INTERNATIONAL JOURNAL OF PHARMACEUTICAL SCIENCES
[ISSN: 0975-4725; CODEN(USA): IJPS00]
Journal Homepage: https://www.ijpsjournal.com

OPEN ACCESS

Review Article

Quantitative Structure–Activity Relationship (QSAR) in Modern Drug Design: A Review of Methods, Validation, Applications, and Future Perspectives

Araj Ranee D.*, Jadhav Mohit M., Adhikari Samidha S., Adhikari Ishawari L., Padwal Krutika B.

Department of Pharmaceutics, Siddhi's Institute of Pharmacy, Nandgaon Murbad Thane 421401 (Affiliated to Dr. Babasaheb Ambedkar Technological University, Lonere Raigad).

ARTICLE INFO

Published: 16 Nov 2026
Keywords:
QSAR, drug design, molecular descriptors, machine learning, deep learning, model validation, applicability domain, virtual screening, ADMET
DOI:
10.5281/zenodo.24271173

ABSTRACT

Quantitative structure–activity relationship (QSAR) modelling connects the chemical structure of compounds described using values to their biological effects how the body handles them or their harmful properties. This review looks at how QSAR has changed since the 1960s starting with the Hansch and Free–Wilson methods and how it has grown into today's use of machine learning. It covers the types of molecular descriptors the modelling techniques used validation practices and how QSAR helps in designing new drugs. The review is based on peer-reviewed studies found in PubMed, Scopus, Web of Science and Google Scholar with attention to guidelines from the OECD and recent improvements in the field. Traditional 2D and 3D-QSAR methods like CoMFA and CoMSIA are still useful especially when working with compounds and when understanding the science behind the results is important. For more varied datasets modern methods such as random forests support vector machines, deep neural networks and graph neural networks show better performance.. Success doesn't come just from using advanced algorithms. The reliability of a QSAR model depends a lot on how well the data is cleaned and prepared how it is tested using both internal and external validation whether y-randomisation is done and how clearly the models applicability domain is defined. QSAR plays a role in many areas of drug development. It is used in screening to find promising compounds in optimizing lead molecules in predicting absorption, distribution, metabolism, excretion and toxicity (ADMET) and in predicting harmful effects.

***Corresponding Author:** Araj Ranee D.
Address: Department of Pharmaceutics, Siddhi's Institute of Pharmacy, Nandgaon Murbad Thane 421401 (Affiliated to Dr. Babasaheb Ambedkar Technological University, Lonere Raigad).
Email ✉: raneearaj@gmail.com
Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

It has also helped in discovering antibiotics like halicin using artificial intelligence. Despite its usefulness QSAR still faces some problems. These include inconsistent data, activity cliffs where small changes lead to big differences in activity, limited understanding of model decisions and the need to estimate uncertainty and do prospective testing. Looking ahead progress will come from AI that makes models easier to understand from transfer learning and multi-task learning that allow models to use knowledge from different tasks and, from combining QSAR with structure-based methods and system-level approaches. Together these advances help keep QSAR as an cost-effective part of modern drug design.

INTRODUCTION

Drug discovery is a costly and high-attrition process and computational methods are now used routinely to prioritise compounds before synthesis and testing [1,2]. Among these quantitative structure–activity relationship (QSAR) modelling is one of the oldest and widely applied. QSAR rests on the similar-property principle: similar molecules tend to exhibit similar biological effects so activity can be expressed as a mathematical function of molecular structure Activity = f (descriptors) [3,4].

The field began in the 1960s, when Hansch and coworkers related the potency of phenoxyacetic acids to Hammett electronic constants and the octanol–water partition coefficient [5,6] and Free and Wilson proposed a group-contribution model [7]. The introduction of three- field-based methods CoMFA [8] and CoMSIA [9] brought spatial information into the models. Over the two decades large public bioactivity repositories such as ChEMBL and PubChem [10,11] and advances in machine learning [12–14] have transformed QSAR into a throughput, data-intensive discipline [15,16].

Despite this progress published QSAR models vary widely in quality. Common pitfalls include curated structures, overfitting, misuse of the cross-validated coefficient q^2 as a proof of predictivity and the absence of a defined applicability domain [17–19]. This review therefore (i) summarises the evolution and theory of QSAR, (ii) compares descriptor types and modelling algorithms (iii) discusses validation and regulatory principles, (iv) highlights applications, in drug design and (v) outlines challenges and future directions.

