Revolutionizing Drug Safety Assessment via QSAR and q-RASAR-based toxicity prediction to protect human health

Presented

by

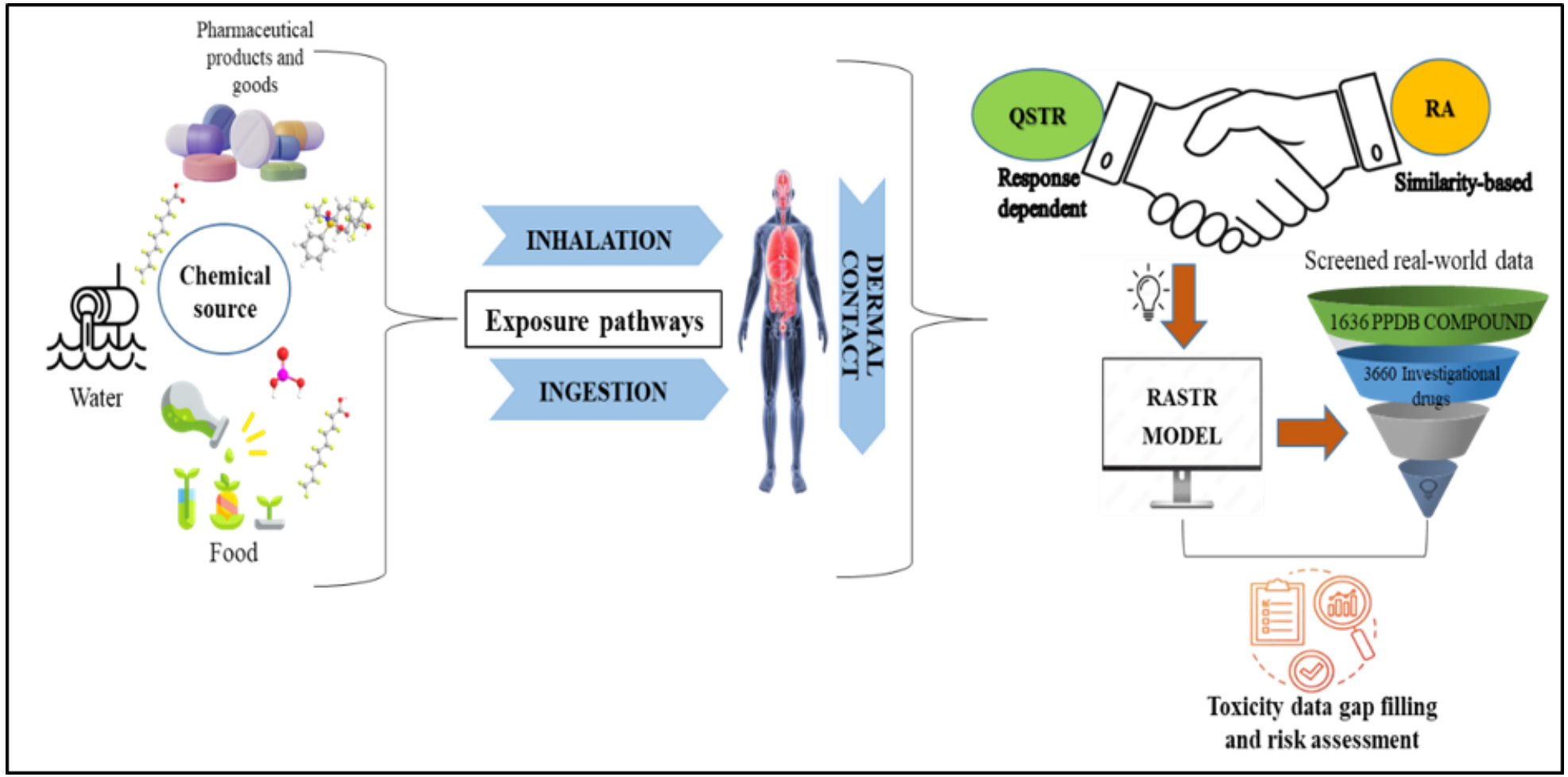
Shubha Das
PhD scholar
Pharmaceutical Chemistry

Under the guidance of

Dr. Probir Kumar Ojha, Associate Professor
Dept. of Pharmaceutical Technology,
Jadavpur University,
Kolkata

2025

Graphical Abstract



Introduction (why this study matters)

- **Massive chemical exposure**

Over 204 million chemicals are registered worldwide, with extensive pharmaceutical and pesticide exposure posing serious threats to human health.

- **Escalating health impacts**

Chemical exposure is strongly linked to endocrine disruption, cancer, neurotoxicity, and multi-organ damage-creating an urgent public health challenge.

- **Drug development failures**

81% of drug candidates fail clinical trials; 33% are terminated specifically due to toxicity, and 90% of market withdrawals result from severe toxicity detected in late stages.

- **Traditional testing limitations**

Conventional methods cost $14 billion annually, use over 100 million animals, show poor human correlation, and, critically, detect toxicity too late after billions are invested.

