

Artificial Neural Network (ANN)-Based Optimization of Chitosan Nanoparticles Loaded with Curcumin for Enhanced Oral Bioavailability: A Machine Learning Assisted Quality by Design (QbD) Approach

Nitin Shekapure¹, Dr. Pramod Kumar Soni², Dr. Shaina³, Ms. Sapna Satija⁴, Ms. Pramila Chandel⁵, Mr. Rishu Singh⁶

¹Associate Professor, Department of Robotics and Automation, AISSMS COE, Pune

Email: ngshekapure@aiissmscoe.com

²Assistant Professor, Department of Computer Science and Engineering, Sharda University, Greater Noida, UP, India

Email: pramod.apr10@gmail.com

³Associate Professor, Department of AI & DS, IIMT College of Engineering, Greater Noida

Email: shaina.arora22@gmail.com

⁴Assistant Professor, Department of BCA, JIMS Engineering Technical Campus, Greater Noida, UP, India

Email: sapna_3023@yahoo.com

⁵Assistant Professor, Department of BCA, JIMS Engineering Technical Campus, Greater Noida, UP, India

Email: pramilachandel01@gmail.com

⁶Assistant Professor, School of Computer Applications, SDGI Global University

Email: rishusingh261119@gmail.com

Received: 5th January, 2026; Revised: 18th January 2026; Accepted: 25th January, 2026; Available Online: 28th January, 2026

ABSTRACT

Curcumin is a polyphenol molecule with pharmaceutical potential whose oral bioavailability is limited by its relatively low water solubility, chemical instability under certain conditions, high metabolic clearance, and lack of systemic bioavailability. This study introduced a machine-learning based Quality by Design (QbD) approach to computational optimization of curcumin loaded chitosan nanoparticles. The primary Dataset consisted of 10,000 observations from 9,000 design groups, 11 formulation/process predictors, and 11 critical quality attributes (CQAs). Given that there was no explicit provenance field in the workbook and that all of the observation identifiers were in the SIM convention, the dataset was treated conservatively as computational/simulated-like, not experimentally measured. The group-aware split generated 6,980 training observations, 1,496 validation observations, and 1,524 independent test observations with no overlap in the design group. Optimized and trained a multi-output PyTorch Artificial Neural Network (ANN) with Optuna on NVIDIA Tesla T4 with CUDA 12.8. The final architecture consisted of one 192-unit GELU hidden layer, dropout 0.1782, AdamW optimization, and a five-seed equal-weight ensemble. Independent-test R^2 values were 0.969 for drug loading, 0.877 for particle size, 0.855 for zeta potential, 0.768 for entrapment efficiency, 0.395 for PDI, and 0.407-0.478 for the six cumulative release endpoints. The ANN ensemble yielded the lowest test RMSE for all 11 CQAs among models tested in this study, and the moderate R^2 values for PDI and release endpoints showed some predictive uncertainty remained. Although the work shows a reproducible computational strategy on candidate formulation prioritization and a definition of a model-derived design region, experimental CQA confirmation and pharmacokinetic evaluation would be required before any claim of enhanced oral bioavailability could be made.

Keywords: Curcumin; chitosan nanoparticles; artificial neural network; Quality by Design; machine learning; ionic gelation; multi-objective optimization; drug delivery.

How to cite this article: Nitin Shekapure¹ Dr. Pramod Kumar Soni² Dr. Shaina³ Ms. Sapna Satija⁴ Ms. Pramila Chandel⁵ Mr. Rishu Singh⁶. Artificial Neural Network (ANN)-Based Optimization of Chitosan Nanoparticles Loaded with Curcumin for Enhanced Oral Bioavailability: A Machine Learning Assisted Quality by Design (QbD) Approach. *Int J Drug Deliv Technol.* 2026;16(85s):782-789. DOI: 10.25258/ijddt.16.85s.78.

Source of support: Nil.

Conflict of interest: None

1. INTRODUCTION

1.1 Background of the Study

Curcumin is a polyphenolic compound with a broad range of studies, although translation to oral dosage forms is limited due to its low aqueous solubility, instability under certain conditions, rapid

metabolism, and consequently poor systemic exposure (Anand et al., 2007; Nelson et al., 2017). Therefore, research is conducted on nanoparticle systems for improving dispersion, protecting the drug cargo, modifying the release, and enhancing interaction with gastrointestinal environment.

