Insight Brief

HINT: Hierarchical Interaction Network for Clinical Trial Outcome Prediction

CAO XIAO, Senior Director, Amplitude
KEXIN HUANG, Research Student, IQVIA and PhD Student, Stanford University
TIANFAN FU, Department of Computational Science and Engineering, Georgia Institute of Technology
LUCAS GLASS, Vice President, Analytics Center of Excellence, IQVIA
JIMENG SUN, Chief AI Scientist, IQVIA and Professor, University of Illinois Champaign Urban

Introduction

The global clinical trial market has reached $44.3 billion in 2020 and is expected to grow to $69.3 billion by 2028. The costs of conducting clinical trials are extremely expensive, and the time to run a trial takes multiple years with a low probability of success.

Clinical trials are crucial for drug development but often face uncertain outcomes due to safety, efficacy, or patient recruitment problems. In a scientific paper recently published in *Cell Patterns*, IQVIA experts Lucas Glass and Jimeng Sun proposed using a Hierarchical INTeraction network (HINT) to predict clinical trial outcomes. A HINT model first encodes multi-modal data, such as the drug molecule, target disease and trial eligibility criteria, into embeddings, then trains knowledge embedding modules using drug pharmaco-kinetic data and historical trial data. Finally, the hierarchical interaction graph connects all the embeddings to capture their interactions and predict the trial outcome. HINT was trained and validated on 1,160 Phase I trials, 4,449 Phase II trials and 3,436 Phase III trials. It significantly outperforms the best baseline method on most metrics (benchmark dataset and codes are released at https://github.com/futianfan/clinical-trial-outcome-prediction.

Can we predict trial outcomes in an in silico manner?

Vast amounts of historical clinical trial data and massive knowledge bases on passed and failed drugs bring a new opportunity for using machine learning models to tackle the critical question: **Can one predict the trial outcome before the trial starts?** We at IQVIA are among the first to study the general trial outcome prediction problem across different trial phases for multiple diseases.

Various public data sources can provide vital information for predicting trial outcome. For example, the clinicaltrials.gov database has 369.7K historical clinical trials with valuable information to mine. In addition to this database, IQVIA's team utilized the standard medical codes of the diseases and their natural language description through the National Institutes of Health website. Another source, DrugBank Database, contains the biochemical description of many drugs, which allows for the computational modeling of drug molecules.

There is great interest in developing a general method for trial outcome prediction. Previous research attempted to predict drug approvals based on drug and clinical trial features using classical machine learning methods. Several limitations impeded the utility of those existing trial outcome prediction models including:

- **Lack of benchmark data.** Large datasets in the clinical trial domain are often not available, which severely affects data science efforts on clinical trial-related research.
- **Limited task definition and study scope.** Existing works either focus on predicting individual components of trials, such as patient-trial matching or only a subset of disease groups. Although helpful for part of the trial design, they do not predict the trial outcome for a broad set of target diseases.
- **Limited features used for prediction.** Due to their limited task definition and study scope, existing works often only leverage restricted disease-specific parts, which cannot be generalized for other diseases.
- **Ignoring the complex relations among trial components.** Due to the limited data and task scope, existing methods often simplify their predictions by limited input features and rely on classical classification methods not explicitly designed for modeling the interaction of different trial components.

Our approach:

To provide accurate trial outcome predictions for all trials, we proposed a HINT model that is trained on a multi-modal dataset, including molecule information of the drugs, the target disease information, the trial eligibility criteria, and biomedical knowledge. HINT first encodes these multi-modal data into latent embedding vectors of the drug molecule, the target disease, and the trial risk, where an imputation module is designed to handle missing data.

Next, we trained a knowledge embedding module from external knowledge on pharmaco-kinetic properties for improving drug embedding. We also trained another trial risk embedding module using historical trial data for improving trial risk embedding and presented an interaction graph module to connect all the embeddings to capture various interaction effects from different trial components. Finally, HINT learned a dynamic attentive graph neural network to predict trial outcomes.

