningful improvement to patient's quality of life. *Without dosing human patients, how can we accurately predict whether a drug will improve patient's quality of life?* This is like asking how to accurately predict the future. We can never accurately predict the future, except we can get tolerably accurate in certain well-understood systems, such as weather forecasting. We are designing this pipeline to maximize our understanding of the disease biology using non-human models. We hope this understanding will lead to a more accurate prediction of clinical outcome of a drug.

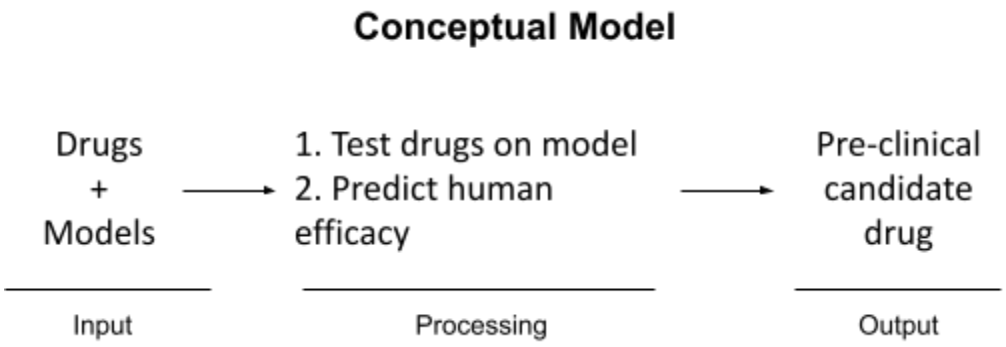
The following diagram illustrates different stages of the pipeline and their relationships.



Conceptual Model: Input-Processing-Output

Conceptually, a drug development pipeline is a simple Input-Processing-Output system. Input is a library of drugs (including small molecules, gene therapy, antisense oligonucleotides etc) and disease models (including cells, iPSCs, animal models etc). Processing is testing each drug on the models and selecting the one drug with maximum potential to improve patient's life. Output is the drug identified by processing. Although simplified, this conceptual model helped us

identify the four key variables we can tweak to improve the outcome of the pipeline: Drug library, Disease Models, Testing, Prediction.



Drug library

Success of the pipeline to identify a lead drug is limited by the selection of drugs in this library. We want a scientifically rigorous criteria to include a drug in the library, but at the same time, we want to leave room for exploration. We will start the library with a set of small selection of FDA approved small molecule drugs to repurpose, add combinations of approved drugs or novel molecules and include drugs based on gene therapy or other novel technologies.

Disease Models

We want models that accurately represent the human disease in the biological system or process they represent ex: biochemical, cellular, whole organism etc. We also want models to be sensitive enough to show a measurable difference when intervened with a drug. We will be creating the following models with the goal of covering as many biological systems as possible:

- GPX4 recombinant protein
- CRISPR created GPX4 variant in reference cell line
- Patient-derived fibroblasts
- Patient-derived iPS Cell lines
- Brain and Bone organoids derived from patient-derived iPSCs
- GPX4 mutant worm model
- GPX4 mutant zebrafish model
- GPX4 conditional complete knockout mice
- GPX4 transgenic mice

Testing

For genetic conditions, animal models are built by recreating the genetic variation in the animal's genome or silencing the gene entirely. These are good approximations of the human condition but seldom sufficient to predict the clinical outcome of a drug. Some might argue that patient derived cells, fibroblasts or iPSCs are good predictors of clinical outcome. This might be true in

hindsight, but there is no way to determine the ideal model a priori. Based on this observation, we will test every drug on at least three different models.

Metrics
The measurements we observe on the model as just as important as the model itself. In addition to standard measurements for respective model, will measure biomarkers known to be relevant in humans for this condition.

Calibration
Drug development pipelines typically select drugs based on the drug's ability to restore wildtype phenotype to a disease model. *Would the same drug also restore phenotype in human patients*? It is impossible to predict. We can make an educated guess if we had existing data about transferability of model observations to humans. To establish this baseline data, we will test a panel of drugs known to produce positive, negative and neutral impact on healthy humans. In addition to positive phenotype changes, we will prefer drugs that produce a positive activity compared to the baseline.

Preliminary Toxicology Evaluation
Why test a drug if it is toxic? We will establish a series of simple, low cost, fast turnaround toxicology assays to study understand toxicity. Drugs that pass this stage will be tested on models.

Results Database
Disease models are, at best, a good approximation of the human condition. Every drug we test on the model offers a different insight into the disease. We will store the results of every test in a database in a standardized format. It allows us to analyze the data independent of the tests. We will make parts of the database public to allow the data scientists and research community to search for hidden.

Prediction

Goal of the prediction step is to identify one or two drugs from the results database that can significantly improve the quality of life of patients. There is no good algorithm to predict patient outcome. We will try a variety of known statistical analysis and data mining techniques to identify an appropriate pre-clinical lead drug. We will also anonymize the data and make it public for any data scientist to re-analyze the data using different algorithms or by augmenting with different datasets. We will continue with toxicology study and N-of-1 clinical trial based on the lead drug identified by us or someone in the public.