2. Scope and Review Methodology

The searches used combinations of the terms "QSAR" "3D-QSAR" " descriptors" "machine learning" "deep learning" "model validation" "applicability domain" "virtual screening" and "ADMET".

Priority was given to (a) methodological papers, (b) validation and best-practice guidelines including the OECD principles [20] (c) benchmark and application studies in drug discovery and (d) authoritative reviews.

Table 1 summarises the scope, in a PICO-style format adapted for modelling studies.

Element	Definition	Criteria applied
Population	Chemical datasets	Small-molecule series with measured activity, ADMET, or toxicity endpoints
Intervention	QSAR modelling	Hansch, Free–Wilson, 3D-QSAR, classical ML, deep learning, consensus models
Comparison	Alternative methods	Algorithm-to-algorithm and descriptor-to-descriptor benchmarks; docking and pharmacophore models
Outcome	Model quality	Internal (Q^2) and external (R^2ext, CCC, RMSE) validation; applicability domain; prospective experimental confirmation
Study type	Literature	Methodological papers, benchmarks, guidelines, and application case studies

Table 1: ***Scope of the review (PICO-style framework adapted to computational studies)***

3. Evolution and Fundamentals of QSAR

QSAR has evolved through four overlapping phases. In the phase of the 1960s QSAR used linear free-energy relationships, where activity was modelled with lipophilic, electronic and steric parameters [5–7]. The second phase, during the 1980s to 1990s brought 3D-QSAR. In 3D-QSAR molecules are aligned on a grid and interaction energies are linked to activity using

partial least squares regression [8,9,21]. The third phase in the 2000s saw QSAR move to descriptor spaces and ensemble learning and a stronger focus, on validation [20,22,23]. Today QSAR is dominated by learning and graph-based methods and by approaches that aim to be explainable [13,14,24].



Figure 1: ***Timeline of major milestones in the development of QSAR methodology, from classical Hansch analysis to graph neural networks and explainable artificial intelligence***

A QSAR model is built on three components. The first component is a training set of compounds that have comparable biological measurements. The second component is a representation of the chemical structure, such as descriptors or learned embeddings. The third component is an machine-learning algorithm that links the representation to the response. The response variable can be continuous for example pIC_{50} or pK_i for regression models or categorical such as /inactive or toxic/non-toxic for classification models. Figure 2 shows the generalised workflow. In this workflow curation, validation and applicability-domain assessment are as important, as model fitting [3,4].

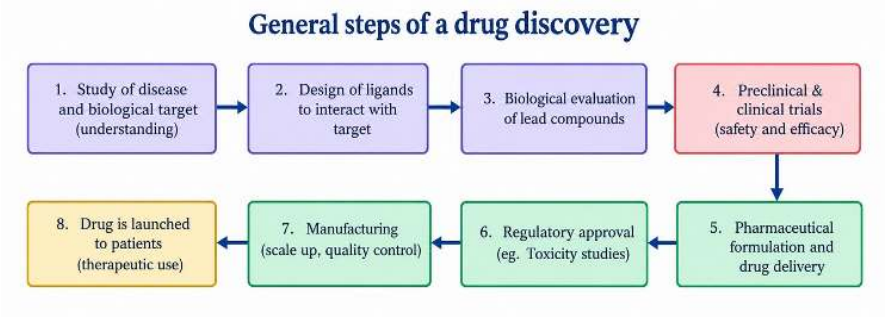


Figure 2: ***Generalised QSAR modelling workflow, from data curation through model building, internal and external validation, applicability-domain definition, and application in drug-design decision making.***

4. Molecular Descriptors and Data Curation

4.1 Data curation

Model quality is limited by data quality. Structural errors including stereochemistry, salts, tautomers and duplicates and varying assay data are common in public databases. Fourches et al. Found that curation significantly improves predictivity [18]. Recommended steps include standardizing structures removing inorganic and mixture records handling duplicates and activity cliffs and when combining data from labs limiting to similar assay conditions [10,11,25].