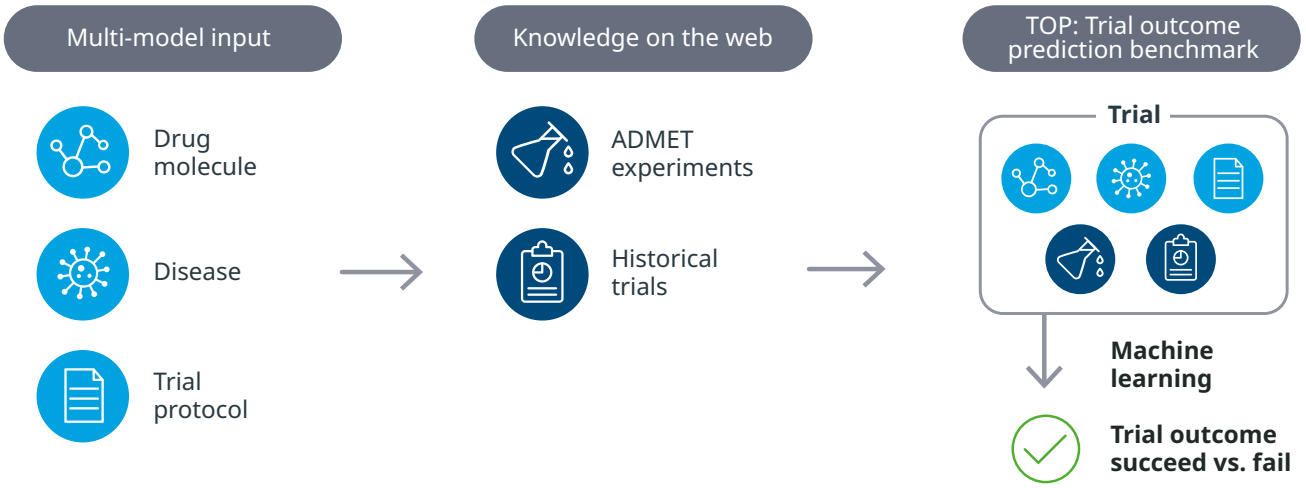
Our main contributions:

- **Problem** - We formally defined a model framework for a general clinical trial outcome prediction task, which not only models various trial risks, including drug safety, treatment efficiency, and trial recruitment, but also models a wide range of drugs and indications (e.g., diseases). Our model framework generalized over new trials given the drug molecule, target disease, and trial eligibility criteria.

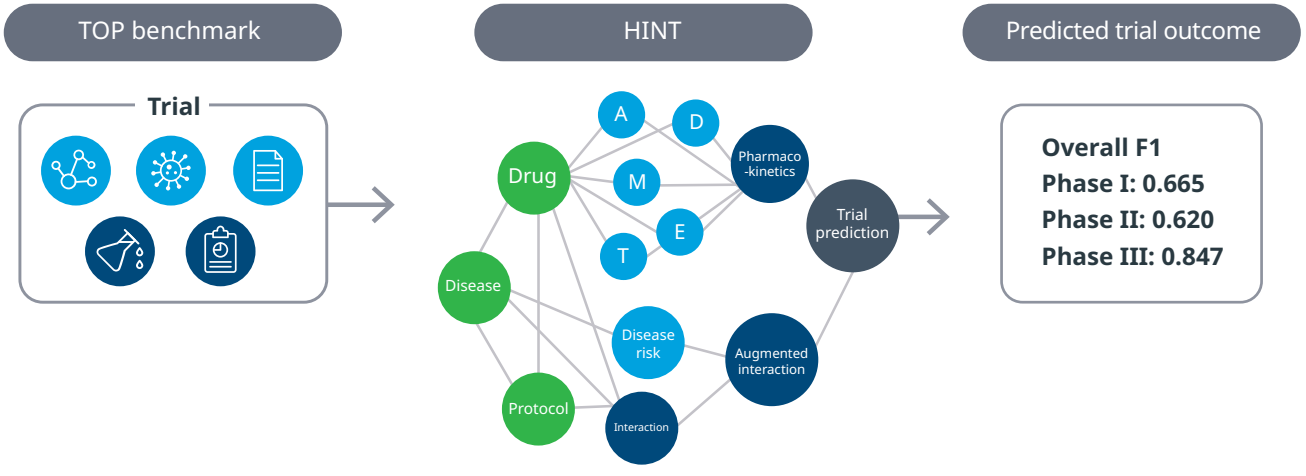
- **Benchmark** - To enable general clinical trial outcome predictions, we leveraged a comprehensive set of datasets, including DrugBank, disease codes, and clinical trial records to curate a Trial Outcome Prediction dataset (TOP).
- **Results** - We evaluated HINT against state-of-the-art baselines using real-world data. HINT achieved statistically significant improvement compared with the best baseline method (COMPOSE). We also conducted an ablation study to evaluate the importance of key components of clinical trials to the prediction power and the effectiveness of the hierarchical formulation of a trial interaction graph. Finally, we conducted a case study to show the potential real-world impact of HINT by successfully predicting some prominent trial outcome.
- **Method** - We designed a graph neural network method that explicitly simulates different clinical trial components and their interaction relations for predicting trial outcome.

