From data exploration to predictive models: advanced Machine Learning and Artificial Intelligence techniques for cardiotoxicity analysis

E.L. Viganò,[1] A. Roncaglioni,[1] D. Ballabio[2]

[1]Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Milan; [2]Milano Bicocca University, Italy

In recent years, cardiovascular toxicity has attracted considerable attention from scientists and clinicians since Cardiovascular Disease (CVD) is one of the leading causes of mortality worldwide. However, except for drugs, the evaluation of the potential cardiotoxic effects of chemicals is poorly addressed and regulated. *In silico* methodologies are rapidly emerging as an essential tool in toxicology and pharmaceutical research. These approaches comprise a series of methodologies, which can play an important role in the reduction, and replacement of *in vivo* experiments that are much more laborious, time-consuming, and expensive. Our work focuses on developing predictive models using state-of-the-art Artificial Intelligence (AI) and Machine Learning (ML) techniques to assess the cardiotoxicity of drugs, pesticides, and industrial products. These models can help identify potential safety concerns early in the drug discovery process. As AI approaches are "data-hungry," we have collected thousands of data points for different molecular initiating events and key events related to cardiotoxicity, following the concept of the Adverse Outcome Pathway (AOP) network.

To effectively represent chemical information for AI/ML, we curated the collected data and conducted an analysis to evaluate different types of chemical representations such as Quantitative Structure-Activity Relationships (QSARs) descriptors, fingerprint, embeddings, graph, and combinations of these in more complex AI architectures. Our aim was to determine the most effective descriptors to represent chemical information in a way that enables high-performance models. Indeed, we tested various AI architectures to consider multiple aspects that could lead to potential toxic effects, such as multimodal and multitask approaches. By considering these different approaches, we aim to develop models that can accurately predict potential hazard cardiotoxic effects of drugs, pesticides, and industrial products.

Correspondence: E.L. Viganò
E-mail: edoardo.vigano@marionegri.it

Conference presentation: this paper was presented at the Fourth Centro 3R Annual Meeting - The role of 3Rs in the age of One Health: where we are and where we're going - 13-15 September 2023, Università degli Studi Milano-Bicocca.

Biomedical Science and Engineering 2023; 4:226
doi:10.4081/bse.2023.226