4.2 Descriptor classes

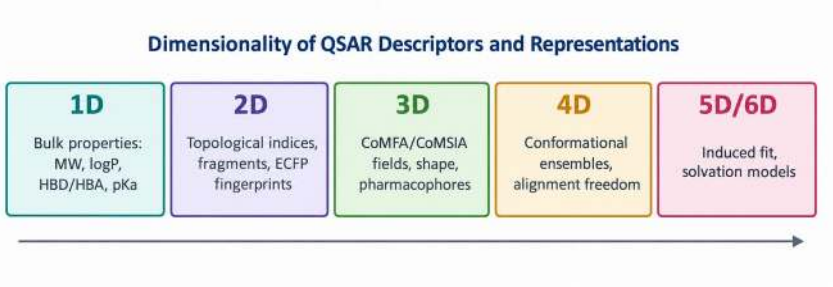


Figure 3: ***Hierarchy of QSAR descriptors and representations by dimensionality, with increasing structural information and computational cost from 1D bulk properties to multidimensional (5D/6D) approaches.***

Descriptors are grouped based on the dimensionality of the information they represent (Figure 3). Many descriptors can be calculated using software like PaDEL-Descriptor and Mordred [26,27]. The theory is explained in detail by Todeschini and Consonni [28]. Circular fingerprints such, as ECFP are some of the commonly used 2D representations [29]. Table 2 shows the types of descriptors.

Class	Examples	Advantages	Limitations
1D / constitutional	Molecular weight, logP, H-bond donors/ acceptors, pKa [30]	Fast, interpretable; rule-of-five filters	Low discrimination of close analogues.
2D topological	Connectivity indices, ECFP fingerprints, fragments [28,29]	Alignment-free; robust; scalable	No conformational information.
3D field-based	CoMFA steric/electrostatic fields; CoMSIA similarity indices [8,9]	Mechanistic insight; contour maps	Alignment- and conformer-dependent.
Quantum-chemical	HOMO/LUMO energies, partial charges, dipole moment	Capture electronic effects	Costly for large libraries.
Learned embeddings	Graph-network features [14,31,32]	Data-driven; no manual selection	Large data needs; lower transparency.

Table 2: Major Classes of Molecular Descriptors Used in QSAR

I find that when many descriptors are correlated with compounds we must use feature selection methods such as algorithms, stepwise selection, LASSO or random-forest importance. We also need dimensionality reduction techniques, like PCA or PLS. These steps are essential to limit overfitting [4,21].

5. Modelling Algorithms

regression, especially multiple linear regression (MLR) and PLS continues to be useful for small similar datasets because the coefficients can be understood directly [4,21]. Non-linear machine-learning methods are now common for datasets. Random forests use separate decision trees and are good with noise need very little adjustment and show which variables are most important [12,22]. Support vector machines (SVM) use functions to handle non-linear patterns and work well with datasets that are not too big [33]. Multi-task deep neural networks, trained with thousands of features across targets did better than random forests in a big industry test [13] and DeepTox won the Tox21 challenge [34].

In times graph neural networks (GNNs) take information directly from the molecular graph so there is no need for manual feature creation [31,32]. Directed message-passing networks reached the results, on many tests [14,25] and a GNN-based search found a new type of antibiotic called halicin [35]. However when looking closely it is clear that on noisy data, simple methods and groups of models often do just as well as deep models [16,36]. Table 3 shows the strengths of the different algorithms.

Algorithm	Type	Strengths	Weaknesses
MLR / PLS [21]	Linear	Interpretable; handles collinearity (PLS)	Cannot capture non-linearity
Random forest [12,22]	Ensemble trees	Robust, little tuning, importance ranking	Limited extrapolation
SVM [33]	Kernel method	Good for small/medium data	Kernel choice; poor scalability
Deep neural network [13,34]	Multi-layer perceptron	Multi-task learning; high accuracy on big data	Data-hungry; hyper-parameters
Graph neural network [14,32]	Message passing	Learns representation from the graph	Needs large data; opaque

Table 3: ***Comparison of commonly used QSAR modelling algorithms***

6. Validation and the Applicability Domain

The OECD principles state that a QSAR model must have an endpoint a well-defined algorithm, a defined domain of applicability good measures of fit, robustness and predictivity and a mechanistic explanation when possible [20]. Validation of a QSAR model involves steps [19,23].

Internal validation uses methods like leave-one-out or k-fold cross-validation to calculate Q^2. A high Q^2 value is important. It is not enough to ensure that the model can predict well [17]. To check for correlations y-randomisation is used. This means the activity values are randomly shuffled. A good model should not work well when the response is scrambled [37].

The important step is external validation, which uses a separate test set that was never involved in building or tuning the model. This test set shows how well the model performs on data. Common metrics, for this include R^2ext, RMSE and the concordance correlation coefficient (CCC) [17,38,39]. Acceptance criteria that are often used are shown in Table 4.