- **Study objective**

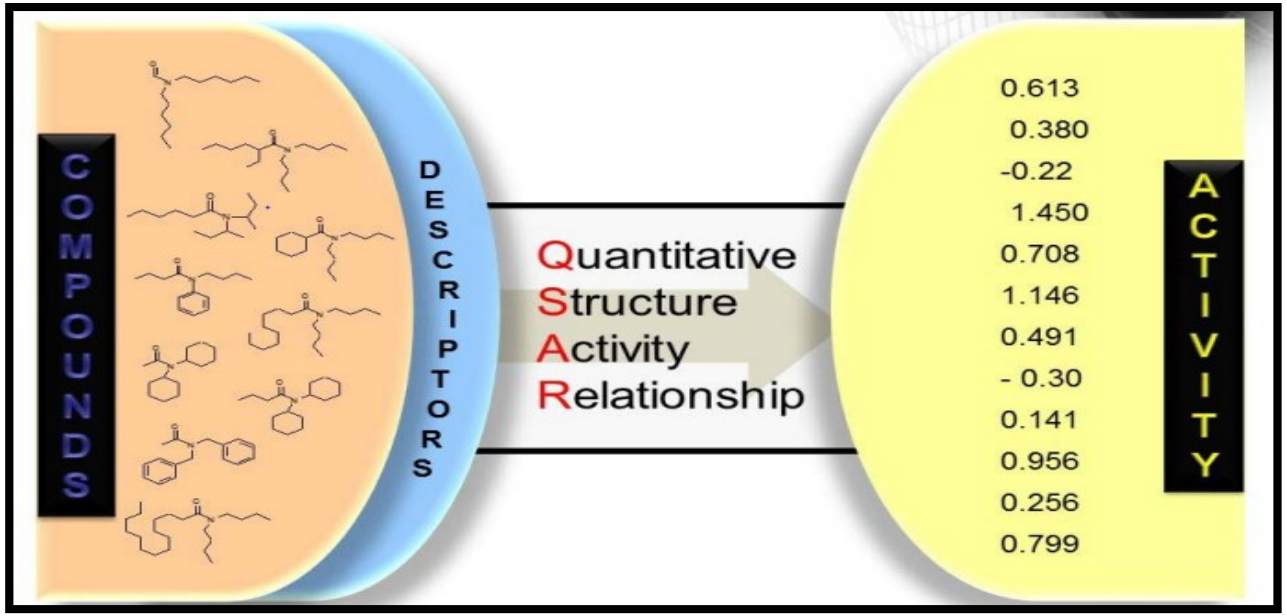
To address this critical data gap, we developed the first-ever chemometric model using the pTDLo endpoint (negative log of Toxic Dose Low) the lowest dose causing any toxic effect in humans, enabling direct human toxicity prediction without animal extrapolation.

QSAR (Quantitative Structure-Activity Relationship)

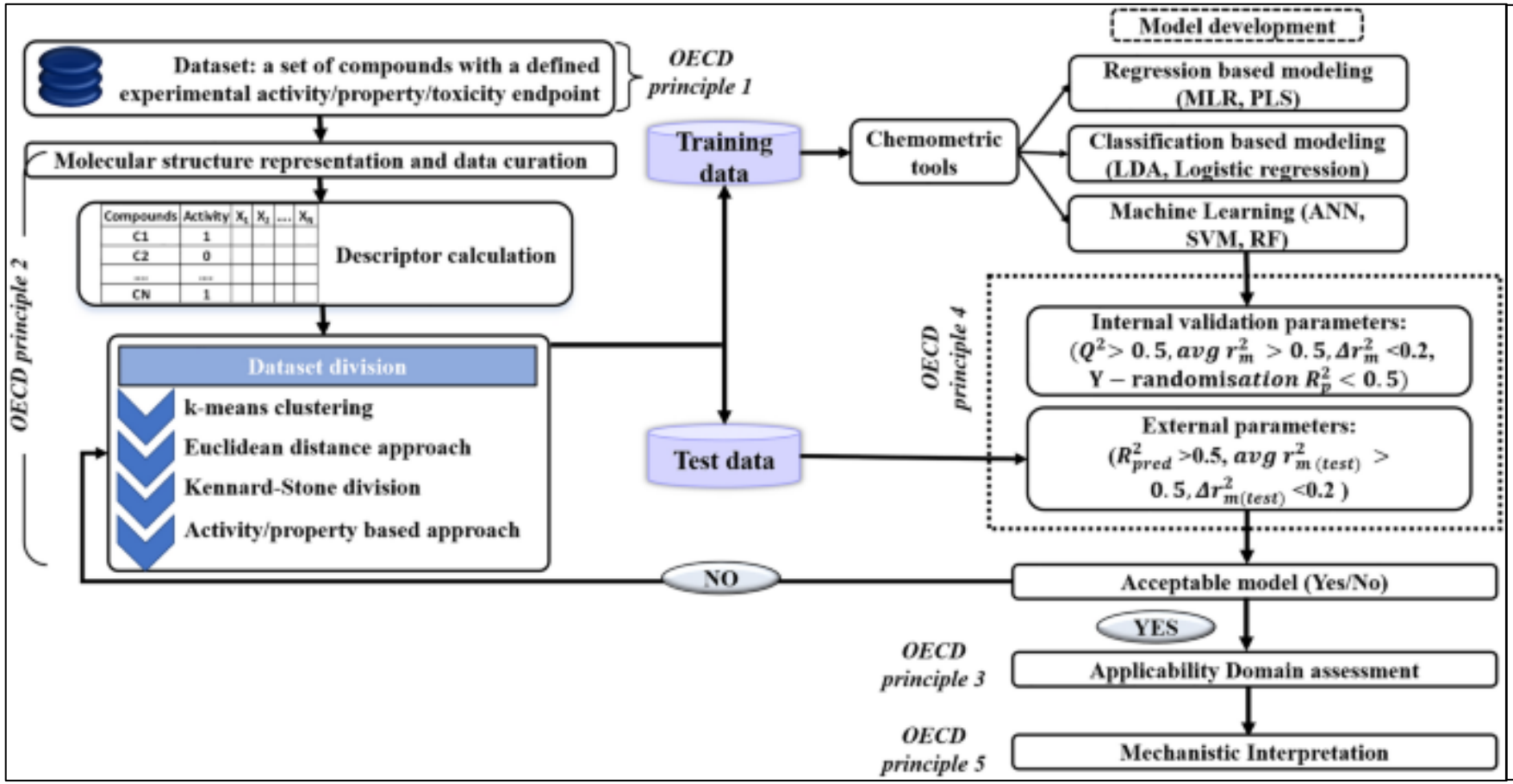
- QSAR is a mathematical relationship that can correlate chemistry with the properties (physicochemical/biological/toxicological) of molecules using various computationally or experimentally derived quantitative parameters.

