International Journal of Pharmacy and Pharmaceutical Research (IJPPR)
Volume 32, Issue 7, July 2026 ijppr.humanjournals.com ISSN: 2349-7203

Integrating Design of Experiments and Machine Learning in Reversed-Phase High Performance Liquid Chromatography Method Development: A Comprehensive Review

Pandiyan Govindasamy*[1], Rakshana Subramanian[2], Negalya Somasundaram[3], Sunitha PG[4], Deattu Narayanan[5]

*[1,2,3]Post Graduate Student, Department of Pharmaceutical Chemistry, College of Pharmacy, Madras Medical College, Chennai, Tamil Nadu 600 003, India
[4,5]Assistant Professor, Department of Pharmaceutical Chemistry, College of Pharmacy, Madras Medical College, Chennai, Tamil Nadu 600 003, India
Affiliated to The Tamil Nadu Dr.M.G.R. Medical University, Chennai 600 003, Tamil Nadu, India

Received: 27 May 2026 Revised: 15 June 2026 Accepted: 28 June 2026

ABSTRACT

Reversed-phase high-performance liquid chromatography (RP-HPLC) continues to be the primary analytical technique for identifying, quantifying, and evaluating the stability of active pharmaceutical ingredients and their formulations. Conventional method development using the one-factor-at-a-time (OFAT) approach is laborious, consumes significant resources, and frequently results in methods with limited operational flexibility. In response, Design of Experiments (DoE), aligned with Quality by Design (QbD) principles, and more recently, machine learning (ML) and artificial intelligence (AI), have gained attention as effective complementary approaches to streamline and enhance RP-HPLC method development. This review brings together current understanding of RP-HPLC fundamentals, DoE strategies, ML-based development techniques, and the growing integration of DoE with ML, while also assessing existing challenges and future directions. A narrative synthesis was conducted based on peer-reviewed literature sourced from PubMed, Google Scholar, ScienceDirect, and Springer, supplemented by a focused selection of primary and review articles related to RP-HPLC, DoE, QbD, and ML applications in chromatographic analysis. DoE methodologies including Plackett-Burman, factorial, Box-Behnken, and central composite designs facilitate structured exploration and optimisation of chromatographic conditions within an established Operable Design Region. Meanwhile, ML models such as random forest, support vector machines, and artificial neural networks leverage molecular features and past chromatographic data to predict retention times, resolution, and mobile phase composition. Emerging hybrid workflows combining DoE with ML, along with Bayesian optimisation, are increasingly shown to reduce experimental effort while enhancing method robustness and alignment with regulatory standards. The convergence of DoE and ML marks a significant advancement in RP-HPLC development, yet broader implementation in routine pharmaceutical quality control will depend on resolving issues related to data harmonisation, model transparency, and regulatory validation.

Keywords: RP-HPLC; Design of Experiments; Quality by Design; Machine Learning; Retention time prediction; Chemometrics

1. INTRODUCTION

HPLC is a separation technique that is being extensively used in all fields due to its versatility. It is used for the qualitative and quantitative analysis of the API, impurities, excipients and others in Pharmaceutical, biomedical, food, environmental and forensic laboratories(1,2). RP-HPLC, being a type of HPLC, has a nonpolar stationary phase (C18 or C8) and a polar mobile phase, which has the reproducibility and versatility for most of the analysis ranges(1,3). RP-HPLC is the method among HPLC techniques mostly mentioned in the pharmacopoeia monograph and regulatory submissions(3,4). The traditional method, which is trail and error method, uses a change in the chromatographic parameters, such as pH, temperature and mobile phase etc are changed in a sequence order by holding the other conditions(4,5). Thus, the main concern arises, which is study of the interaction between the parameters becomes impossible. Another setback for this method is that it is time-consuming and resource-intensive(5,6)[5,6]. To overcome the main concern, two upgrades have been made. Those two are DoE and Integration of ML into the chromatographic method development. DoE was initially formalised by Fisher. DOE is a statistical tool embedded within the QbD framework that aids in maintaining product quality throughout the process.(7) Formalised was later upgraded into an advanced version like Box, Behnken,

Plackett, Burman and others(8–11). These extended versions allow us to change multiple factors not randomly but on an experimental matrix. With this analyst can study the main effects and interaction effects. Even with a small set of experimental runs, all these can be made successful (12–15). When it is used as part of the Quality by Design (QbD) approach, following the rules described in the ICH Q8, Q9, Q10, and Q14 guidelines(16–20). Defining the risk assessment, ATP and MODR, this becomes the risk-based framework tool that acts as an engine of the QbD(21–25).

Secondly, the use of AI and ML in the development of chromatographic methods. Large publicly available data sets, libraries that contain the molecular descriptors, and the computational power have made this a strong base for the use of AI & ML(26–30). QSSR models reduce the experimental runs indirectly aids in the reduced solvent usage, manpower, experimental cost, etc. QSSR models are built in a way of algorithm like random forest, SVM/SVR, gradient boosting and ANN/DNN. The work of these model algorithms is to predict the output based on the molecular and instrumental descriptors themselves. Output includes Peak retention time, peak resolution, optimal mobile phase composition (31–38). DOE backs the statistical work, whereas the ML handles the data-driven modelling. Both when infused together for the optimisation (as they are on one Motto), an outperforming optimisation can be achieved. A growing framework that fuses two i. DoE generated training data feeding machine learning surrogate models. ii. Bayesian optimisation guided autonomous method development(27,39,40). Aim of this review is to summarise the (fundamentals of RPHPC on method development and validation) (describe DOE & QBD methodology applied to RP-HPLC) (ML & AI application survey in chromatographic method development, (iv) examine the emerging integration of DoE and ML, and (v) discuss the applications, challenges, and future perspectives of this evolving field, drawing on both foundational and contemporary literature retrieved from PubMed, Google Scholar, and allied databases.

2. Fundamentals of RP-HPLC

2.1 Principle and mechanism: Separation of the mixture in the RP-HPLC is based on the difference in their respective hydrophobic partitioning of the analyte.C18, C8 or phenyl bonded silica are the frequently and most used non-polar stationary phases. Polar mobile phase usually consists of the composition of the buffer/water with the organic modifier. (1,2,41). In RP-HPLC, where the stationary phase is non-polar, thus the analyte that are non-polar (hydrophobic) has affinity towards the stationary phase and take time to elute, or in other words, take longer retention time. Polar/ionisable analyte, being the contrast to the stationary phase, elutes more quickly and takes much less time than the latter. However, it can be modified using pH control or ionising agents (1,42). Classical parameters that govern the overall performance of the system are k'-capacity factor, a-selectivity, Rs-Resolution, and Column efficiency/N-theoretical plates(2,41).

2.2 Instrumentation: Parts of a modern RP-HPLC contain the following parts in the system they are pump, which can elute both the isocratic and the gradient, a sample injector, which can either be manual or an autosampler, a thermostated analytical column, a detector and a data acquisition system. Most commonly used detectors are UV-Vis photodiode array, fluorescence and mass spectrometric. (1,2,43). UPLC is the advanced instrument of the HPLC where 2 µ m particles are employed, and high pressure with shorter analysis time and resolution time is used and considered. They are widely used in pharmaceutical analysis. (6,44,45).

2.3 Method development variables: for optimising or even for developing, choosing the factors that have the key role is mandatory for the output results. Thus, the factors/variables that can influence the performance are organic modifier type and percentage, type of buffer used, concentration, pH and ion pairing agents(mobile phase related), II-column chemistry, particle size, column length and internal diameter (stationary phase related), III-flow rate, column temperature, injection volume, gradient profile, and detector wavelength(41,46–48). All the factors are non-additive, which is that they don't increase with increasing the values, as they are based on the structure of the analyte and the other variables associated with factors, not on the assumption(46,49,50).