We focused on clinical trials that aim at discovering new indications of drug candidates. The trials that did not involve drug molecules, such as surgery techniques and medical devices, were out of the scope.

Curating Clinical Trial Outcome Prediction (TOP) Benchmark



Curating Clinical Trial Outcome Prediction (TOP) Benchmark



BENCHMARK DATASET

To standardize the clinical trial outcome prediction, we created a benchmark dataset for Trial Outcome Prediction (TOP), which incorporates rich data components about clinical trials, including drug, disease and eligibility criteria. We first described the data components and then reported the processing steps to construct this benchmark dataset.

For each clinical trial, we produced four data items:

1. **Drug molecule information**, including SMILES strings and molecular graphs for the drug candidates used in the trials.
2. **Disease information**, including ICD-10 codes (disease codes), disease descriptions and disease hierarchy in terms of CCS codes.
3. **Trial eligibility criteria** in unstructured natural language that contains inclusion and exclusion criteria.
4. **Trial outcome information** including an indicator of trial success or failure, trial phase, start and end date, sponsor and trial size. Auxiliary datasets were provided: pharmacokinetics (PK) data, which consists of wet lab experiment results, and the drug SMILES strings, provided in MoleculeNet. Disease risk data contains the historical success rate of the target disease and the disease descriptions, from ClinicalTrials.gov.

DATA CURATION PROCESS

We created the TOP benchmark for trial outcome prediction from multiple data sources, including drug knowledge bases, disease codes (ICD-10 code), historical clinical trials and manually curated trial outcome labels. We applied a series of selection filters to ensure the chosen trials had high-quality outcome labels and started with 369.7K raw clinical trial records from ClinicalTrials.gov.

RESULTS:

We compared HINT with several baseline approaches, covering conventional machine learning models and deep learning-based models. HINT was evaluated on different disease groups, including neoplasm (oncology/cancer/tumor), respiratory system disease, digestive system disease and nervous system disease. We also evaluated the prediction performance on high/low prevalence disease trials, finding that low prevalence disease-related trials still achieved good prediction performance, validating the generalization of the proposed method.

We used a missing data imputation module to estimate molecule embedding. HINT with imputation outperformed the other methods due to its ability to impute missing molecule embeddings. We concluded that the trials with unseen disease codes (not in the training set) are more likely to be mis-predicted by HINT, which is more likely to happen on rare diseases.

Trial Outcome Prediction

Existing works often focus on predicting individual patient outcomes in a trial instead of a general prediction about the overall trial success. They predict at the patient level, whereas HINT focuses on the trial level. Some evaluate various conventional machine learning models for clinical trial outcome prediction, however, they do not leverage rich trial features, such as drug molecules, trial eligibility criteria. By contrast, HINT takes multi-modal data sources into account.

LIMITATIONS:

Supporting more trial types: HINT can handle interventional trials involving small molecules. Other trial types, such as medical devices and biologics trials are not covered by the current model due to the molecule encoding.

Supporting rare diseases: Like any machine learning model, HINT requires sufficient training data to train accurate predictive models. However, low prevalence diseases, especially rare diseases, are difficult to handle due to the lack of sufficient historical trials as the training data for HINT.

Enhancement of model interpretability: HINT is a graph neural network model, which integrates comprehensive data sources to predict trial outcomes. Due to the complex interaction patterns, it can be difficult to explain those predictions.

Trial outcome labels: HINT assumes a simple binary label of success or not. However, there might be more granular classes of trial outcomes, especially for failed trials. It would be more useful if the model could classify the trial to more specific failure reasons. Trial publications are often disproportionally skewed towards successful trials, although the detailed explanation of failed trials can benefit future trials.

Conclusion

In the complete scientific paper, a machine learning benchmark for trial outcome prediction was created. We designed a graph neural network-based method, HINT, to leverage multi-sourced data and incorporate multiple factors in a hierarchical interaction graph for predicting trial outcome. Empirical studies indicate that HINT outperforms multiple baseline methods in several prediction metrics on phase-level trial outcome prediction. Future works include expanding to other trial types beyond intervention trials of small molecules and expanding the binary trial outcome labels. Ultimately, it was learned that HINT can predict the success probability of the trial accurately, reassuring the drug developers for the prospect of the treatment.

While HINT is not yet fully integrated into IQVIA's solutions, IQVIA's Asset Intelligence tool utilizes AI/ML to predict clinical trial success and provide enhanced, objective exploration of existing clinical development pathways. This solution allows you to screen therapies via our extensive knowledge base, enabling you to identify a shortlist of drugs and understand their probability for technical and regulatory success.

* References can be found in the complete scientific paper featured in the open access journal Patterns.

CONTACT US

iqvia.com