- Descriptors are the numerical values that characterize specific information about a studied molecule.

Response (activity/property/toxicity)=*f*(Information in form of chemical structure or properties) =*f*(Descriptors)



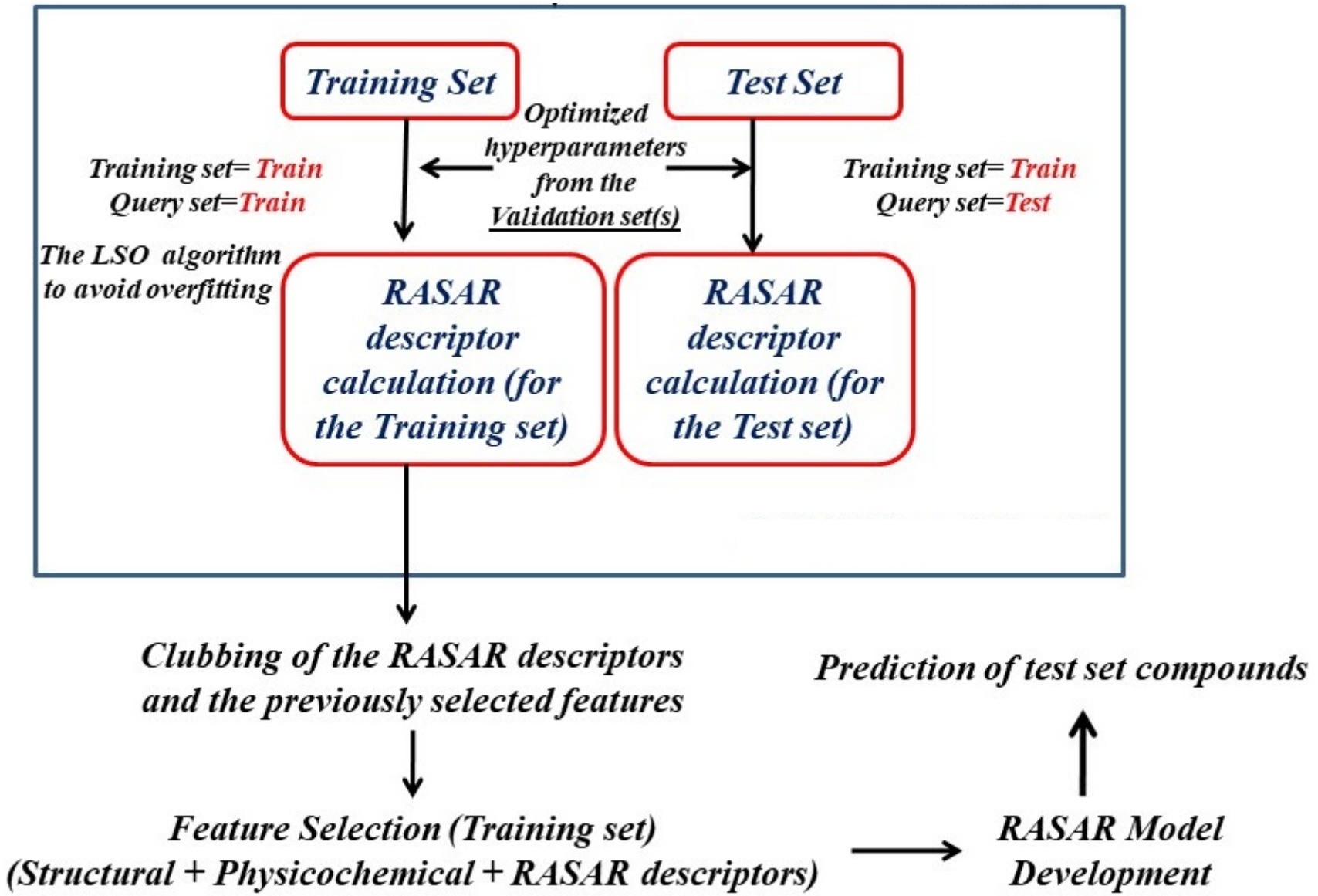
Standard algorithm of QSAR modeling: The OECD guideline