2.4 Method validation: Once a candidate RP-HPLC method has been developed, it must be validated according to regulatory guidance, principally ICH Q2(R2)(18) (harmonising with the newly introduced ICH Q14 on analytical procedure development), which specify the validation characteristics summarised in Table 1: specificity/selectivity, linearity and range, accuracy, precision (repeatability, intermediate precision), limit of detection (LOD), limit of quantitation (LOQ), and robustness. Forced-degradation (stress) studies are additionally used to establish stability-indicating capability, where the method is intended to monitor degradation products (51,52).

International Journal of Pharmacy and Pharmaceutical Research (IJPPR)
Volume 32, Issue 7, July 2026 ijppr.humanjournals.com ISSN: 2349-7203

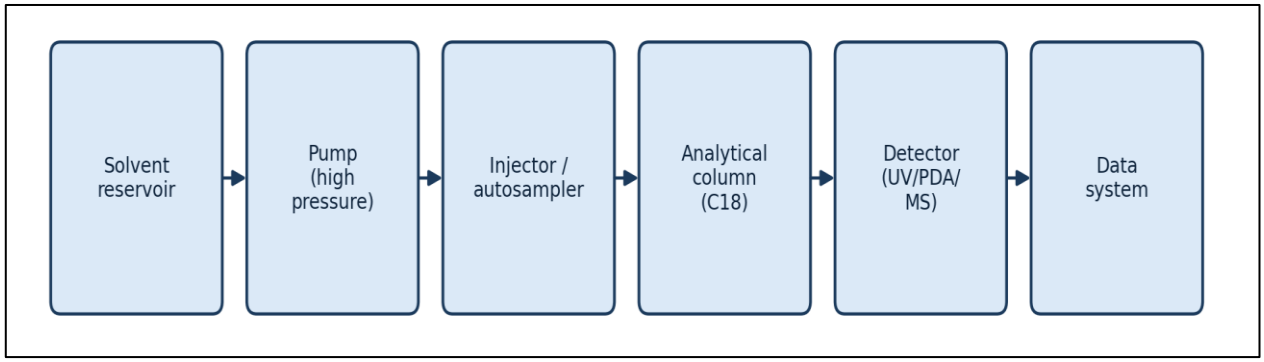


Figure 1:Schematic layout of a modern RP-HPLC system, comprising solvent delivery, injection, column, detection, and data-handling modules.

Table 1.Key ICH Q2(R2)-based validation parameters for RP-HPLC methods

Validation parameter	Definition	Typical acceptance criterion
Specificity/selectivity	Ability to unequivocally assess the analyte in the presence of impurities, degradants, or matrix components	Resolution ≥ 2.0 between critical peak pairs; no interference at analyte retention time
Linearity/range	Proportionality of detector response to analyte concentration over the working range	Correlation coefficient (r) ≥ 0.999 over ≥ 5 concentration levels
Accuracy	Closeness of the measured value to the true (spiked/reference) value	Recovery 98-102% (drug substance); 95-105% (matrices)
Precision (repeatability)	Agreement among replicate measurements under the same conditions	%RSD ≤ 2.0
Intermediate precision	Agreement across analysts, days, or instruments within the same laboratory	%RSD ≤ 2.0-3.0
LOD/LOQ	Lowest concentration reliably detected/quantified	S/N ≈ 3 (LOD); S/N ≈ 10 (LOQ)
Robustness	The capacity of the method to remain unaffected by small, deliberate variations in method parameters	No significant change in system suitability parameters within the design space

3. Design of Experiments (DoE) and Quality by Design in RP-HPLC

3.1 Statistical basis of DoE: DoE is a structured approach to planning experiments so that the effect of multiple independent variables (factors) on one or more dependent responses can be estimated efficiently, including main effects, two-factor interactions and, where required, curvature, using far fewer runs than an exhaustive combinatorial or OFAT search (10,12,14,53). The theoretical foundation traces to Fisher's principles of randomisation, replication, and blocking and to the response-surface methodology introduced by Box and Wilson (9).

3.2 Screening designs: When the number of potentially influential factors is large and it is unclear which are critical, screening (resolution III/IV) designs such as the Plackett-Burman design(11,15) or fractional factorial designs are used to identify the "vital few" factors with a minimum number of experiments, typically N = 8, 12, or 16 runs for up to 7-11 factors (15,54). Factors found to be statistically insignificant (e.g., via Pareto charts or half-normal probability plots) are fixed at practical levels, while significant factors are advanced to optimisation (15,54).

3.3 Optimisation designs: Once the critical factors (typically 2-4) have been identified, response-surface designs are used to model curvature and locate the optimum. The most commonly applied designs in RP-HPLC method development are the central composite design (CCD), Box-Behnken design (BBD), and, less frequently, the Doehlert design(8,49,50). Figure 3 illustrates the geometric arrangement of experimental points for the full factorial, Box-Behnken, and central composite designs. BBD is a spherical, rotatable design that avoids extreme (corner) combinations of factors, which is often desirable in chromatography, where extreme conditions may damage the column or cause co-elution (23,50,55). CCD augments a factorial design with axial (star) points and centre points, permitting estimation of a full second-order (quadratic) polynomial model relating chromatographic responses (e.g., resolution,

tailing factor, retention time, peak area) to the factors under study(43,50,56). Multiple responses are frequently reconciled using desirability-function optimisation as proposed by Derringer and Suich (57), which combines individual response desirabilities into a single composite score to identify a global optimum.

3.4 Analytical Quality by Design (AQbD): QbD, introduced into pharmaceutical development through the ICH Q8(R2), Q9(R1), and Q10 guidelines (16,17,20,55) and more recently extended specifically to analytical procedures in ICH Q14, is "a systematic approach to development that begins with predefined objectives and emphasises product and process understanding and process control, based on sound science and quality risk management"(55) [34]. Its analytical-method-specific counterpart, AQbD, follows a workflow (Figure 2) comprising: (i) definition of the Analytical Target Profile (ATP), (ii) identification of Critical Quality Attributes (CQAs) and Critical Method Parameters (CMPs) through risk assessment tools such as Ishikawa (fishbone) diagrams or Failure Mode and Effects Analysis (FMEA), (iii) a screening DoE to shortlist significant CMPs, (iv) an optimisation DoE to build a predictive response-surface model, (v) definition of the Method Operable Design Region (MODR)/design space within which the method reliably meets the ATP, (vi) establishment of a control strategy, and (vii) formal method validation and lifecycle management (21–24,58–60). Numerous case studies illustrate this workflow for RP-HPLC methods, including AQbD-driven development of stability-indicating assays for buserelin acetate (61), nevirapine and its degradation products (62), luteolin (63), glimepiride/linagliptin (43), fexuprazan hydrochloride (64), and telmisartan/amlodipine/bisoprolol combinations (65), among others (23,66–75).

3.5 Regulatory and risk-management dimension: Beyond method optimisation, AQbD explicitly incorporates quality risk management principles from ICH Q9(R1)(20) [36], enabling analytical scientists to justify the acceptable operating ranges of method parameters to regulators and to build in continuous verification across the method's lifecycle, consistent with ICH Q10 (16) and the newer ICH Q14 analytical lifecycle concept(19,60). This risk-based justification is a key advantage of DoE/QbD over empirically validated OFAT methods, which typically lack a documented multidimensional operating region (24,46,72,76,77).



Figure 2: An eight-step Analytical Quality by Design (AQbD) workflow applied to RP-HPLC method development, from Analytical Target Profile definition through validation and lifecycle management.