Chitosan is of special relevance, since it is a biodegradable and cationic polysaccharide that has been reported to be useful in particulate drug delivery systems and can be used to generate nanoparticles via ionic interaction with tripolyphosphate (TPP) (Agnihotri et al., 2004; Bernkop-Schnürch and Dünnhaupt, 2012). The preparation of curcumin-loaded chitosan nanoparticles has also been experimentally established, validating the possible use of this carrier-drug combination (Khan et al., 2016).

It is a multivariate formulation problem. The interactions of chitosan concentration, molecular weight and degree of deacetylation, TPP concentration, polymer-crosslinker and drug-polymer ratio, ethanol content, pH, stirring conditions, addition rate and mixing time can affect particle size, polydispersity index (PDI), zeta potential, entrapment efficiency, drug loading and drug release behavior. Quality by Design (QbD) offers a structured approach to organizing these inputs as Critical Material Attributes (CMAs) and Critical Process Parameters (CPPs), and then relating them to Critical Quality Attributes (CQAs) (ICH, 2009; Rawal et al., 2019). However, standard response models may not capture nonlinear interactions across a wide formulation space. Artificial neural networks (ANNs) are a flexible surrogate-modeling approach that has a long history of application in pharmaceutical formulation optimization (Takayama et al., 1999; Wang et al., 2022).

1.2 Problem Statement

It is difficult to optimize several QAs of a nanoparticle simultaneously, since making one response better may make another worse and the relationships may be nonlinear. Predictive evaluation may also be overly optimistic when replicates or closely related design conditions are distributed between training and test data. A strict workflow is then required, with either leakage-safe group partition, preprocessing done only after training, separate testing, comparison of models, interpretability, and separation between model-derived optimization and experimental validation. The present work solves this problem by computational approach instead of assuming any pharmacokinetic advantage by formulation CQAs alone.

1.3 Objectives of the Study

The aims were to:

- (i) Structure formulation/process variables and CQAs in a QbD context;
- (ii) Develop and independently test a multi-output ANN;
- (iii) Use model-interpretability analyses to identify influential formulation/process variables;
- (iv) Run multi-objective Pareto optimization; and

- (v) Select an ANN-predicted formulation candidate and a model-derived design region for experimental validation.

1.4 Research Questions

How well can the ANN predict the modeled CQAs for future design groups? What are the most important formulation/process variables to predict? Is the ANN a good surrogate model for traditional baseline regressors? What are the presumed formulation conditions that are likely to lead to a favorable multi-CQA compromise without experiments?

2. REVIEW OF LITERATURE

2.1 Curcumin and Chitosan Nanoparticles

Curcumin is a compound with broad biological interest, but poor biopharmaceutical attributes. Low water solubility and fast transformation and/or metabolism are cited repeatedly as challenges for oral delivery (Anand et al., 2007; Hewlings and Kalman, 2017; Nelson et al., 2017). Although favorable particle properties can be demonstrated, encapsulation of nanoparticles is one way to modify the physicochemical context in which curcumin is dispersed and released. Chitosan is mucoadhesive and cationic, and has been applied to the micro- and nanoparticulate delivery of drugs and macromolecules (Agnihotri et al., 2004; Bernkop-Schnürch and Dünnhaupt, 2012). The formulation concept is further supported by the curcumin loaded chitosan nanoparticles reported by Khan et al. (2016) who highlighted that the size and surface charge, loading, and release should not be considered in isolation.

2.2 Ionic Gelation, QbD, and Formulation Variables

Electrostatic interaction between protonated chitosan and multivalent TPP is usually adopted for ionic gelation. The interaction between the chains is influenced by the polymer concentration and molecular properties, the ratio of chitosan to TPP affects the network formation, the pH affects the ionization of chitosan, and the stirring speed, addition speed, and post addition mixing influence the dispersion evolution. These factors can thus influence the formation of the particles as well as drug incorporation and release. By connecting a TPP to CQAs and correlating the properties of the materials and process parameters that can influence them, QbD formalizes these relationships (ICH, 2009). This framework can be useful in nanotechnology product development because it is not just a single endpoint (Rawal et al., 2019; Grangeia et al., 2020).