- ***OECD Principle 1***: A defined End point
- ***OECD Principle 2***: An unambiguous algorithm
- ***OECD Principle 3***: A defined domain of applicability
- ***OECD Principle 4***: Appropriate measure of goodness of fit, robustness, and predictivity
- ***OECD Principle 5***: A mechanistic interpretation, if possible

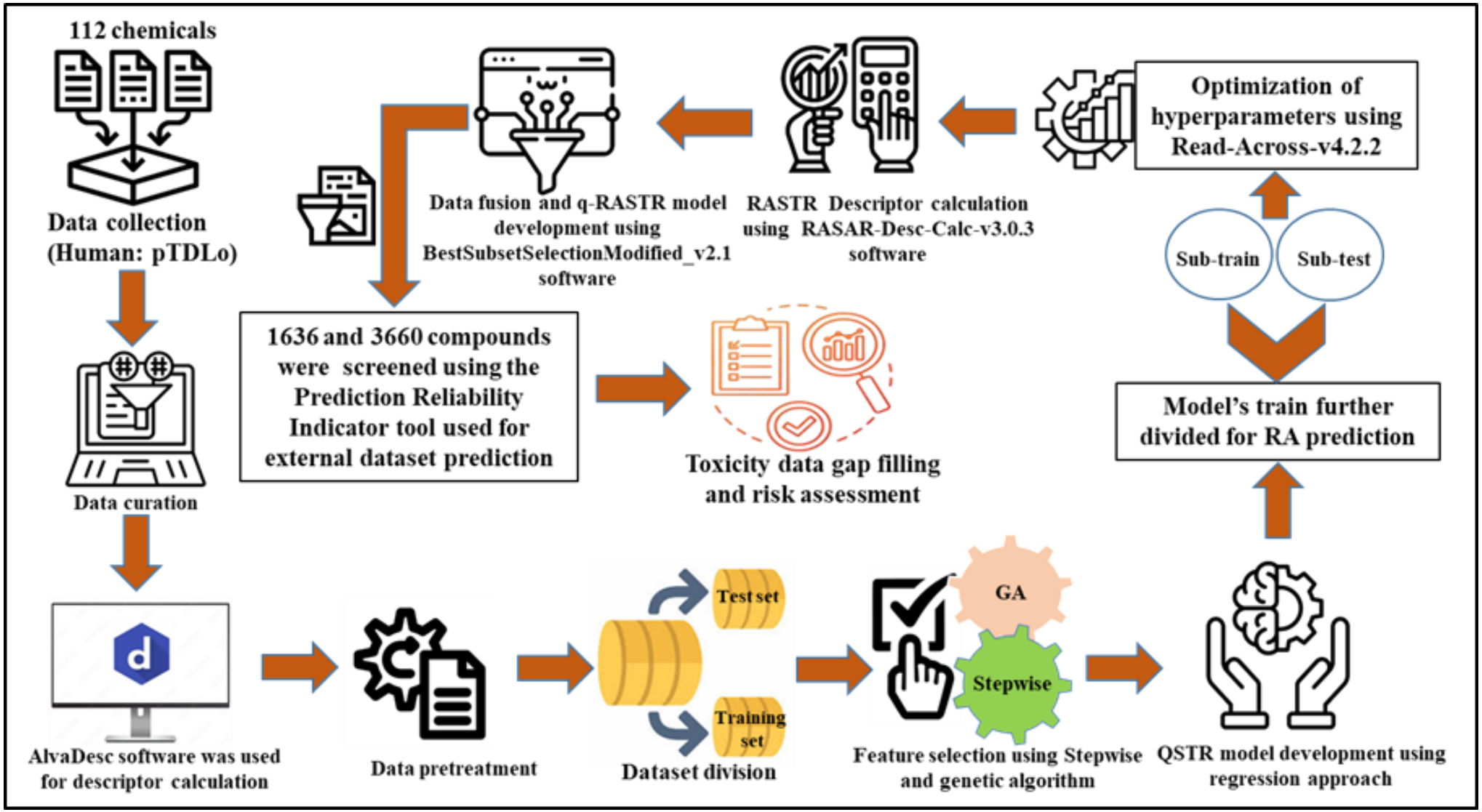
OECD, **2004**. Validation of (Q)SAR Models. URL https://www.oecd.org/chemicalsafety/risk-assessment/validationofqsarmodels.htm

q-RASAR (Quantitative Read Across Structure-Activity Relationship)



Available at: https://sites.google.com/jadavpuruniversity.in/dtc-lab-software

Methods and Materials (present work)



Results and discussion

PLS-based q-RASAR model : $\mathrm{pTDLo} = 3.93766 - 1.64409 \times CVact(GK) + 1.10047 \times CVsim(GK) + 1.08861 \times gm * Avg.Sim - 3.30300 \times nN = C - N < +0.79645 \times B05[C - C] + 0.67774 \times minaaaC + 0.52104 \times B08[C - C]$

Table 1. Statistical quality of QSTR and q-RASTR models.