International Journal of Pharmacy and Pharmaceutical Research (IJPPR)
Volume 32, Issue 7, July 2026 ijppr.humanjournals.com ISSN: 2349-7203

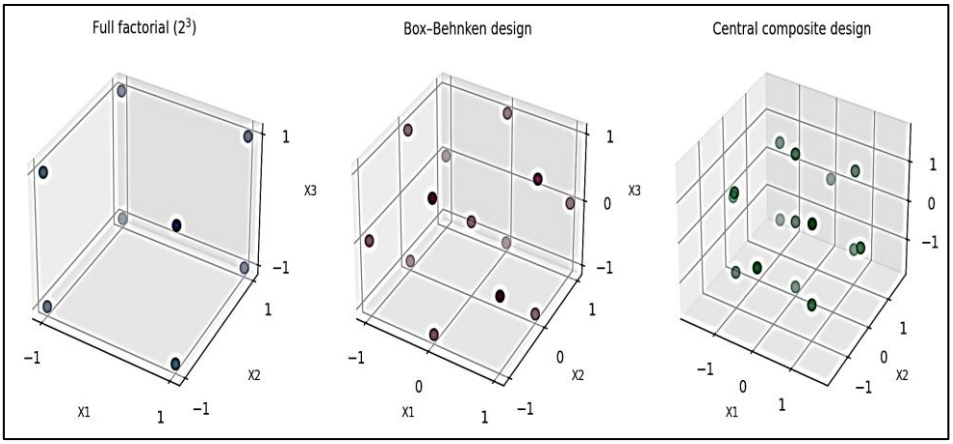


Figure 3: Geometric representation of experimental points for a full factorial (2^3) design, a Box-Behnken design, and a central composite design, the three DoE structures most frequently used to optimise RP-HPLC critical method parameters.

Table 2. Comparative overview of commonly used DoE designs in RP-HPLC method development. Comparative overview of commonly used DoE designs in RP-HPLC method development

Design	Typical purpose	No. of factors	Approx. runs	Key features
Plackett-Burman	Screening	7-11 (or more)	8-16	Resolution III; two-level; estimates main effects only; confounded interactions
Full/fractional factorial	Screening/optimisation	2-5	4-32	Two or more levels; can estimate interactions (full) or main effects with aliasing (fractional)
Box-Behnken design (BBD)	Optimisation	3-5	13-46	Three levels; spherical, rotatable; avoids extreme corner combinations
Central composite design (CCD)	Optimisation	2-5	13-52	Factorial + axial (star) + centre points; full quadratic model; rotatable
D-optimal design	Optimisation (constrained space)	Variable	Variable	Computer-generated; efficient when factor space is irregular or constrained
Mixture design	Formulation/mobile-phase composition	2-4 components	Variable	Used when factors are proportions of a mixture summing to 100%

4. Machine Learning and Artificial Intelligence in RP-HPLC

4.1 Rationale: Even a well-executed DoE campaign is limited to the specific analyte(s), column, and factor ranges studied; each new molecule or matrix, in principle, requires a fresh set of experiments. Machine learning offers a complementary, data-driven route: by learning statistical relationships between molecular/instrumental descriptors and chromatographic outcomes across many historical experiments, ML models can predict retention behaviour, resolution, or suitable mobile-phase compositions for new analytes with markedly fewer, or in some cases no, new experiments (26–30,33).

4.2 Quantitative structure-retention relationships (QSRR): QSRR modelling correlates molecular descriptors (e.g., logP, polar surface area, molecular weight, hydrogen-bond donor/acceptor counts, topological indices) with experimentally observed retention times or retention indices(33,35). Historically, QSRR relied on multiple linear regression and partial least squares; the field has since been extended to non-linear machine-learning algorithms, including random forest (RF), support vector regression (SVR), gradient boosting, extreme gradient boosting (XGBoost), and, most recently, artificial neural networks (ANN), deep neural networks (DNN), convolutional neural networks (CNN), and graph neural networks (GNN) (31,32,34,36,78–84). Random forest, an ensemble of decision trees introduced by Breiman (81), has repeatedly demonstrated favourable robustness to noise and computational efficiency for both spectral and chromatographic regression tasks(26,29), while deep-learning approaches trained on large public

retention-time repositories such as the METLIN small-molecule retention-time (SMRT) dataset (38) and PredRet (85) have achieved mean absolute errors as low as ~0.5 min for structurally diverse compound sets(31,32).

4.3 Retention-time and mobile-phase prediction in pharmaceutical RP-HPLC. Several recent studies apply ML directly to pharmaceutical RP-HPLC problems: random forest regression has been used to predict RP-HPLC retention time from molecular descriptors for drug-like compounds(29); random forest regression has also been benchmarked against bagging CART, PLS, and non-linear SVM for multivariate spectral and chromatographic calibration of cardiovascular drugs including hydrochlorothiazide and valsartan (26); and a machine-learning framework coupling analytical method validation data with predictive modelling has been proposed to forecast suitable mobile-phase composition for RP-HPLC assays(30). Beyond retention time, ML models have also been used to predict chromatographic stationary-phase selectivity across columns (37), to model gradient retention behaviour with ensembles of random forest, gradient boosting, and AdaBoost(86), and to support QSRR-based identification of unknowns in non-targeted analysis(36,53).

4.4 Autonomous and AI-guided method development: Another category of applications employs machine learning not just to generate predictions, but to actively guide the experimental process. For instance, Bayesian optimization algorithms have enabled autonomous optimization of gradient HPLC separations for complex mixtures, achieving effective results in as few as 35 runs(40). This concept has been expanded to a fully automated, operator-free method development for impurity profiling, where the system independently balances resolution and analysis time(39). However, critical assessments note that deterministic retention modelling tools such as commercial gradient optimization software are sometimes incorrectly labelled as "AI," even though they differ fundamentally from true machine learning, deep learning, or reinforcement learning methods(27,28). When comparing AI-predicted HPLC methods with those optimized using traditional design-of-experiments (DoE) approaches for multi-drug formulations, studies indicate that while AI can speed up the initial phase of method development, expert input is still typically necessary to finalise, validate, and refine the method to meet performance and sustainability standards (87).

4.5 Chemometrics and neural networks in pharmaceutical analysis: Artificial neural networks and associated chemometric methods have long been applied in pharmaceutical analysis, with uses going beyond chromatography to include spectrophotometry and multivariate calibration. These approaches laid the conceptual groundwork for modern deep-learning QSRR models (88). Comprehensive overviews of AI in pharmaceutical R&D place chromatographic method development within a broader context, highlighting the increasing integration of artificial intelligence across various stages such as drug discovery, formulation, and quality assurance. (28,89)

Table 3. Machine learning algorithms reported for retention-time/method prediction in liquid chromatography

Algorithm	Type	Reported strengths	Reported limitations
Multiple linear regression (MLR)/PLS	Linear	Simple, interpretable, low data requirement	Cannot capture non-linear retention behaviour
Random forest (RF)	Ensemble (bagging)	Robust to noise, handles non-linearity, low overfitting risk, feature-importance ranking	Less accurate than deep models on very large datasets; limited extrapolation
Support vector machine/regression (SVM/SVR)	Kernel-based	Effective in high-dimensional descriptor space, good generalisation with small datasets	Sensitive to kernel/hyperparameter choice; computationally intensive for large datasets
Gradient boosting/XGBoost	Ensemble (boosting)	High predictive accuracy, handles heterogeneous data well	Prone to overfitting without careful regularisation; less interpretable
Artificial/deep neural networks (ANN/DNN)	Neural network	Captures complex non-linear relationships; scales well with very large datasets (e.g., SMRT)	Requires large training data; limited interpretability (“black box”); risk of overfitting on small pharmaceutical datasets
Graph neural networks (GNN)	Neural network (graph-based)	Learns directly from molecular graph structure; strong performance on structurally diverse sets	Computationally demanding; relatively new to pharmaceutical QC applications