The key CQAs are also interdependent for nanoparticle products. While smaller particles may be beneficial for dispersion and interaction with surfaces, excessive crosslinking or processing may change PDI, surface charge, entrapment, and release. There is a relationship but not equivalence between entrapment efficiency and drug loading: one formulation may achieve a high entrapment efficiency and yet load a relatively small amount of drug per total mass of nanoparticles. Similarly, a

positive zeta potential can help with colloidal behavior and interaction with negatively charged mucosal surfaces, and must be considered in conjunction with size distribution and release. These compromises make the multi-response optimization strategy more appropriate than optimizing each endpoint individually.

2.3 ANN and Machine-Learning-Assisted Formulation Optimization

ANNs can also approximate nonlinear multivariate functions without assuming a fixed polynomial structure. Pharmaceutical applications span from initial formulation optimization to modern approaches based on QbD, including the prediction of complex product responses (Takayama et al., 1999; Simões et al., 2020; Wang et al., 2022). For example, Hashad et al. (2016) used neural-network techniques to optimize the formulation variables of chitosan-TPP nanoparticles. However, the usefulness of ANN modeling is dependent on proper validation. Hyperparameters should be tuned based on training/validation data, the independent test set should be kept until final evaluation, and model importance should be understood as predictive association, not causality. The present group-aware, multi-seed, baseline compared workflow was motivated by these principles.

2.4 Research Gap

While existing literature supports curcumin nanocarriers, chitosan-TPP formulation, QbD, and ANN-based optimization, the present study integrates all of these into a single reproducible computational framework that explicitly separates design groups during testing, evaluates all modeled CQAs, quantifies multi-seed stability, compares conventional baselines, and carries the frozen ANN forward into Pareto optimization and design-space exploration. The contribution is thus methodological and formulation rather than proof of experimentally improved oral bioavailability.

3. RESEARCH METHODOLOGY

3.1 Research Design and Dataset

The study used Computer-aided QbD-ANN design. The primary dataset included 10,000 observations and 25 columns—three Identifiers/Grouping Fields, 11 scientifically meaningful columns of formulation/process predictors, and 11 CQAs. These predictors were chitosan stock concentration, chitosan molecular weight, degree of deacetylation, TPP stock concentration, chitosan:TPP mass ratio, curcumin:chitosan mass ratio, final ethanol percentage, chitosan-phase pH, stirring speed, TPP addition rate, and post-addition mixing time. The CQAs were particle size, PDI, zeta potential, entrapment efficiency, drug loading and cumulative release at 1, 2, 4, 6, 12 and 24 h. The observations lacked a dedicated provenance field or metadata sheet, but all observation identifiers employed the SIM convention. Therefore, this information was considered to be a computational development data, and not confirmed measured laboratory data.

3.2 QbD Classification, Validation, and Data Splitting

Material/ formulation variables were categorized as CMAs, process variables as CPPs, and the 11 measured response variables as CQAs (Table 1). Validated dimensions, data types, missing values, duplicate rows and IDs, replicate structure, numerical ranges, impossible or infinite values, and monotonicity of cumulative release. There were no missing values, no duplicate rows, and release profiles were monotonic with the checks performed. There were 500 groups that had triplicate observations, giving a total of 9,000 design groups. All records sharing design_group_id were placed in a single partition to avoid replicate leakage. The final split comprised 6,980 training observations, 1,496 validation observations and 1,524 independent-test observations.

Table 1. QbD classification of modeled variables

QbD category	Variables
CMA / formulation-material variables	Chitosan stock concentration; chitosan molecular weight; degree of deacetylation; TPP stock concentration; chitosan:TPP mass ratio; curcumin:chitosan mass ratio; final ethanol % v/v
CPPs	Chitosan-phase pH; stirring speed; TPP addition rate; post-addition mixing time
CQAs	Particle size; PDI; zeta potential; entrapment efficiency; drug loading; cumulative release at 1, 2, 4, 6, 12 and 24 h

3.3 Preprocessing, ANN Development, and Evaluation

Only the training partition was used for the preprocessing. The same transformation path was applied to new formulations after predictors and targets were scaled independently on their separate training sets. The main model used was a PyTorch feed-forward multi-output regressor using 11 inputs and 11 outputs. Optuna conducted 40 hyperparameter trials with training/validation data only. Selected architecture included one hidden layer of 192 units using the GELU activation function with dropout of 0.178244 and no batch normalization. The learning rate was set to 0.00298609, the weight decay was set to 2.073×10^{-8} , batch size was 64, early stopping patience was 45 epochs, and the maximum number of epochs was 400. Final training is performed with seeds 42, 1337, 2026, 31415, and 8675309; the prediction is an average of all the seeds, rather than the best one.