Validation Metrics	PLS-based QSAR model	PLS-based q-RASAR model
No of des/No of LVs	7/2	7/3
N_{train}/N_{test}	84/28	84/28
R^2(train)	0.657	0.710
Q^2_{LOO} (train)	0.613	0.658
$Q^2_{F1\ (test)}$	0.764	0.812
$Q^2_{F2\ (test)}$	0.764	0.812
$Q^2_{F3\ (test)}$	0.771	0.818
MAE_{test}	0.579	0.545
CCC	0.872	0.898
$\overline{r^2_{m(test)}}$	0.679	0.741
Δr^2_m(test)	0.100	0.087
MAE-based prediction quality	Moderate	Good

Predictive accuracy and robustness assessment

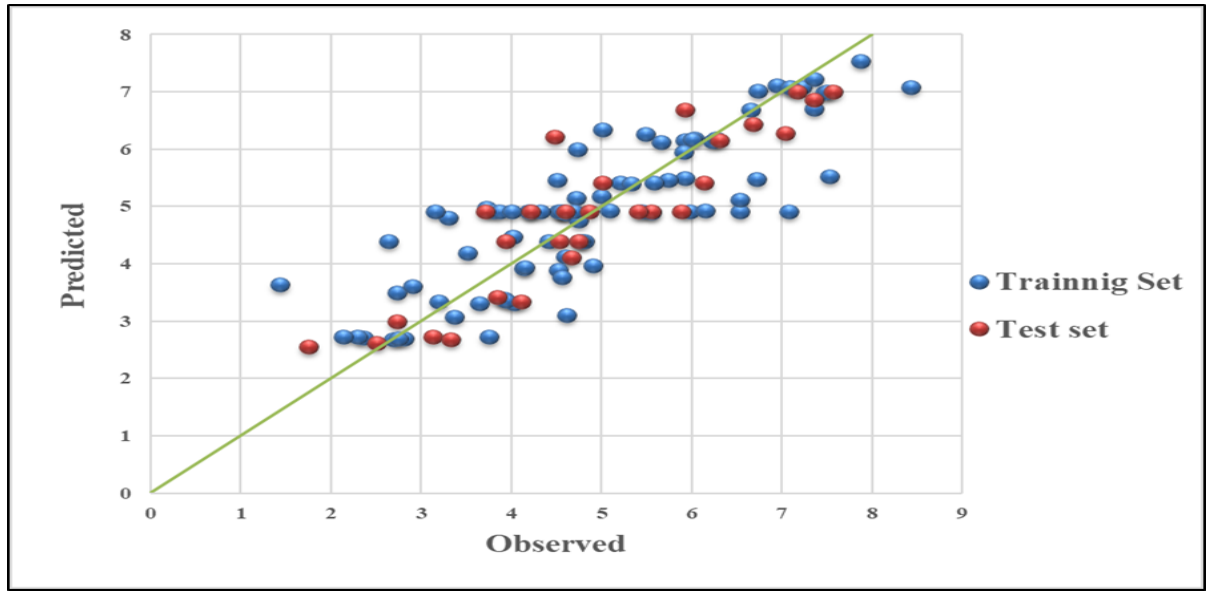


Figure 1. Scatter plot of the developed model.

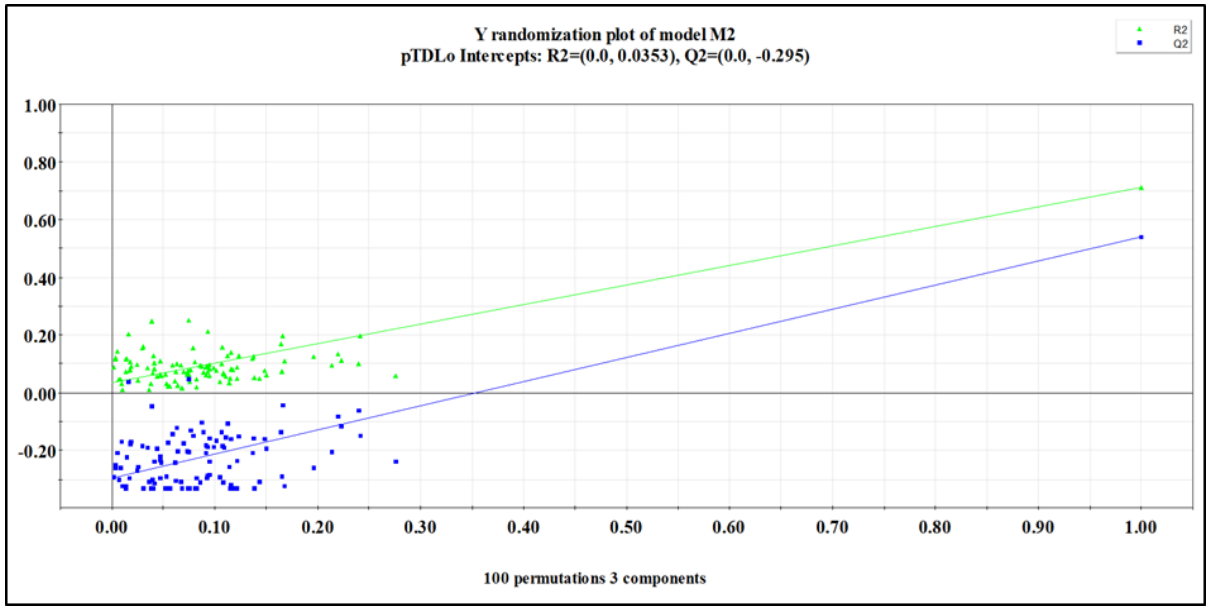


Figure 2. Y randomization plot of the developed model.

