REVOLUTIONIZING BIOMEDICAL RESEARCH: INTEGRATING CUTTING-EDGE AI/ML TO UNLEASH INNOVATION IN DRUG DISCOVERY AND THERAPEUTICS DEVELOPMENT

Toxicity Read-across Tools in CLEAR for Assessment of Emerging Threat Chemicals

Nathaniel Stemmler SRC Inc. **Stephen Houghton** SRC Inc. **Mario Citra** SRC Inc. **Lauren Cassidy** SRC Inc.

SRC, Inc is working with the Threat Agent Science team in DTRA's Advanced and Emerging Threat Division to develop the Combined Logical Estimation Application for Rapid Results (CLEARR) software tool. CLEARR is a comprehensive chemical decision support tool that can be used before, during, and after an incident to facilitate rapid characterization of known and emerging threat chemicals. The software tool combines both in silico QSAR/QSPR predictive algorithms and a large experimental database to provide a rapid and comprehensive risk, exposure, and hazard assessment using the chemical structure. CLEARR is a component of DTRA's Computational Rapid Identification & Scientific Threat AnaLysis (CRISTAL) program which seeks to reduce the time-consuming and costly approach of generating chemical exposure assessments traditionally involving animal testing development from years to months.

One feature of CLEARR is the ability to conduct read across assessments on ovel emerging threat chemicals or chemicals lacking experimental physicochemical or toxicological data. Analog identification (ID) and read-across strategies can help bridge the gaps in information about these substances. An approach that is often employed uses structural analog identification to identify the "nearest neighbors" with experimental data to fill important data gaps. This method of using data from a close structural analog is typically referred to as a read-across strategy. In the past, read-across relied heavily upon a subject matter expert (SME) to scan a long list of chemicals looking for those that were similar enough to use data as a surrogate for the chemical lacking data. This is an extremely time consuming, somewhat subjective and error prone process. To automate and standardize this process, SRC has implemented an analog identification tool that relies upon fingerprinting methods from the Chemical Developers Kit (CDK).

CLEARR currently uses three fingerprinting methods: Extended; MACCS; Klekota-Roth. In CLEARR, the Endpoint Specific read-Across operates in the following manner: the user submits a chemical, navigates to the analog ID feature. The application encodes the submitted SMILES into a fingerprint and then automatically compares the submitted chemicals fingerprint to a fingerprint library stored the CLEARR database. A list of structural analogs is identified and ranked on similarity score and available experimental data.

In addition to generating structurally similar analogs for individual endpoints, CLEARR can also provide a global list of structurally similar compounds using strictly the structural fingerprints from the entire database (354,000 chemicals) as the discriminator rather than a specific endpoint. This allows the user to see the closest structural analogs from the entire database and all available endpoints.

The application enables read-across strategies in CLEARR by rapidly identifying or generating surrogates that can be utilized to refine experimental pursuits, develop better hazard prediction results with additional inputs, and identify viable medical countermeasures. This is a capability in development but will be incrementally shared on the CLEARR demo server as new features become available.

Special thanks to Dr. Don Cronce of DTRA CBS for funding. Also, thanks to Kristen Hoffman and Sara Peacock also of DTRA, and Kevin Ulmes from DHA for technical discussions. Also, thanks to Dr. Chris Ellis, Dr. Kyle Glover, Dr. Morgan Minyard, and Dr. Michael Feasel of DEVCOM CBC for facilitating the collaboration.