Animal-GAN: A Generative AI Alternative to Animal Clinical Pathology Testing

U.S. FOOD & DRUG ADMINISTRATION
NATIONAL CENTER FOR TOXICOLOGICAL RESEARCH

Xi Chen[1], Auerbach, Scott S [2], Weida Tong[1*], Zhichao Liu[1*]

[1] Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, Food and Drug Administration, Jefferson, AR, USA
[2] Division of the National Toxicology Program, National Institute of Environmental Health Sciences, Durham, NC, USA

Weida.Tong@fda.hhs.gov
Zhichao.Liu@fda.hhs.gov

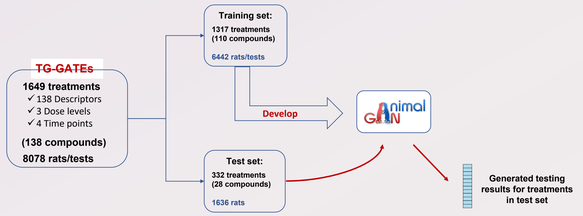
Abstract

Clinical pathology testing of laboratory animals is crucial for chemical safety evaluation and risk assessment. As the toxicology community and regulatory agencies are moving towards replacement, reduction, and refinement (3Rs) of animal studies, and meanwhile, abundant animal data are available from the public domain, we are exploring an Artificial Intelligence (AI) approach to learn from the existing animal testing results to generate the new animal data without conducting additional animal experiments. The AnimalGAN was developed based on a generative adversarial network (GAN)-based framework. Specifically, we constructed AnimalGAN by learning clinical pathology data from multiple experimental treatments (i.e., compound/time/dose combinations) available in TG-GATEs. We also challenged the AnimalGAN with data from DrugMatrix. AnimalGAN successfully inferred hematological and biochemical parameters with high similarity (0.998±0.002) to the corresponding animal testing values under the same experimental design. Furthermore, the generated hematological and biochemical parameters by AnimalGAN could yield similar toxicity identification results comparable with those from animal samples, with an average concordance rate over 90%. Moreover, we asked AnimalGAN to generate clinical pathology data for treatments reported in DrugMatrix. We found that the generated data can be used to assess toxicity as their corresponding animal data did, with an average concordance rate around 70%, which is comparable with the concordance of toxicity assessment results by using TG-GATEs data and DrugMatrix data for their overlapped treatments. The proposed AnimalGAN framework and its applications demonstrated the potential of utilizing advanced AI approaches to produce non-animal models as alternatives to animal studies based on the existing data.

Disclaimer

The information in this poster represents the opinions of the speaker and does not necessarily represent FDA's position or policy.

AnimalGAN Development and Evaluation



AnimalGAN is a generative adversarial network-based framework, involving two networks – generator and discriminator. It receives treatment information, including molecular descriptors, dose levels and time points of compounds, as model condition.

The dataset extracted from the rat repeated dose studies in the open TG-GATEs database was used to develop our AnimalGAN.

We split the dataset into two sets, using about 80 percent, 110 compounds treated rats under 1317 treatment conditions as training set to construct the model. Then we asked the AnimalGAN to generate the clinical pathology testing results for the 332 treatment conditions of the remaining 28 compounds in the test set. Due to those are unseen data in model development process, we can use them to evaluate the performance and generalization of the AnimalGAN.

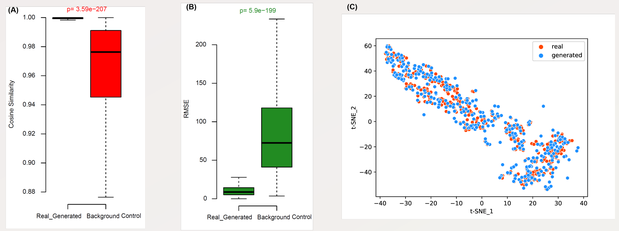
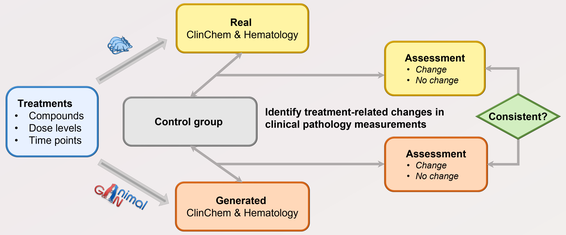


Figure 1. Boxplot of (A) Cosine Similarity, (B) RMSE between generated data and their corresponding animal testing data in the test set, and (C) t-SNE plot of test set.

AnimalGAN-generated data-based Toxicity Assessment



For a given treatment, we compared the difference between the vehicle control-treated group and compound-treated group for each measurement to assess if there is change between the two group. If the test result based on generated data and their corresponding real data were both indicating there is no change between the treated group and control group, or both indicate treatment-related changes, we consider that the generated data and the real data are consistent in toxicity assessment. Otherwise, we consider they are not consistent.

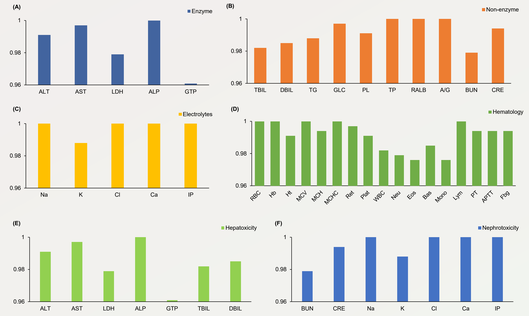
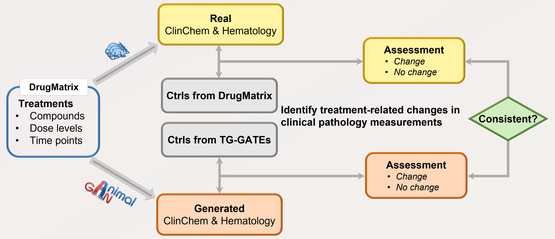


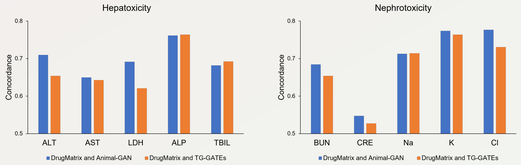
Figure 2. Concordance between Real and Generated Measurements on Toxicity Assessment.

AnimalGAN model generated clinical pathology data can achieve equivalent performance as real animal data in toxicity assessment.

External Validation with DrugMatrix



Due to AnimalGAN was developed using TG-GATEs data, we used the vehicle-control treated group in TG-GATEs to identify treatment-related changes for the generated data.



The generated data can be used to assess toxicity as their corresponding animal data did, with an average concordance rate around 70%, which is comparable with the concordance of toxicity assessment results by using TG-GATEs data and DrugMatrix data for their overlapped treatments.

Summary

- We proposed a GAN-based framework named AnimalGAN, which learns from the existing animal data and holds promise for generating new animal data just from chemical information in combination of dosage and treatment durations.
- Learning from the pre-existing data from the TG-GATEs, AnimalGAN successfully generated the clinical pathology testing results with high similarity to the corresponding animal testing results.
- AnimalGAN generated clinical pathology testing results can be used to assess toxicity, and the assessment results are highly consistent with the animal testing results-based assessment.