5. Integration of DoE and Machine Learning

DoE and machine learning (ML) tackle the optimization of RP-HPLC methods from complementary perspectives. DoE offers a statistically efficient and mechanistically transparent approach for exploring a defined, limited factor space around a given method, producing high-quality experimental data with minimal noise and well-characterized design features such as orthogonality and rotatability. (8,12,14,49). In contrast, machine learning performs particularly well at leveraging large and diverse historical datasets—covering numerous analytes, columns, and laboratories—to make broad predictions across chemical space. However, it usually requires substantial amounts of data and offers less straightforward insight into underlying mechanisms.(27,31,33). A number of aligned approaches have begun to take shape (Figure 4). One key strategy involves using data generated through Design of Experiments (DoE) as high-quality input for training localized machine learning models or response-surface surrogates. This integrates the experimental efficiency inherent in DoE with the adaptability of machine learning regression techniques. While conceptually similar to traditional response-surface methods, this approach expands on them by incorporating non-linear learning algorithms.(40,49). Second, Bayesian optimization—a sequential machine learning approach that employs a probabilistic surrogate model, typically a Gaussian process, to identify the most informative next experiment—has been applied to optimize multi-linear HPLC gradients for complex mixtures. This method achieves effective results with significantly fewer experimental runs compared to a full factorial design of experiments and has further been adapted to enable fully autonomous, "operator-free" workflows for impurity detection.(39,40). Third, retention time prediction models based on QSRR or deep learning—trained on extensive public datasets—can help identify a promising initial region within the design space (such as an estimated organic modifier concentration or gradient duration). This preliminary selection can guide a subsequent, more focused design of experiments (DoE) on a smaller scale, aimed at refining and verifying the method in accordance with ICH Q2(R2)/Q14 guidelines. (31,32,34,38,60,90). Critical assessments highlight that hybrid DoE-ML approaches maintain the regulatory clarity of traditional Design of Experiments (DoE), offering a traceable and statistically supported design space. At the same time, they cut down the number of experimental trials required to define this space—especially beneficial in complex separations involving multiple analytes or when transferring methods between columns or labs.(27,28,37). At the same time, these reviews note that the interpretability, validation, and regulatory acceptance of machine learning components within a formally structured AQbD submission are still far less developed compared to traditional DoE approaches. They also highlight that existing "AI-assisted" method development tools differ significantly in how much they rely on true machine learning as opposed to deterministic retention modeling algorithms. (27,28,87).

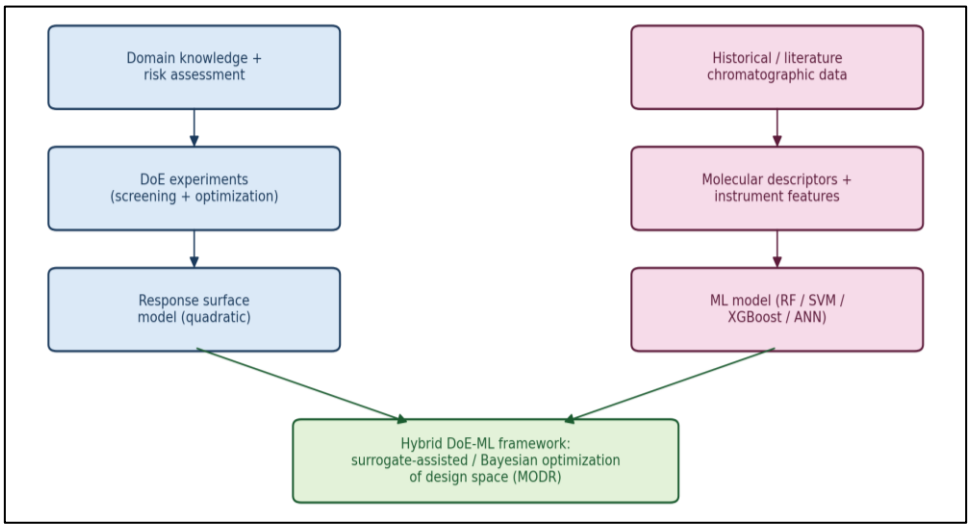


Figure 4: Conceptual convergence of Design of Experiments and machine learning in RP-HPLC method development, illustrating parallel classical (DoE) and data-driven (ML) pipelines feeding a hybrid, surrogate-assisted optimisation of the design space/Method Operable

6. Applications

Table 4 summarises representative published applications of DoE/QbD, ML, and combined approaches to RP-HPLC method development across different analyte classes, illustrating the breadth of the field from small-molecule antihypertensives and antivirals to biologics-adjacent formulations, forced-degradation/impurity profiling, and bioanalytical assays in biological matrices.

Table 4. Representative RP-HPLC method development studies employing DoE/QbD, ML, or hybrid approaches

Analyte(s)/context	Approach	Design/algorithm	Key outcome	Ref.
Glimepiride and linagliptin (synthetic mixture)	AQbD	Central composite design	Robust simultaneous RP-HPLC assay within a validated design space	(43)
Fexuprazan hydrochloride (stability-indicating)	AQbD	Integrated DoE approach	Stability-indicating RP-HPLC method with defined MODR	(64)
Telmisartan, amlodipine besylate, bisoprolol fumarate	AQbD + green chemistry	Box-Behnken design	Simultaneous, environmentally benign RP-HPLC assay	(65)
Buserelin acetate in polymeric nanoparticles	AQbD	Central composite design	Validated RP-HPLC release-testing method	(61)
Nevirapine and degradation products	AQbD	Screening + optimisation DoE	Stability-indicating method with resolved degradants	(62)
Luteolin (fractional factorial)	QbD-driven	Fractional factorial design	Optimised RP-HPLC estimation method	(63)
Aprepitant and EP impurities	Forced degradation + LC-MS/MS	Not DoE-specific (mechanistic)	Toxicity prediction of degradation products	(52)
Linezolid and ethionamide in human plasma	Bioanalytical validation	Conventional development	Validated bioanalytical RP-HPLC method	(91)
Cardiovascular drugs (hydrochlorothiazide, valsartan) - spectral calibration	Machine learning	Random forest regression	RF outperformed/matched PLS and SVM for multivariate calibration	(26)
Generic drug-like compounds	Machine learning	Random forest regression on molecular descriptors	Predicted RP-HPLC retention time from structure	(29)
Mobile-phase composition prediction	Machine learning	ML coupled to method-validation data	Predictive framework for RP-HPLC mobile-phase selection	(30)
Complex dye/impurity mixtures	Autonomous/Bayesian optimisation	Bayesian optimisation (Gaussian process)	Near-optimal gradient separation in ~35 experiments	(39,40)
Amlodipine, hydrochlorothiazide, candesartan	AI vs. DoE comparison	AI-predicted vs. in-lab optimised conditions	In-lab (DoE-refined) method greener and more efficient than raw AI output	(87)

7. Challenges and Limitations

Despite significant advancements, the widespread use of integrated DoE-ML approaches in RP-HPLC method development is still hindered by several challenges. (i) Data standardization and accessibility: Machine learning models—especially deep-learning QSRR models—depend on large volumes of retention time data that are uniformly annotated. However, differences in chromatographic parameters such as column chemistry, mobile phase composition, temperature, and gradient profiles across labs reduce the generalizability of these models, often requiring retraining or recalibration for new settings.(34,38,92). (ii) Model interpretability: Ensemble and deep-learning models often function as "black boxes," making it difficult to provide the mechanistic explanations that regulators require in an AQbD submission. In contrast, classical DoE produces polynomial models that are transparent and statistically interpretable. (28,46,83). (iii) Regulatory pathways for validating AI components remain a work in progress. While ICH Q14 establishes design of experiments (DoE)-based analytical quality by design (AQbD), it currently lacks specific, detailed guidance on machine learning-based method development or ongoing verification driven by AI.(19,60). (iv) The risk of overfitting is significant when working with pharmaceutical RP-HPLC datasets for a given analyte, as these datasets are typically small. This makes high-capacity machine learning models prone to poor generalisation unless measures such as proper regularisation, cross-validation, or transfer learning from larger public datasets are implemented(31,83). (v) There is often confusion between deterministic simulation and authentic artificial intelligence: many commonly used tools for developing "intelligent" methods are based on mechanistic retention models—such as the linear solvent strength theory—rather than on data-driven machine learning. The scientific literature frequently fails to clearly differentiate between these approaches, making it difficult to assess the real contributions of machine learning accurately. (28). (vi) Sustainability considerations: automated systems or AI-generated recommendations do not inherently reduce solvent use or shorten analysis time. Environmental impact must be independently evaluated using assessment tools like the Analytical Eco-Scale, GAPI, or AGREE. Optimization goals should therefore clearly incorporate green or white analytical chemistry criteria in addition to chromatographic performance metrics. (87,93–101).