Parameter	Typical criterion	Purpose / reference
R^2 (training)	> 0.6	Goodness of fit [17]
Q^2 (cross-validation)	> 0.5	Internal robustness; not sufficient alone [17,23]
R^2_{ext} (test set)	> 0.6	External predictivity; sensitive to test-set range [39]
CCC (test set)	> 0.85	Agreement between observed and predicted values[38]
Y-randomisation	R^2 of scrambled models << true R^2	Excludes chance correlation [37]
Applicability domain	Leverage $h < h^*$; |std. residual| < 3	Reliable-prediction region [40,41]

Table 4: ***Frequently used statistical criteria for QSAR model validation***

The applicability domain, abbreviated AD sets the chemical space where predictions can be trusted. Methods used to determine this domain include geometric techniques, distance-to-model calculations and the leverage approach that is shown in the Williams plot [40,41]. If a compound lies outside the AD it should be flagged, not reported as reliable. Figure 4 shows the diagnostic plots that help validate the model.

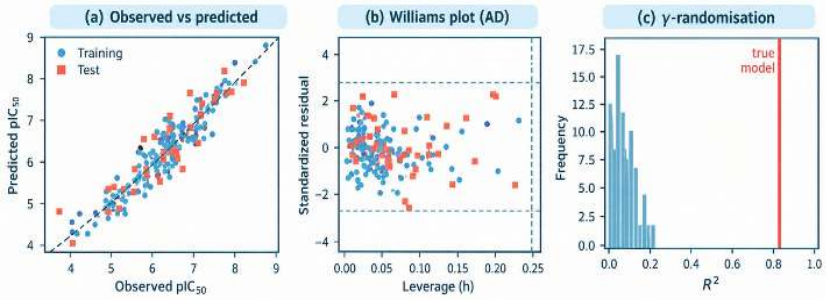


Figure 4: ***Illustrative validation diagnostics (simulated data for demonstration only): (a) observed versus predicted activity for training and test sets; (b) Williams plot defining the applicability domain by leverage (h*) and standardised residuals (±3); (c) y-randomisation test, in which the true model (red line) is well separated from models built on scrambled activities.***

7. Applications in Drug Design

7.1 screening and lead optimisation

QSAR models help scientists sort through huge virtual collections of molecules. QSAR models reduce the number of compounds that need to be built or tested making the search faster. QSAR models also show which parts of a molecule should change to improve the lead. 3D-QSAR contour maps reveal where large groups, hydrogen-bond donors or charges are beneficial giving clues for medicinal chemists [8,9]. When the target shape is partly known, QSAR models, together with docking and pharmacophore models give extra support based on how the ligand looks [2,15].

7.2 ADMET and toxicity prediction

Poor pharmacokinetics and high toxicity still cause many drug projects to fail later. Simple physicochemical rules, such as Lipinski's rule of five are early filters for oral absorption [30]. Multi-task deep learning models predict toxicity points from the Tox21 programme. QSAR models are also accepted by regulators as a way to reduce animal testing when they are validated following OECD principles [3,20].

7.3 Learning-enabled discovery

A directed-message-passing neural network trained on growth-inhibition data from a few thousand molecules was used to screen more than 100 million compounds. The network found halicin and other antibacterial candidates, with different structures and these were later confirmed in experiments. This example shows how modern QSAR can go beyond tweaking known scaffolds and can hop between different scaffolds across a large chemical space [1,16].

8. Future Perspectives

Data quality and quantity. I find that experimental noise, assay heterogeneity and publication bias reduce the accuracy that can be achieved and that benchmark datasets may not match the realities of discovery settings [18,25,36]. Activity. Extrapolation. I see that small structural changes can lead to differences in potency and that most models struggle with this effect. Interpretability. Black-box models need explanations after the fact and tools like SHAP and attribution maps help medicinal chemists make sense of predictions [24,42]. Uncertainty quantification. I believe that calibrated confidence estimates and defined applicability domains are crucial when making decisions [16,40]. Prospective validation. I think that retrospective statistics need to be paired with synthesis and testing just like the halicin study showed [35].

Promising directions include transfer learning and multi-task learning that make the most of data the blending of QSAR with structure-based design and molecular dynamics federated learning that lets organisations collaborate without revealing proprietary structures and generative models that suggest new molecules while keeping QSAR-predicted properties in check [1,16,36]. I feel that adopting community standards, for curation, reporting and benchmarking will be essential if we want results to be reproducible [3,20].