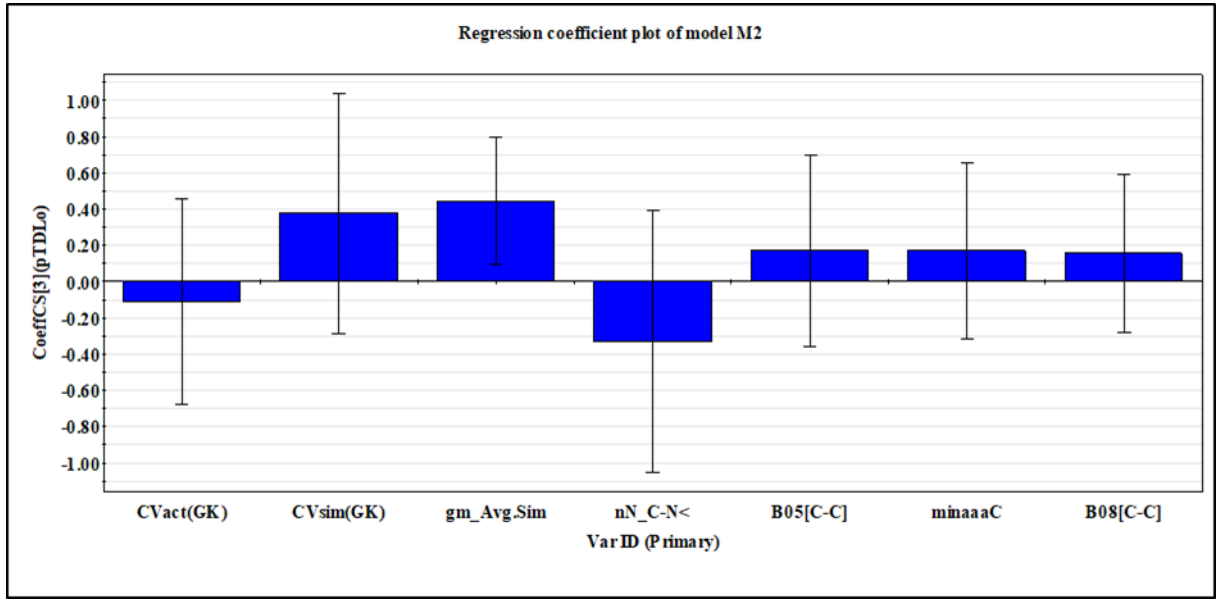


Figure 3. Regression coefficient plot of the developed model.