To verify GPU execution, PyTorch 2.11.0+cu128 with CUDA 12.8 on NVIDIA Tesla T4 was used. Independent-test performance was summarized using R^2 , MAE, RMSE, MSE, MAPE when applicable and using mean error. Baselines were trained on the same

split as linear regression, random forest, and extra trees. Interpretability unified permutation importance, integrated gradients, and one-variable sensitivity analysis. The analyses were used to describe model dependence of predictors, not causal effects.

For the leakage audit, the following identifiers were considered non-predictive and were kept only for grouping, split construction and replicate checks: `observation_id`, `design_group_id`, `replicate_no`. This ensured that the network did not learn random identifiers that could represent network design membership. This independent test set was not used to select hyperparameters or architecture. These controls are important because if repeated design conditions were not present, test performance might look better than it would on actual unseen formulation groups. To prevent reporting a single favorable initialization and to obtain a simple estimate of model stability, multiple-seed training was performed.

3.4 Multi-Objective Optimization and Design-Space Analysis

A 1,500-trial Optuna NSGA-II search was performed using the frozen five-seed ANN ensemble as a surrogate. The particle size and PDI were reduced, and the zeta potential, entrapment efficiency and drug loading were increased. No endpoints were forced to achieve any unsupported therapeutic target; only physically plausible monotonic 0-100% cumulative release was required. Using a nearest-neighbor distance threshold from the validation data, candidates were selected to be similar to the training domain. The feasible Pareto set was reduced to the final compromise using minimum normalized distance to the multivariate utopia point. For each of the two cases, 50,000 combinations were sampled in each training range to visualize an exploratory model derived design region. Those visualizations were based on data - but not on regulatory or externally set product specifications.

4. RESULTS AND DISCUSSION

4.1 Dataset Profile and ANN Development

The Dataset contained 10,000 observations over 9,000 design groups, and the 11 CQAs were modeled along with 11 predictors. The clear structural audit and separation of data into groups eliminated the apparent data-quality and replicate-leakage problems. These records should not be treated as definitive experimental measurements because there was no formal documentation of the provenance and SIM identifiers were universal. The group-aware test set still gave an independent evaluation within the same test set, for a different set of design groups.

A relatively shallower ANN (one 192-unit GELU layer and moderate dropout) was favored over a deep network in hyperparameter optimization. This makes sense if the data is tabular with structure such as a table, in which further depth does not necessarily lead to better generalization. Five seeds were trained to give a stability check and empirical ensemble uncertainty estimate. We also checked computational

reproducibility following execution, that is, re-loading the model using seed-42, the maximum absolute difference is 0.0.

4.2 Predictive Performance and Baseline Comparison

Independent-test performance was materially different by CQA (Table 2). The drug loading showed the best agreement ($R^2 = 0.9689$ and RMSE = 0.5876 percentage points), followed by particle size ($R^2 = 0.8774$ and RMSE = 27.04 nm), zeta potential ($R^2 = 0.8551$ and RMSE = 2.09 mV), and entrapment efficiency ($R^2 = 0.7681$ and RMSE = 3.81 percentage points). PDI was more difficult ($R^2 = 0.3946$; RMSE 0.0439). Release responses were also only moderately accounted for, ranging from $R^2 = 0.4072$ at 24 h to $R^2 = 0.4775$ at 1 h. This pattern rejects the common description of the ANN as a generally good performer: the ANN explained drug incorporation and some particle properties well, but the release dynamics and PDI showed significant unexplained variation.

The ANN ensemble still achieved the lowest independent-test RMSE for each of the four model classes for the same CQA. For particle size, the RMSE for ANN was 27.04 nm, which was lower than that of the linear regression (31.99 nm), random forest (33.06 nm), and extra trees (32.49 nm). For zeta potential, ANN RMSE was 2.09 mV versus 2.14, 4.34, and 4.39 mV, respectively. For some weaker ones, differences were less: ANN RMSE was 0.0439 for PDI, while the baseline RMSE ranged between 0.0454 and 0.0467. The ANN thus gave a stable but target-specific predictive gain, not a universal gain that can be considered near perfect prediction.