8. Future Perspectives

Future RP-HPLC method development could become faster and more reliable by using large, standardized, publicly available databases of retention times and chromatographic conditions, allowing researchers to build on existing experimental data instead of starting from scratch. (38,85), make more accurate, consistent, and widely applicable predictions for pharmaceutical RP-HPLC method development. Second, Future HPLC systems may combine chromatographic theory (mechanistic retention models) with machine learning (data-driven corrections) and use Bayesian Optimization to intelligently select experiments and Reinforcement Learning to improve decisions. Such systems can develop and optimize chromatographic methods automatically with little or no human input, while remaining both accurate and interpretable (27,28,39,40). Third, Green and white analytical chemistry principles should be included while optimizing the HPLC method, instead of checking them only after the method is developed. This helps produce methods that are both efficient and environmentally friendly (93–101). Fourth, explainable AI (XAI) techniques (e.g., SHAP values, feature-importance analysis) applied to ML-derived retention or resolution models may show which chromatographic factors have the greatest effect on the prediction and makes the models easier to understand and more acceptable for regulatory review(28,83). Lastly ICH Q14 recommends continuously monitoring and improving analytical methods throughout their entire lifecycle(19,60), Future regulatory guidelines may need to provide clear requirements for the validation, maintenance, and continuous monitoring of ML or AI-assisted analytical methods, similar to current regulatory efforts for AI & ML based medical software.

9. Conclusion

In the pharmaceutical industry, rapid transition has been achieved by the use of RP-HPLC in method development and also in other aspects, as they offer better analytical workload(powerhouse). The concept of DoE and QbD frameworks has been keenly watched, and the importance were given to the methodology as they are statistically clearer and even based on the risk-based design space. They have been complimented and are encouraged with the convergence of ML along with it. Recent advancements, such as Bayesian optimisation and machine learning models, have been seen in many works as they offer better results than the OFAT method. The use of ML alongside DoE is that both are strong candidates in their own right. Combining them offers us a chance to get the synergistic effect of the desired output. relationship between parameters or variables, prediction based on those factors that govern the output (pH, temp, mobile phase concentration, etc.). Though ML usage has several benefits and outranks most of all the methods that are currently in use and used, they rely heavily on the large dataset, data quality and quantity. DoE & ML offers rapid, more sustainable RP-HPLC methods. Consideration must be on the inclusion of large data sets, quality data and data interpretability.

Acknowledgement

We express our sincere thanks to the Department of Pharmaceutical Chemistry, College of Pharmacy, Madras Medical College(MMC), Chennai for providing necessary facilities for the research work.

Conflict of Interest

The authors declare that there is no conflict of interest.

REFERENCES

1. Basak S, Das P. Reverse Phase High-performance Liquid Chromatography: A Comprehensive Review of Principles, Instrumentation, Analytical Procedures, and Pharmaceutical Applications. Journal of Preventive, Diagnostic and Treatment Strategies in Medicine. 2025 Apr;4(2):83–92. doi:10.4103/jpdtsm.jpdtsm_34_25
2. Supriya Hazra1 SB. A Comprehensive Review on High Performance Liquid Chromatography (HPLC): Introduction, Theory, Instrumentation, Advantages, Limitations and Applications [Internet]. 2025 Dec 16. doi:10.5281/ZENODO.17951616
3. Madhurkar PS, Bhawar DSB, Kolhe MH. A REVIEW ON: RP-HPLC. Vol. 10. 10(8).
4. Manasa Padma Meduri, Pooja Agarwal, G. Vimala, Nafiza Banu. The development and validation studies of RP-HPLC method – A review. World J Pharm Sci. 2016 Jan 1;4(1):85–92.
5. Lavesh Jain, Hitesh Kothari, Shaziya Yasmeen, Anju Goyal. Reversed-Phase High-Performance Liquid Chromatography (RP-HPLC) Based Analytical Method Development and Validation for Pharmaceutical Dosage Forms: A Comprehensive Review [Internet]. Zenodo; 2026 [cited 2026 Jul 9]. Available from: https://zenodo.org/doi/10.5281/zenodo.20688311 doi:10.5281/ZENODO.20688311
6. Ankush Bhalerao, Mr. Satish Shelke, Dr. Vijay Borkar. Advances, Applications, and Challenges in RP HPLC Method Development. IJARSCT. 2023 Apr 3;111–23. doi:10.48175/IJARSCT-9017
7. Internet Archive [Internet]. [cited 2026 Jul 14]. Statistical Method For Research Workers : Fisher, R. A : Free Download, Borrow, and Streaming. Available from: https://archive.org/details/in.ernet.dli.2015.205971