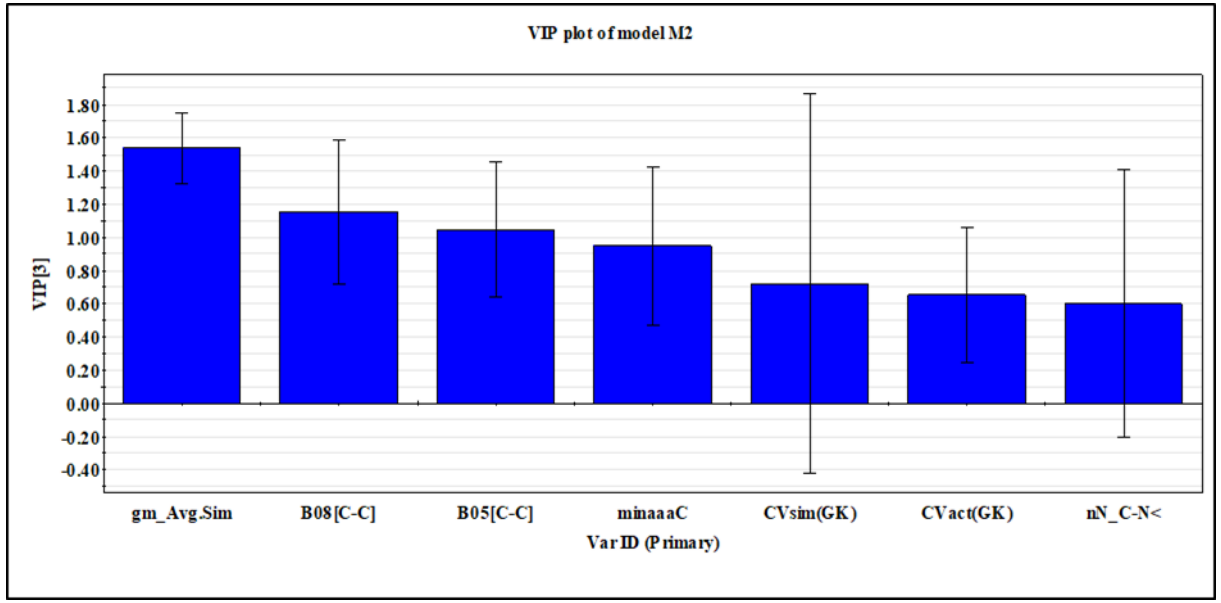
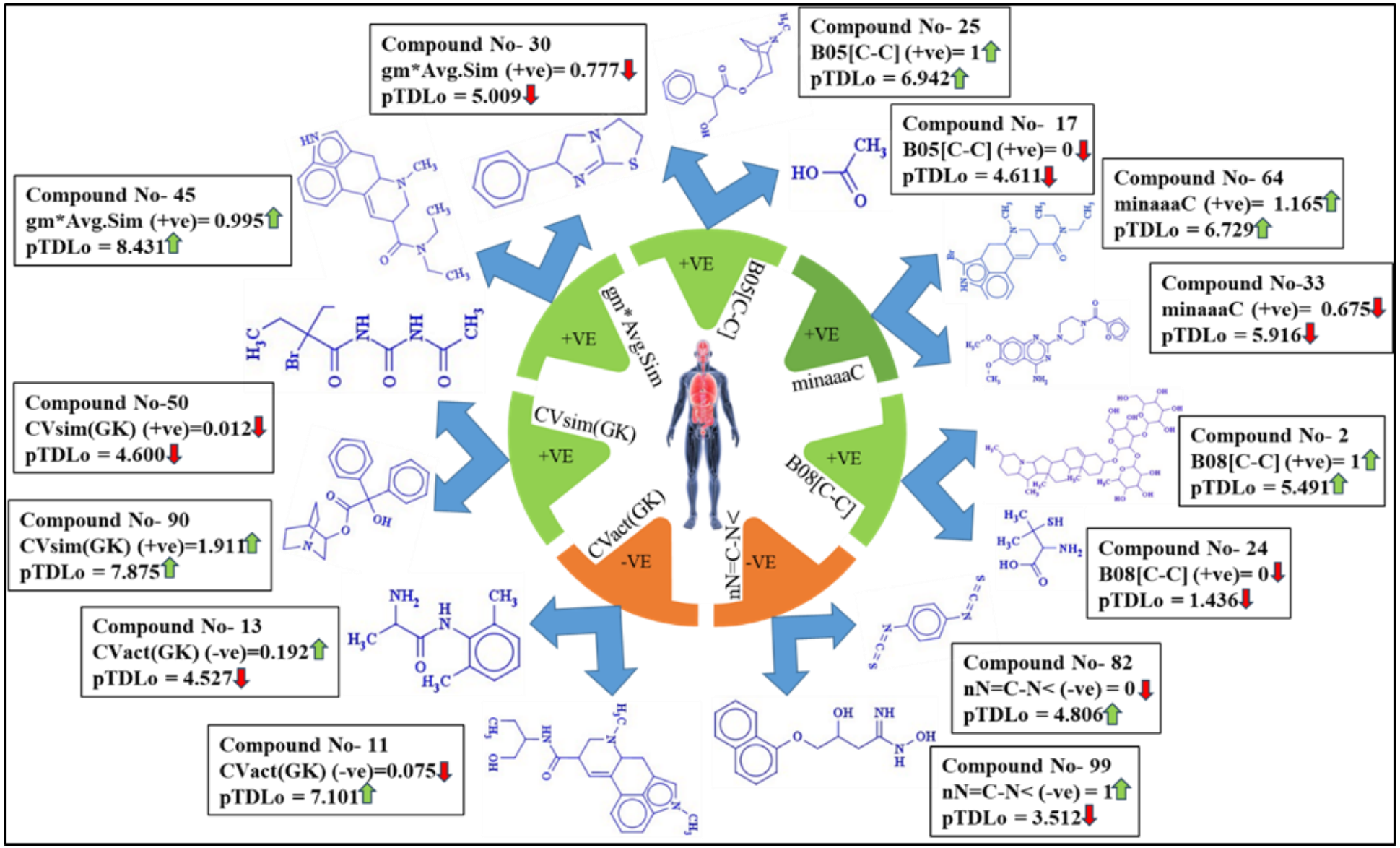


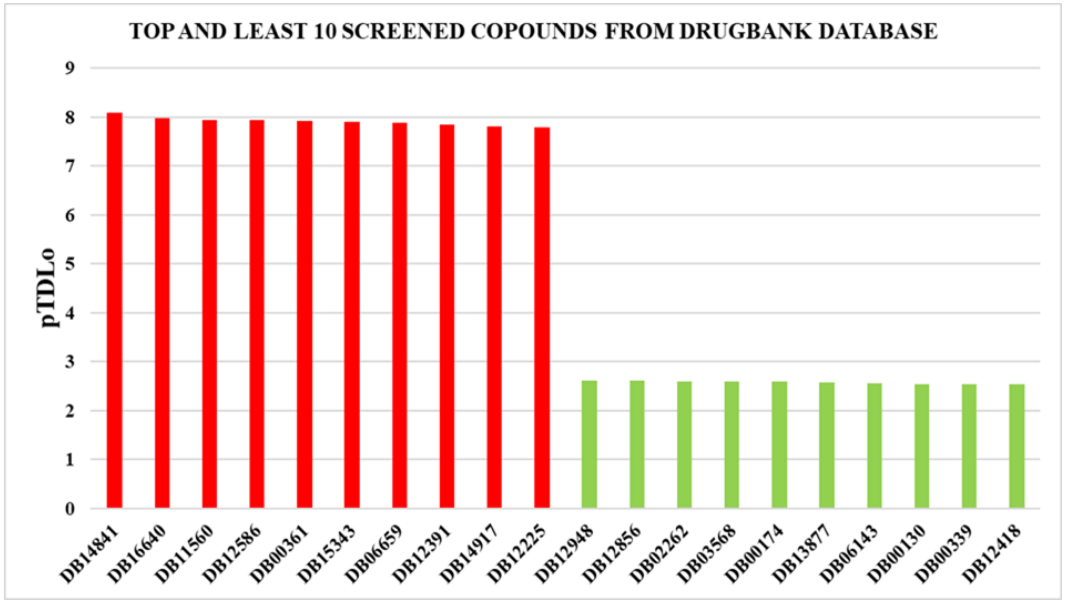
Figure 4. VIP plot of the developed model.