Table 2. Independent-test performance of the five-seed ANN ensemble

CQA	R^2	MAE	RMSE
Particle size (nm)	0.8774	20.8260	27.0397
PDI	0.3946	0.0355	0.0439
Zeta potential (mV)	0.8551	1.6640	2.0925
Entrapment efficiency (%)	0.7681	2.8803	3.8059
Drug loading (%)	0.9689	0.4202	0.5876
Release 1 h (%)	0.4775	1.9086	2.4394
Release 2 h (%)	0.4456	2.7968	3.5569
Release 4 h (%)	0.4211	3.9980	5.0190
Release 6 h (%)	0.4133	4.7021	5.8526
Release 12 h (%)	0.4082	5.2888	6.5226
Release 24 h (%)	0.4072	4.3095	5.3765

Artificial Neural Network (ANN)-Based Optimization of Chitosan Nanoparticles Loaded with Curcumin for Enhanced Oral Bioavailability: A Machine Learning Assisted Quality by Design (QbD) Approach

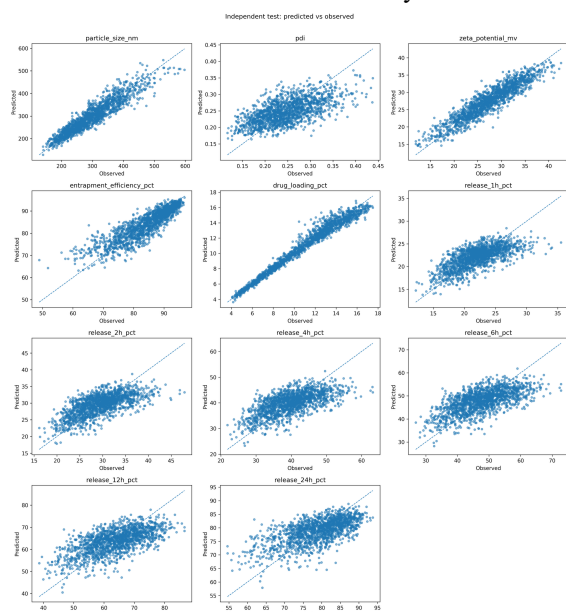


Figure 1. Independent-test predicted-versus-observed plots for the 11 CQAs. Agreement was strongest for drug loading, particle size, zeta potential, and entrapment efficiency.

4.3 Model Interpretation

Permutation importance showed that the most significant and global factors were the mass ratio of curcumin:chitosan, chitosan:TPP mass ratio, chitosan-phase pH, TPP concentration, and chitosan molecular weight. At the target level more specific relationships were observed: TPP stock concentration was the most significant for prediction of particle size, chitosan-phase pH was most important for prediction of zeta-potential, curcumin:chitosan ratio was the most important factor for prediction of entrapment efficiency and drug loading, and chitosan:TPP ratio was the most important factor for all release endpoints. Sensitivity analysis also indicated nonlinear response patterns around the reference formulation. Within the ionic-gelation framework, these findings are scientifically plausible but still model associations. They should be used to guide formulation experiments, not as evidence that increasing one factor will lead to the predicted response in isolation.

Interpretability matched the physical formalization problem. Importance of curcumin:chitosan ratio for drug loading and entrapment corresponds to direct contribution of relative drug input to the modeled response, and the prominence of chitosan:TPP ratio and TPP concentration for size and release are consistent with their direct involvement in the formation of the ionic network. The dominant predictor of zeta potential was chitosan-phase pH, which is consistent with the pH-dependent protonation of chitosan amino groups. But the patterns of the ANN do reflect correlations found in the primary dataset that should be used to inform controlled experiments, not replace mechanistic verification.

4.4 Multi-Objective Optimization and Predicted Formulation

The NSGA-II search yielded 435 non-dominated solutions. 331 Pareto solutions survived nearest-neighbor domain screening, and all Pareto candidates met the cumulative-release consistency check. By normalized distance to the multivariate utopia point, the trade-off between size, PDI, zeta potential, entrapment efficiency, and drug loading was maintained instead of imposing an arbitrary single weighted score. This is significant because no external validated CQA specification or therapeutic release target was provided with the primary Dataset. The formulation selected for the ANN was optimized to 2.8969 mg/mL chitosan, 53.14 kDa molecular weight, 84.09% degree of deacetylation, 0.2696 mg/mL TPP, chitosan:TPP ratio 7.8410, curcumin:chitosan ratio 0.1896, 8.831% v/v ethanol, pH 4.670, 1,029 rpm stirring, TPP addition 0.3830 mL/min, and 50.37 min post-addition mixing (Table 3). The ensemble predicted particle size 207.69 ± 3.57 nm, PDI 0.2035 ± 0.0021 , zeta potential $+35.23 \pm 0.24$ mV, entrapment efficiency $84.25 \pm 0.44\%$, and drug loading $14.20 \pm 0.20\%$. The cumulative release was predicted at 25.07% at 1 h and 84.70% at 24 h, increasing monotonously with time. The uncertainty values quoted are standard deviations over the five ANN seeds, not experimental uncertainty.