8. Box GEP, Behnken DW. Some New Three Level Designs for the Study of Quantitative Variables. Technometrics. 1960 Nov;2(4):455–75. doi:10.1080/00401706.1960.10489912
9. Box GEP, Wilson KB. On the Experimental Attainment of Optimum Conditions. Journal of the Royal Statistical Society Series B: Statistical Methodology. 1951 Jan 1;13(1):1–38. doi:10.1111/j.2517-6161.1951.tb00067.x
10. Montgomery DC. Design and analysis of experiments. Tenth edition. Hoboken, NJ: Wiley; 2020. 1 p.
11. Plackett RL, Burman JP. THE DESIGN OF OPTIMUM MULTIFACTORIAL EXPERIMENTS. Biometrika. 1946;33(4):305–25. doi:10.1093/biomet/33.4.305
12. Dhoot A, Fernandes G, Naha A, Rathnanand M, Kumar L. Design of Experiments in Pharmaceutical Development. Pharmaceutical Chemistry Journal. 2019 Dec 4;53. doi:10.1007/s11094-019-02070-4
13. Fukuda IM, Pinto CFF, Moreira CDS, Saviano AM, Lourenço FR. Design of Experiments (DoE) applied to Pharmaceutical and Analytical Quality by Design (QbD). Braz J Pharm Sci. 2018 Nov 8;54(spe). doi:10.1590/s2175-97902018000001006
14. N. Politis S, Colombo P, Colombo G, M. Rekkas D. Design of experiments (DoE) in pharmaceutical development. Drug Development and Industrial Pharmacy. 2017 Jun 3;43(6):889–901. doi:10.1080/03639045.2017.1291672
15. Vanaja K, Shobha Rani RH. Design of Experiments: Concept and Applications of Plackett Burman Design. Clinical Research and Regulatory Affairs. 2007 Jan;24(1):1–23. doi:10.1080/10601330701220520
16. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. ICH Q10: Pharmaceutical Quality System [Internet]. Geneva, Switzerland: International Council for Harmonisation (ICH); 2008 Jun. Report no.: Q10. Available from: https://www.ich.org/page/quality-guidelines
17. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. ICH Q8(R2): Pharmaceutical Development [Internet]. Geneva, Switzerland: International Council for Harmonisation (ICH); 2009 Aug. Report Q8(R2). Available from: https://www.ich.org/page/quality-guidelines
18. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. ICH Q2(R2): Validation of Analytical Procedures [Internet]. Geneva: International Council for Harmonisation (ICH); 2023. Report Q2(R2). Available from: https://www.ich.org/page/quality-guidelines
19. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. ICH Q14: Analytical Procedure Development [Internet]. Geneva, Switzerland: International Council for Harmonisation (ICH); 2023 Nov. Report no.: Q14. Available from: Official ICH Q14 Guideline (PDF)
20. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. ICH Q9(R1): Quality Risk Management [Internet]. Geneva, Switzerland: International Council for Harmonisation (ICH); 2023 Jan. Report Q9(R1). Available from: https://www.ich.org/page/quality-guidelines
21. Bairagi A, Kothrukar R, Chikhale H, Kosanam S, Borse L. AQbD-novel strategy for analytical methods. Futur J Pharm Sci. 2024 Oct 4;10(1):138. doi:10.1186/s43094-024-00706-1
22. Katekar V, Sangule D, Bhurbhure O, Ingle P, Dhage S, Jadhav K. A Review on Quality by Design Approach in Analytical Methods. J Drug Delivery Ther. 2022 Jun 15;12(3-S):255–61. doi:10.22270/jddt.v12i3-S.5386
23. Peraman R, Bhadraya K, Padmanabha Reddy Y. Analytical Quality by Design: A Tool for Regulatory Flexibility and Robust Analytics. International Journal of Analytical Chemistry. 2015;2015:1–9. doi:10.1155/2015/868727
24. Rozet E, Lebrun P, Michiels JF, Sondag P, Scherder T, Boulanger B. Analytical Procedure Validation and the Quality by Design Paradigm. Journal of Biopharmaceutical Statistics. 2015 Mar 4;25(2):260–8. doi:10.1080/10543406.2014.971176
25. Vogt FG, Kord AS. Development of Quality-By-Design Analytical Methods. Journal of Pharmaceutical Sciences. 2011 Mar;100(3):797–812. doi:10.1002/jps.22325
26. Aggarapu S, Galla R, Shaik NJ, Dasari SJ. Application of Plackett–Burman and Box–Behnken designs to develop a sensitive and robust HPLC method for quantifying hypolipidemic drugs, rosuvastatin and bempedoic acid in tablets. Futur J Pharm Sci. 2025 May 26;11(1):61. doi:10.1186/s43094-025-00814-6
27. Alves E, Gurupadayya BM, Prabhakaran P. Artificial Intelligence in HPLC Method Development: A Critical Review of Technological Integration, Limitations, and Future Directions. Critical Reviews in Analytical Chemistry. 2025 Oct 28;1–43. doi:10.1080/10408347.2025.2575352
28. Dixon TM, Williams J, Besenhard M, Howard RM, MacGregor J, Peach P, et al. Operator-free HPLC automated method development guided by Bayesian optimization. Digital Discovery. 2024;3(8):1591–601. doi:10.1039/D4DD00062E
29. Ghasemi JB, Tavakoli H. Application of random forest regression to spectral multivariate calibration. Anal Methods. 2013;5(7):1863. doi:10.1039/c3ay26338j
30. Selvi VG, Urs AR, B TC. Retention Time Prediction in High-Performance Liquid Chromatography Using Random Forest Regression [Internet]. In Review; 2025 [cited 2026 Jul 9]. Available from: https://www.researchsquare.com/article/rs-6053718/v1 doi:10.21203/rs.3.rs-6053718/v1
31. Bouwmeester R, Martens L, Degroeve S. Generalized Calibration Across Liquid Chromatography Setups for Generic Prediction of Small-Molecule Retention Times. Anal Chem. 2020 May 5;92(9):6571–8. doi:10.1021/acs.analchem.0c00233
32. Fedorova ES, Matyushin DD, Plyushchenko IV, Stavrianidi AN, Buryak AK. Deep learning for retention time prediction in reversed-phase liquid chromatography. Journal of Chromatography A. 2022 Feb;1664:462792. doi:10.1016/j.chroma.2021.462792

33. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning [Internet]. New York, NY: Springer New York; 2009 [cited 2026 Jul 9]. (Springer Series in Statistics). Available from: http://link.springer.com/10.1007/978-0-387-84858-7 doi:10.1007/978-0-387-84858-7
34. Héberger K. Quantitative structure–(chromatographic) retention relationships. Journal of Chromatography A. 2007 Jul;1158(1–2):273–305. doi:10.1016/j.chroma.2007.03.108
35. Kaliszan R. QSRR: Quantitative Structure-(Chromatographic) Retention Relationships. Chem Rev. 2007 Jul 1;107(7):3212–46. doi:10.1021/cr068412z
36. Wen Y, Talebi M, Amos RIJ, Szucs R, Dolan JW, Pohl CA, et al. Retention prediction in reversed phase high performance liquid chromatography using quantitative structure-retention relationships applied to the Hydrophobic Subtraction Model. Journal of Chromatography A. 2018 Mar 16;1541:1–11. doi:10.1016/j.chroma.2018.01.053
37. Xie J, Chen S, Zhao L, Dong X. Application of artificial intelligence to quantitative structure–retention relationship calculations in chromatography. Journal of Pharmaceutical Analysis. 2025 Jan;15(1):101155. doi:10.1016/j.jpha.2024.101155
38. Yang Q, Ji H, Lu H, Zhang Z. Prediction of Liquid Chromatographic Retention Time with Graph Neural Networks to Assist in Small Molecule Identification. Anal Chem. 2021 Feb 2;93(4):2200–6. doi:10.1021/acs.analchem.0c04071
39. Boelrijk J, Ensing B, Forré P, Pirok BWJ. Closed-loop automatic gradient design for liquid chromatography using Bayesian optimization. Analytica Chimica Acta. 2023 Feb;1242:340789. doi:10.1016/j.aca.2023.340789
40. Shi Z, Yi Y, Madrigal E, Hrovat F, Zhang K, Lin J. A generalizable methodology for predicting retention time of small molecule pharmaceutical compounds across reversed-phase HPLC columns. Journal of Chromatography A. 2025 Feb;1742:465628. doi:10.1016/j.chroma.2024.465628
41. Snyder LR, Dolan JW, Gant JR. Gradient elution in high-performance liquid chromatography: I. Theoretical basis for reversed-phase systems. Journal of Chromatography A. 1979 Mar 21;165(1):3–30. doi:10.1016/S0021-9673(00)85726-X
42. HPLC Columns, Theory, Technology, and Practice: U. D. Neue, Wiley-VCH Publishers, New York, 1997, 393 pp., $89.95. Instrumentation Science & Technology. 1998 Sep;26(4):439–40. doi:10.1080/10739149808001913
43. Mehta H, Shaileshbhai PA. Center Composite Design Assisted Optimization of RP-HPLC Method for Simultaneous Estimation of Glimepiride and Linagliptin in Synthetic Mixture Using Aqbd Perspective [Internet]. In Review; 2026 [cited 2026 Jul 9]. Available from: https://www.researchsquare.com/article/rs-9947795/v1 doi:10.21203/rs.3.rs-9947795/v1
44. Borman P, Campa C, Delpierre G, Hook E, Jackson P, Kelley W, et al. Selection of Analytical Technology and Development of Analytical Procedures Using the Analytical Target Profile. Anal Chem. 2022 Jan 18;94(2):559–70. doi:10.1021/acs.analchem.1c03854
45. Pal* M, Singh R, Singh DA. WJPR - Abstract. 1 [Internet]. [cited 2026 Jul 14]. Available from: https://www.wjpr.net/abstract_show/29793
46. Butcher DJ. The Essence of Chromatography By Colin F. Poole (Wayne State University). Elsevier: Amsterdam and New York. 2003. x + 926 pp. $89.00. ISBN 0-444-50199-1 (Paperback). J Am Chem Soc. 2003 Aug 1;125(33):10144–5. doi:10.1021/ja033507t
47. Lavesh Jain, Hitesh Kothari, Shaziya Yasmeen, Anju Goyal. Reversed-Phase High-Performance Liquid Chromatography (RP-HPLC) Based Analytical Method Development and Validation for Pharmaceutical Dosage Forms: A Comprehensive Review [Internet]. Zenodo; 2026 [cited 2026 Jul 7]. Available from: https://zenodo.org/doi/10.5281/zenodo.20688311 doi:10.5281/ZENODO.20688311
48. Manasa Padma Meduri, Pooja Agarwal, G. Vimala, Nafiza Banu. The development and validation studies of RP-HPLC method – A review. World J Pharm Sci. 2016 Jan 1;4(1):85–92.
49. Bezerra MA, Santelli RE, Oliveira EP, Villar LS, Escaleira LA. Response surface methodology (RSM) as a tool for optimization in analytical chemistry. Talanta. 2008 Sep 15;76(5):965–77. doi:10.1016/j.talanta.2008.05.019
50. Ferreira SLC, Bruns RE, Ferreira HS, Matos GD, David JM, Brandão GC, et al. Box-Behnken design: An alternative for the optimization of analytical methods. Analytica Chimica Acta. 2007 Aug;597(2):179–86. doi:10.1016/j.aca.2007.07.011
51. Gurumukhi VC, Bari SB. Quantification and Validation of Stability-Indicating RP-HPLC Method for Efavirenz in Bulk and Tablet Dosage Form using Quality by Design (QbD): A Shifting Paradigm. Journal of Chromatographic Science. 2022 Feb 16;60(2):143–56. doi:10.1093/chromsci/bmab061
52. Ravi Babu K, Reddipalli GS. VALIDATED HPLC METHOD FOR SIMULTANEOUSQUANTIFICATION OF APREPITANT ANDITSEPIMPURITIES WITH LC-MS/MS ELUCIDATIONANDTOXICITY PREDICTION OF ITS FORCEDDEGRADATIONPRODUCTS. RJC. 2026;19(01):231–43. doi:10.31788/RJC.2026.1919524
53. Manager GS. Statistical Experimental Design and its Application to Pharmaceutical Development Problems. Drug Development and Industrial Pharmacy. 1986 Jan;12(8–9):1109–23. doi:10.3109/03639048609065848
54. Deidda R, Orlandini S, Hubert P, Hubert C. Risk-based approach for method development in pharmaceutical quality control context: A critical review. Journal of Pharmaceutical and Biomedical Analysis. 2018 Nov;161:110–21. doi:10.1016/j.jpba.2018.07.050
55. Yu LX. Pharmaceutical Quality by Design: Product and Process Development, Understanding, and Control. Pharm Res. 2008 Apr;25(4):781–91. doi:10.1007/s11095-007-9511-1