CONCLUSION

QSAR has changed over six decades from linear relationships into a group of data-driven methods that help almost every part of drug discovery. The information shown here shows that good predictions mainly rely on collected data the right descriptors, strong outside testing and a clear area where the model can be used not just, on how complicated the algorithm is. Deep and graph-based models help QSAR reach more varied chemical areas but they need to have ways to explain their results and show how sure they are. When used with experiments and other tools that look at structures QSAR will keep being a part of smart and affordable drug development.

REFERENCES

1. Vamathevan J, Clark D, Czodrowski P, et al. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov. 2019;18 (6):463-477. doi:10.1038/s41573-019-0024-5.
2. Ekins S, Mestres J, Testa B. In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling. Br J Pharmacol. 2007;152 (1):9-20. doi:10.1038/sj.bjp.0707305.
3. Tropsha A. Best practices for QSAR model development, validation, and exploitation. Mol Inform. 2010;29 (6-7):476-488. doi:10.1002/minf.201000061.
4. Roy K, Kar S, Das RN. A primer on QSAR/QSPR modeling: fundamental concepts. Cham: Springer; 2015. doi:10.1007/978-3-319-17281-1.
5. Hansch C, Maloney PP, Fujita T, Muir RM. Correlation of biological activity of phenoxyacetic acids with Hammett substituent constants and partition coefficients. Nature. 1962;194:178-180. doi:10.1038/194178b0.
6. Hansch C, Fujita T. ρ-σ-π analysis. A method for the correlation of biological activity and chemical structure. J Am Chem Soc. 1964;86 (8):1616-1626. doi:10.1021/ja01062a035.
7. Free SM, Wilson JW. A mathematical contribution to structure-activity studies. J Med Chem. 1964;7 (4):395-399. doi:10.1021/jm00334a001.
8. Cramer RD, Patterson DE, Bunce JD. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. J Am Chem Soc. 1988;110 (18):5959-5967. doi:10.1021/ja00226a005.
9. Klebe G, Abraham U, Mietzner T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. J Med Chem. 1994;37 (24):4130-4146. doi:10.1021/jm00050a010.
10. Gaulton A, Hersey A, Nowotka M, et al. The ChEMBL database in 2017. Nucleic Acids Res. 2017;45 (D1):D945-D954. doi:10.1093/nar/gkw1074.
11. Kim S, Chen J, Cheng T, et al. PubChem 2023 update. Nucleic Acids Res. 2023;51 (D1):D1373-D1380. doi:10.1093/nar/gkac956.
12. Breiman L. Random forests. Mach Learn. 2001;45 (1):5-32. doi:10.1023/A:1010933404324.
13. Ma J, Sheridan RP, Liaw A, Dahl GE, Svetnik V. Deep neural nets as a method for quantitative structure-activity relationships. J Chem Inf Model. 2015;55 (2):263-274. doi:10.1021/ci500747n.
14. Yang K, Swanson K, Jin W, et al. Analyzing learned molecular representations for property prediction. J Chem Inf Model. 2019;59 (8):3370-3388. doi:10.1021/acs.jcim.9b00237.
15. Cherkasov A, Muratov EN, Fourches D, et al. QSAR modeling: where have you been? Where are you going to? J Med Chem. 2014;57 (12):4977-5010. doi:10.1021/jm4004285.
16. Muratov EN, Bajorath J, Sheridan RP, et al. QSAR without borders. Chem Soc Rev. 2020;49 (11):3525-3564. doi:10.1039/D0CS00098A.
17. Golbraikh A, Tropsha A. Beware of q2! J Mol Graph Model. 2002;20 (4):269-276. doi:10.1016/S1093-3263 (01)00123-1.
18. Fourches D, Muratov E, Tropsha A. Trust, but verify: on the importance of chemical structure curation in cheminformatics and QSAR modeling research. J Chem Inf Model. 2010;50 (7):1189-1204. doi:10.1021/ci100176x.

