Abstract

Enhancing Chemical-Induced Human Carcinogenic Risk Evaluation through Advanced AI Technologies [†]

Nelly Babayan [1,2,*], Helga Stopper [3], Lusine Khondkaryan [1,2], Ani Tevosyan [2,4], Gohar Tadevosyan [1], Lilit Apresyan [1], Lusine Adunts [2] and Zaven Navoyan [2]

[1] Institute of Molecular Biology, National Academy of Sciences, Yerevan 0014, Armenia; l_khondkaryan@mb.sci.am (L.K.); g_tadevosyan@mb.sci.am (G.T.); l_apresyan@mb.sci.am (L.A.)

[2] Toxometris.ai Inc., Glendale, CA 91204, USA; atevosyan@toxometris.ai (A.T.); ladunts@toxometris.ai (L.A.); znavoyan@toxometris.ai (Z.N.)

[3] Institute of Pharmacology and Toxicology, University of Würzburg, 97070 Würzburg, Germany; stopper@toxi.uni-wuerzburg.de

[4] YerevaNN, 0025 Yerevan, Armenia

[*] Correspondence: n_babayan@mb.sci.am

[†] Presented at the 1st International Electronic Conference on Toxics, 20–22 March 2024; Available online: https://sciforum.net/event/IECTO2024.

Keywords: carcinogenicity; in silico; risk-score; small molecules

Introduction: Efficient cancer risk assessment is vital for sustainable practices in pharma, agriculture, and environmental protection. Traditional animal tests for chemical carcinogenicity are time- and cost-consuming. Ongoing efforts focus on pioneering alternative approaches to improve accessibility and reliability in cancer risk assessment.

Objectives: This study aimed to develop a scoring function that can rank chemical compounds based on their potential human carcinogenicity through in silico methods.

Materials and methods: An ensemble of diverse AI/ML models, including Boosting Machines, Graph Neural Networks, and Large Language Models, was used to predict endpoints associated with carcinogenicity, including in vitro mutagenicity, in vitro and in vivo clastogenicity/aneugenicity, and rodent carcinogenicity. A risk score function was developed by applying a weighting strategy to every endpoint. Datasets of human carcinogenic and non-carcinogenic chemicals were used to evaluate the performance of the risk score; the p-value was estimated to indicate the significance of the difference.

Results: It was shown that the mean risk score values differed significantly ($p < 0.0001$) between human carcinogens and non-carcinogens. Human carcinogens were able to be predicted with an accuracy rate of 73%, which was slightly lower than the 76% accuracy achieved in experimental carcinogenicity studies in mice and significantly surpassed the 65% accuracy obtained in studies with rats.

Conclusion: The devised risk score evaluates the potential of chemicals to induce cancer in humans in silico by integrating information from diverse cancer-related test results, providing an approach that is nearly as accurate as in vivo experiments. Due to its speed and efficiency, the approach developed can effectively be employed for screening large quantities of chemicals. The risk score developed focuses on genotoxic carcinogens. It is anticipated to enhance the versatility and applicability of the approach through the inclusion of additional endpoints associated with non-genotoxic carcinogenesis, as well as the implementation of more sophisticated AI/ML technologies, such as multi-task learning.

Author Contributions: Conceptualization, N.B. and Z.N.; methodology, A.T. and L.A. (Lusine Adunts); software, A.T.; validation, L.K. and Z.N.; formal analysis, G.T. and L.A. (Lilit Apresyan); investigation, G.T. and L.A. (Lilit Apresyan); resources, Z.N.; data curation, L.K.; writing—original draft preparation, N.B.; writing—review and editing, H.S.; visualization, A.T.; supervision, N.B. and H.S.; project

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is available from the corresponding author upon request.

Conflicts of Interest: The authors and Toxometris.ai Inc. declare no conflict of interest.