56. Suyash Gupta, Preeti Verma, Abhinav Prasoon Mishra, Nainshi Omar, Riya Mathur. A Review on Novel Analytical Method Development and Validation by RP-HPLC Method. Indian Journal of Forensic Medicine & Toxicology. 2021 Sep 5;15(4):3479–86. doi:10.37506/ijfmt.v15i4.17798
57. Derringer G, Suich R. Simultaneous Optimization of Several Response Variables. Journal of Quality Technology. 1980 Oct;12(4):214–9. doi:10.1080/00224065.1980.11980968
58. Karmarkar S, Garber R, Genchanok Y, George S, Yang X, Hammond R. Quality by Design (QbD) Based Development of a Stability Indicating HPLC Method for Drug and Impurities. Journal of Chromatographic Science. 2011 Jul 1;49(6):439–46. doi:10.1093/chrsci/49.6.439
59. Little TA. Design of experiments for analytical method development and validation. Vol. 27. 2014 Jan 1;27:56–60.
60. Prajapati PB, Patel A, Shah SA. Risk and DoE-Based DMAIC Principle to the Multipurpose-RP-HPLC Method for Synchronous Estimation of Anti-Hypertensive Drugs Using AQbD Approach. Journal of AOAC INTERNATIONAL. 2021 Sep 27;104(5):1442–52. doi:10.1093/jaoacint/qsab079
61. Tambare RS, Shahi SR, Gurumukhi VC, Kakade SM, Tapadiya GG. Quality by design (QbD) based development and validation of RP-HPLC method for buserelin acetate in polymeric nanoparticles: Release study. Heliyon. 2024 Oct;10(20):e39172. doi:10.1016/j.heliyon.2024.e39172
62. Jesus Do Nascimento Lopes ID, Kazumasa Fujimori S, De Carvalho Mendes T, Adrielle Dias De Almeida R, Furtado De Mendonça De Sousa F, Areias De Oliveira C, et al. Application of analytical quality by design (AQbD) for stability-indicating method development for quantification of nevirapine and its degradation products. Microchemical Journal. 2024 Apr;199:109939. doi:10.1016/j.microc.2024.109939
63. Patil VA, Mali AJ. QbD-Guided RP-HPLC Method for Luteolin Featuring Full Factorial Design and Response Surface Optimization. J Pharm Innov. 2026 Jun;21(3):264. doi:10.1007/s12247-026-10558-2
64. Mehta H, Mhaske S, Tandel P. Integrated Design of Experiments Approach for a Stability-indicating Rp-hplc Method of Fexuprazan Hydrochloride [Internet]. In Review; 2026 [cited 2026 Jul 9]. Available from: https://www.researchsquare.com/article/rs-9946357/v1 doi:10.21203/rs.3.rs-9946357/v1
65. Ghughuskar P, Musale R, Mehta H, Tandel P. AQbD-Assisted Development and Validation of a Green RP-HPLC Method for Simultaneous Quantification of Telmisartan, Amlodipine Besylate, and Bisoprolol Fumarate [Internet]. In Review; 2026 [cited 2026 Jul 9]. Available from: https://www.researchsquare.com/article/rs-10025123/v1 doi:10.21203/rs.3.rs-10025123/v1
66. Bangera PD, Keerikkadu M, Tippavajhala VK, Rathnanand M. ANALYTICAL QBD INTEGRATED RP-HPLC METHOD DEVELOPMENT AND OPTIMIZATION FOR HIGH-PRECISION QUANTIFICATION OF IBRUTINIB IN HYBRID NANOCARRIER SYSTEMS. Int J App Pharm. 2025 Sep 7;361–77. doi:10.22159/ijap.2025v17i5.54683
67. Bhatt D, Rane SI. QbD approach to analytical RP-HPLC method development and its validation. International Journal of Pharmacy and Pharmaceutical Sciences. 2011 Jan 1;3:179–87.
68. Fan Y, Deng Y, Yang Y, Deng X, Li Q, Xu B, et al. Modelling and predicting liquid chromatography retention time for PFAS with no-code machine learning. Environ Sci: Adv. 2024;3(2):198–207. doi:10.1039/D3VA00242J
69. Gustavo González A, Ángeles Herrador M. A practical guide to analytical method validation, including measurement uncertainty and accuracy profiles. TrAC Trends in Analytical Chemistry. 2007 Mar;26(3):227–38. doi:10.1016/j.trac.2007.01.009
70. Hashem H, El-Sayed HM. Quality by design approach for development and validation of a RP-HPLC method for simultaneous determination of co-administered levetiracetam and pyridoxine HCl in prepared tablets. Microchemical Journal. 2018 Dec;143:55–63. doi:10.1016/j.microc.2018.07.031
71. Karageorgou EG, Kalogiouri NP, Samanidou VF. Green Approaches in High-Performance Liquid Chromatography for Sustainable Food Analysis: Advances, Challenges, and Regulatory Perspectives. Molecules. 2025 Aug 31;30(17):3573. doi:10.3390/molecules30173573
72. Krishna MV, Dash RN, Jalachandra Reddy B, Venugopal P, Sandeep P, Madhavi G. Quality by Design (QbD) approach to develop HPLC method for eberconazole nitrate: Application oxidative and photolytic degradation kinetics. Journal of Saudi Chemical Society. 2016 Sep;20:S313–22. doi:10.1016/j.jscs.2012.12.001
73. Prajapati PB, Jayswal K, Shah SA. Application of Quality Risk Assessment and DoE-Based Enhanced Analytical Quality by Design Approach to Development of Chromatography Method for Estimation of Combined Pharmaceutical Dosage Form of Five Drugs. Journal of Chromatographic Science. 2021 Aug 18;59(8):714–29. doi:10.1093/chromsci/bmaa118
74. Tome T, Žigart N, Časar Z, Obreza A. Development and Optimization of Liquid Chromatography Analytical Methods by Using AQbD Principles: Overview and Recent Advances. Org Process Res Dev. 2019 Sep 20;23(9):1784–802. doi:10.1021/acs.oprd.9b00238
75. Zhang L, Mao S. Application of quality by design in the current drug development. Asian Journal of Pharmaceutical Sciences. 2017 Jan 1;12(1):1–8. doi:10.1016/j.ajps.2016.07.006
76. Musters J, Van Den Bos L, Kellenbach E. Applying QbD Principles To Develop a Generic UHPLC Method Which Facilitates Continual Improvement and Innovation Throughout the Product Lifecycle for a Commercial API. Org Process Res Dev. 2013 Jan 18;17(1):87–96. doi:10.1021/op300292a