19. Tropsha A, Gramatica P, Gombar VK. The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models. QSAR Comb Sci. 2003;22 (1):69-77. doi:10.1002/qsar.200390007.
20. Organisation for Economic Co-operation and Development (OECD). Guidance document on the validation of (quantitative) structure-activity relationship [(Q)SAR] models. OECD Series on Testing and Assessment No. 69. Paris: OECD Publishing; 2007.
21. Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. Chemom Intell Lab Syst. 2001;58 (2):109-130. doi:10.1016/S0169-7439 (01)00155-1.
22. Svetnik V, Liaw A, Tong C, Culberson JC, Sheridan RP, Feuston BP. Random forest: a classification and regression tool for compound classification and QSAR modeling. J Chem Inf Comput Sci. 2003;43 (6):1947-1958. doi:10.1021/ci034160g.
23. Gramatica P. Principles of QSAR models validation: internal and external. QSAR Comb Sci. 2007;26 (5):694-701. doi:10.1002/qsar.200610151.
24. Jiménez-Luna J, Grisoni F, Schneider G. Drug discovery with explainable artificial intelligence. Nat Mach Intell. 2020;2 (10):573-584. doi:10.1038/s42256-020-00236-4.
25. Wu Z, Ramsundar B, Feinberg EN, et al. MoleculeNet: a benchmark for molecular machine learning. Chem Sci. 2018;9 (2):513-530. doi:10.1039/C7SC02664A.
26. Yap CW. PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. J Comput Chem. 2011;32 (7):1466-1474. doi:10.1002/jcc.21707.
27. Moriwaki H, Tian YS, Kawashita N, Takagi T. Mordred: a molecular descriptor calculator. J Cheminform. 2018;10:4. doi:10.1186/s13321-018-0258-y.
28. Todeschini R, Consonni V. Molecular descriptors for chemoinformatics. 2nd ed. Weinheim: Wiley-VCH; 2009. doi:10.1002/9783527628766.
29. Rogers D, Hahn M. Extended-connectivity fingerprints. J Chem Inf Model. 2010;50 (5):742-754. doi:10.1021/ci100050t.
30. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev. 1997;23 (1-3):3-25. doi:10.1016/S0169-409X (96)00423-1.
31. Duvenaud D, Maclaurin D, Aguilera-Iparraguirre J, et al. Convolutional networks on graphs for learning molecular fingerprints. Adv Neural Inf Process Syst. 2015;28:2224-2232.
32. Gilmer J, Schoenholz SS, Riley PF, Vinyals O, Dahl GE. Neural message passing for quantum chemistry. In: Proceedings of the 34th International Conference on Machine Learning. PMLR. 2017;70:1263-1272.
33. Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995;20 (3):273-297. doi:10.1007/BF00994018.
34. Mayr A, Klambauer G, Unterthiner T, Hochreiter S. DeepTox: toxicity prediction using deep learning. Front Environ Sci. 2016;3:80. doi:10.3389/fenvs.2015.00080.
35. Stokes JM, Yang K, Swanson K, et al. A deep learning approach to antibiotic discovery. Cell. 2020;180 (4):688-702.e13. doi:10.1016/j.cell.2020.01.021.
36. Bender A, Cortés-Ciriano I. Artificial intelligence in drug discovery: what is realistic, what are illusions? Part 1: ways to make an impact, and why we are not there yet. Drug Discov Today. 2021;26 (2):511-524. doi:10.1016/j.drudis.2020.12.009.
37. Rücker C, Rücker G, Meringer M. y-Randomization and its variants in QSPR/QSAR. J Chem Inf Model. 2007;47 (6):2345-2357. doi:10.1021/ci700157b.
38. Chirico N, Gramatica P. Real external predictivity of QSAR models: how to evaluate it? Comparison of different validation criteria and proposal of using the concordance correlation coefficient. J Chem Inf Model. 2011;51 (9):2320-2335. doi:10.1021/ci200211n.

39. Alexander DLJ, Tropsha A, Winkler DA. Beware of R2: simple, unambiguous assessment of the prediction accuracy of QSAR and QSPR models. J Chem Inf Model. 2015;55 (7):1316-1322. doi:10.1021/acs.jcim.5b00206.
40. Sahigara F, Mansouri K, Ballabio D, Mauri A, Consonni V, Todeschini R. Comparison of different approaches to define the applicability domain of QSAR models. Molecules. 2012;17 (5):4791-4810. doi:10.3390/molecules17054791.
41. Jaworska J, Nikolova-Jeliazkova N, Aldenberg T. QSAR applicability domain estimation by projection of the training set descriptor space: a review. Altern Lab Anim. 2005;33 (5):445-459. doi:10.1177/026119290503300508.
42. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Adv Neural Inf Process Syst. 2017;30:4765-4774.

HOW TO CITE: Araj Ranee D., Jadhav Mohit M., Adhikari Samidha S., Adhikari Ishawari L., Padwal Krutika B., Quantitative Structure–Activity Relationship (QSAR) in Modern Drug Design: A Review of Methods, Validation, Applications, and Future Perspectives, Int. J. of Pharm. Sci., 2026, Vol 4, Issue 11, 2285-2293. https://doi.org/10.5281/zenodo.24271173