Table 3. ANN-predicted optimal formulation and predicted CQAs

Categor y	Variable	Value	Ensembl e SD
Input	Chitosan stock (mg/mL)	2.8969	—
Input	Chitosan MW (kDa)	53.1429	—
Input	Chitosan DDA (%)	84.0926	—
Input	TPP stock (mg/mL)	0.2696	—
Input	Chitosan:TPP mass ratio	7.8410	—
Input	Curcumin:chitosan mass ratio	0.1896	—
Input	Final ethanol (% v/v)	8.8314	—
Input	Chitosan-phase pH	4.6703	—
Input	Stirring speed (rpm)	1029	—
Input	TPP addition (mL/min)	0.3830	—

Category	Variable	Value	Ensemble SD
Input	Post-addition mix (min)	50.3720	—
Predicted CQA	Particle size (nm)	207.6946	3.5722
Predicted CQA	PDI	0.2035	0.0021
Predicted CQA	Zeta potential (mV)	35.2314	0.2374
Predicted CQA	Entrapment efficiency (%)	84.2507	0.4407
Predicted CQA	Drug loading (%)	14.2049	0.1971
Predicted CQA	Release 1 h (%)	25.0654	0.1107
Predicted CQA	Release 2 h (%)	33.8225	0.1686
Predicted CQA	Release 4 h (%)	45.8508	0.2361
Predicted CQA	Release 6 h (%)	54.6778	0.2782
Predicted CQA	Release 12 h (%)	71.1265	0.3031
Predicted CQA	Release 24 h (%)	84.7046	0.2326

Note. All CQA values are model-derived predictions. The optimum has not been experimentally confirmed.

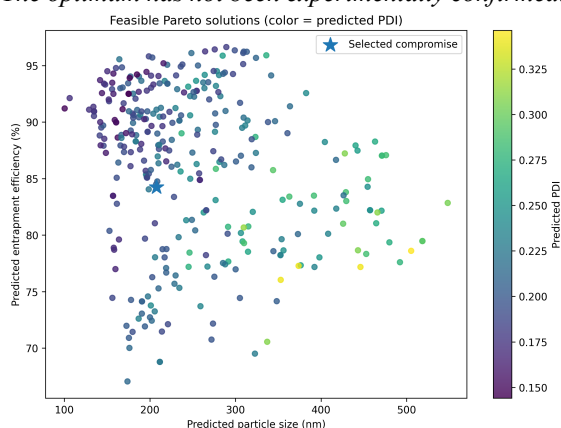


Figure 2. In-domain Pareto solutions from multi-objective ANN optimization. The selected compromise balances the five modeled optimization objectives and is not an experimentally validated optimum.

4.5 Model-Derived Design Space and Oral-Delivery Implications

The design space exploration considered 50,000 combinations within the training ranges. Of these, 45,958 passed the nearest neighbour in-domain

screen based on validation data, and all produced physically monotonic release profiles. When additional median based CQA filters for visualization were applied, 559 samples were considered feasible. The region is an exploratory model-derived design space, not a regulatory or experimentally verified design space, as these filters were internally derived from the primary dataset and not an externally justified QTPP. It is a useful tool for finding areas in which it is necessary to conduct confirmation experiments, and not to extrapolate beyond its training domain.

Selected candidate has a predicted size at the nanoscale, relatively narrow predicted PDI, a positive surface charge, significant amount of curcumin incorporation, and a prolonged monotonic release. These properties are important for formulation development, and the mucoadhesive property of chitosan is supported in the literature (Bernkop-Schnürch and Dünnhaupt, 2012). The primary dataset did not, however, have any AUC data, Cmax data, Tmax data, plasma concentration time data, or other pharmacokinetic endpoints. The study therefore only provides support for a model-predicted formulation profile that can be considered favorable for investigational oral delivery. The oral bioavailability hypothesis has not been confirmed by experimental or in-vivo studies.