77. Sahu PK, Ramisetti NR, Cecchi T, Swain S, Patro CS, Panda J. An overview of experimental designs in HPLC method development and validation. Journal of Pharmaceutical and Biomedical Analysis. 2018 Jan;147:590–611. doi:10.1016/j.jpba.2017.05.006
78. Breiman L. Random Forests. Machine Learning. 2001 Oct;45(1):5–32. doi:10.1023/A:1010933404324
79. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining [Internet]. San Francisco California USA: ACM; 2016 [cited 2026 Jul 9]. p. 785–94. Available from: https://dl.acm.org/doi/10.1145/2939672.2939785 doi:10.1145/2939672.2939785
80. Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995 Sep;20(3):273–97. doi:10.1007/BF00994018
81. Jagani K, Singh YR, Patel E, Jani R, Shah S, Patel P, et al. Coupling analytical method validation with machine learning for mobile phase prediction in RP-HPLC. Discov Chem. 2026 Jun 27;3(1):371. doi:10.1007/s44371-026-00828-w
82. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015 May 28;521(7553):436–44. doi:10.1038/nature14539
83. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature. 1986 Oct;323(6088):533–6. doi:10.1038/323533a0
84. Wiczling P, Kamedulska A. Comparison of Chromatographic Stationary Phases Using a Bayesian-Based Multilevel Model. Anal Chem. 2024 Jan 23;96(3):1310–9. doi:10.1021/acs.analchem.3c04697
85. Domingo-Almenara X, Guijas C, Billings E, Montenegro-Burke JR, Uritboonthai W, Aisporna AE, et al. The METLIN small molecule dataset for machine learning-based retention time prediction. Nat Commun. 2019 Dec 20;10(1):5811. doi:10.1038/s41467-019-13680-7
86. Lotfy H, Erk N, Genc A, Obaydo R, Tiris G. Artificial Intelligence in Chromatography: Greenness and Performance Evaluation of AI-Predicted versus In-Lab Optimized HPLC Methods for Simultaneous Separation of Amlodipine, Hydrochlorothiazide, and Candesartan. Talanta Open. 2025 May 1;11:100473. doi:10.1016/j.talo.2025.100473
87. Xu H, Lin J, Zhang D, Mo F. Retention time prediction for chromatographic enantioseparation by quantile geometry-enhanced graph neural network. Nat Commun. 2023 May 29;14(1):3095. doi:10.1038/s41467-023-38853-3
88. Agatonovic-Kustrin S, Beresford R. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. Journal of Pharmaceutical and Biomedical Analysis. 2000 Jun;22(5):717–27. doi:10.1016/S0731-7085(99)00272-1
89. Kalariya PD, Namdev D, Srinivas R, Gananadhamu S. Application of experimental design and response surface technique for selecting the optimum RP-HPLC conditions for the determination of moxifloxacin HCl and ketorolac tromethamine in eye drops. Journal of Saudi Chemical Society. 2017 Jan;21:S373–82. doi:10.1016/j.jscs.2014.04.004
90. Beg S, Sharma G, Katare OP, Lohan S, Singh B. Development and Validation of a Stability-Indicating Liquid Chromatographic Method for Estimating Olmesartan Medoxomil Using Quality by Design. J Chromatogr Sci. 2015 Aug;53(7):1048–59. doi:10.1093/chromsci/bmu165
91. Vyavahare R, Soni U. DEVELOPMENT AND VALIDATION OF A REVERSE PHASE HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY (RP-HPLC) BASED BIOANALYTICAL METHOD FOR THE ESTIMATION OF LINEZOLID AND ETHIONAMIDE IN HUMAN PLASMA. JAB. 11(14).
92. Stanstrup J, Neumann S, Vrhovšek U. PredRet: Prediction of Retention Time by Direct Mapping between Multiple Chromatographic Systems. Anal Chem. 2015 Sep 15;87(18):9421–8. doi:10.1021/acs.analchem.5b02287
93. Ahmed M, Abdullah, Eiman E, Al-Ahmary KM, Aftab F, Sohail A, et al. Advances in green liquid chromatography for pharmaceutical analysis: A comprehensive review on analytical greenness to sustainable chemistry approaches. Microchemical Journal. 2024 Oct;205:111400. doi:10.1016/j.microc.2024.111400
94. Gałuszka A, Migaszewski Z, Namieśnik J. The 12 principles of green analytical chemistry and the SIGNIFICANCE mnemonic of green analytical practices. TrAC Trends in Analytical Chemistry. 2013 Oct;50:78–84. doi:10.1016/j.trac.2013.04.010
95. Mehta M, Mehta D, Mashru R. Recent application of green analytical chemistry: eco-friendly approaches for pharmaceutical analysis. Futur J Pharm Sci. 2024 Jul 8;10(1):83. doi:10.1186/s43094-024-00658-6
96. Nowak PM, Wietecha-Posłuszny R, Pawliszyn J. White Analytical Chemistry: An approach to reconcile the principles of Green Analytical Chemistry and functionality. TrAC Trends in Analytical Chemistry. 2021 May;138:116223. doi:10.1016/j.trac.2021.116223
97. Oliveira M, Yoshida M, Belinelo V, Valotto R. Degradation Kinetics of Atorvastatin under Stress Conditions and Chemical Analysis by HPLC. Molecules. 2013 Jan 24;18(2):1447–56. doi:10.3390/molecules18021447
98. Pena-Pereira F, Wojnowski W, Tobiszewski M. AGREE—Analytical GREEnness Metric Approach and Software. Anal Chem. 2020 Jul 21;92(14):10076–82. doi:10.1021/acs.analchem.0c01887
99. Płotka-Wasylka J. A new tool for the evaluation of the analytical procedure: Green Analytical Procedure Index. Talanta. 2018 May;181:204–9. doi:10.1016/j.talanta.2018.01.013
100. Saleh SS, Obaydo RH, El Hamd MA, Rostom Y, Mohamed D, Lotfy HM. Guidelines for accurate application of green and white analytical concepts: Merits Versus demerits with insights of significant milestones of assessment tools applied for antiviral drugs. Microchemical Journal. 2024 Apr;199:109917. doi:10.1016/j.microc.2024.109917
101. Sulthana H, Jays J, B PK, Goudanavar P. Green analytical evaluation of anticancer drug analysis: A multi-tool assessment of HPLC and LC–MS methods. Sci Rep. 2026 Jan 13;16(1):3803. doi:10.1038/s41598-025-33868-w

How to cite this article:
Pandiyan Govindasamy et al. Ijppr.Human, 2026; Vol. 32 (7):410-423.

Conflict of Interest Statement: All authors have nothing else to disclose.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.