Mechanistic Interpretation



DrugBank database screening

Sl.No	DRUGBANK_ID	GENERIC_NAME	Predicted values (pTDLo)
Top 5 screened compounds from the DrugBank database (Investigational drugs)			
1	DB14841	Centanafadine	8.090
2	DB16640	Almonertinib	7.969
3	DB11560	Lesinurad	7.936
4	DB12586	Anhydrovinblastine	7.929
5	DB00361	Vinorelbine	7.920
Least 5 screened compounds from the DrugBank database (Investigational drugs)			
1	DB12418	Saccharin	2.545
2	DB00339	Pyrazinamide	2.545
3	DB00130	L-Glutamine	2.546
4	DB06143	Aminoimidazole carboxamide	2.553
5	DB13877	Iniparib	2.582



Conclusions

- Developed the first validated computational model for predicting human-specific acute toxicity using the pTDLo endpoint
- q-RASAR approach outperformed traditional QSAR, achieving robust validation metrics ($R^2 = 0.710$, $Q^2_{F1} = 0.812$)
- Models confirmed through Y-randomization, statistical testing, and external validation using PPDB (1,636 compounds) and DrugBank (3,660 drugs)
- Identified key structural features linked to toxicity:

Increased toxicity: C-C bonds at topological distances 5 & 8, higher E-state indices, similarity variations

Decreased toxicity: Amidine derivatives, consistent toxicological profiles

- Provides cost-effective, ethical alternative to animal testing for early pharmaceutical screening and regulatory applications
- Enables proactive identification of hazardous chemicals, protecting human health and the environment

Limitations & Future Prospects

Limitations:

- ❖ Gender-neutral model - Does not distinguish male/female toxicity responses
- ❖ Acute toxicity only – Chronic and long-term effects not predicted
- ❖ Limited chemical space – Organic compounds only (metals/nanomaterials excluded)
- ❖ Single endpoint – No organ-specific toxicity predictions
- ❖ Data dependency – Limited by available published human toxicity data

Future Prospects:

- ❖ Gender-specific models (male/female)
- ❖ Organ-specific & chronic toxicity
- ❖ Expanded chemical coverage
- ❖ AI/ML integration
- ❖ Web platform development

References

Brumovský, M., Bečanová, J., Kohoutek, J., Borghini, M. and Nizzetto, L., 2017. Contaminants of emerging concern in the open sea waters of the Western Mediterranean. Environmental pollution, 229, pp.976-983. https://doi.org/10.1016/j.envpol.2017.07.082.

CAS, 2023. CAS [WWW Document]. CAS. https://www.cas.org/cas-data/cas-registry.

DiMasi, J.A., Feldman, L., Seckler, A. and Wilson, A., 2010. Trends in risks associated with new drug development: success rates for investigational drugs. Clinical Pharmacology & Therapeutics, 87(3), pp.272-277. https://doi.org/10.1038/clpt.2009.295

Wong, C.H., Siah, K.W. and Lo, A.W., 2019. Estimation of clinical trial success rates and related parameters. Biostatistics, 20(2), pp.273-286. https://doi.org/10.1093/biostatistics/kxy072.

Hu, Y., Ren, Q., Liu, X., Gao, L., Xiao, L. and Yu, W., 2023. In silico prediction of human organ toxicity via artificial intelligence methods. Chemical Research in Toxicology, 36(7), pp.1044-1054. https://pubs.acs.org/doi/full/10.1021/acs.chemrestox.2c00411.

Bender, A., Scheiber, J., Glick, M., Davies, J.W., Azzaoui, K., Hamon, J., Urban, L., Whitebread, S. and Jenkins, J.L., 2007. Analysis of pharmacology data and the prediction of adverse drug reactions and off-target effects from chemical structure. *ChemMedChem: Chemistry Enabling Drug Discovery*, *2*(6), pp.861-873. https://doi.org/10.1002/cmdc.200700026.

Hartung, T., 2009. A toxicology for the 21st century—mapping the road ahead. Toxicological Sciences, 109(1), pp.18-23. https://doi.org/10.1093/toxsci/kfp059.

Gellatly, N. and Sewell, F., 2019. Regulatory acceptance of in silico approaches for the safety assessment of cosmetic-related substances. Computational Toxicology, 11, pp.82-89. https://doi.org/10.1016/j.comtox.2019.03.003.


Thank
you