The design-space exercise also helps to clarify the distinction between the interpolation and evidence of robustness. Only a small number of the sampled points satisfied the exploratory response filters, but most were sufficiently near to the training domain. This narrowing is helpful for prioritization, but is not enough to define an operating range, since no experimental batches were produced in the proposed range. Therefore, the confirmation should not only involve the one compromise formulation, but also carefully selected neighboring conditions to assess the reliability of the predicted response surface in the neighborhood.

5. CONCLUSIONS, LIMITATIONS AND FUTURE SCOPE

5.1 Conclusions

This study developed a leakage-aware, GPU accelerated QbD-ANN framework for curcumin loaded chitosan nanoparticles. Group-aware partitioning and training-only preprocessing led to a clear evaluation of the model's performance, and five-seed training, baseline comparisons, interpretability, and independent testing reinforced that assessment. The ANN was comparably successful in predicting drug loading, particle size, zeta potential, and entrapment efficiency, and was less successful at predicting release endpoints and PDI. The ratio of curcumin:chitosan, chitosan:TPP, pH, and TPP concentration turned out to be important predictive variables, while chitosan molecular weight had no significant relationship. Multi-objective Pareto analysis resulted in a model-derived candidate with predicted particle size of 207.69 nm, PDI 0.2035, zeta potential of +35.23 mV, entrapment

efficiency of 84.25%, drug loading of 14.20%, and monotonic 24-h release. The framework offers a rational approach to prioritize the formulation experiments but the optimized formulation and its oral delivery implications are still predictions until experiments confirm them.

5.2 Limitations of the Study

The major drawbacks are the computational nature of the primary dataset, the questionable transposition of the derived optimum to external experiments, and the lack of pharmacokinetic or direct oral-bioavailability data. A moderate performance for PDI and release endpoints also suggests some unexplained variability remaining. Design space filters derived from the model are not product specifications but rather assumptions made based on the data.

5.3 Future Scope of Research

Future work should involve experimentally preparing the predicted formulation and comparing the observed sets of CQAs with the ANN predictions and uncertainty ranges. Formulations should be tested for external validation across the proposed region, followed by stability, robustness and scale up studies independently. Oral exposure should be directly evaluated in biorelevant release, permeability and in-vivo pharmacokinetic studies, if physicochemical performance has been confirmed. Weaker PDI and release models may also benefit from expanded experimentally verified datasets.

REFERENCES

1. Anand, P., Kunnumakkara, A. B., Newman, R. A., & Aggarwal, B. B. (2007). Bioavailability of curcumin: Problems and promises. *Molecular Pharmaceutics*, 4(6), 807-818. <https://doi.org/10.1021/mp700113r>
2. Hewlings, S. J., & Kalman, D. S. (2017). Curcumin: A review of its effects on human health. *Foods*, 6(10), 92. <https://doi.org/10.3390/foods6100092>
3. Nelson, K. M., Dahlin, J. L., Bisson, J., Graham, J., Pauli, G. F., & Walters, M. A. (2017). The essential medicinal chemistry of curcumin. *Journal of Medicinal Chemistry*, 60(5), 1620-1637. <https://doi.org/10.1021/acs.jmedchem.6b00975>
4. Agnihotri, S. A., Mallikarjuna, N. N., & Aminabhavi, T. M. (2004). Recent advances on chitosan-based micro- and nanoparticles in drug delivery. *Journal of Controlled Release*, 100(1), 5-28. <https://doi.org/10.1016/j.jconrel.2004.08.010>
5. Calvo, P., Remuñán-López, C., Vila-Jato, J. L., & Alonso, M. J. (1997). Novel hydrophilic chitosan-polyethylene oxide nanoparticles as protein carriers. *Journal of Applied Polymer Science*, 63(1), 125-132. [https://doi.org/10.1002/\(SICI\)1097-4628\(19970103\)63:1<125::AID-APP13>3.0.CO;2-4](https://doi.org/10.1002/(SICI)1097-4628(19970103)63:1<125::AID-APP13>3.0.CO;2-4)
6. Bernkop-Schnürch, A., & Dünnhaupt, S. (2012). Chitosan-based drug delivery systems. *European Journal of Pharmaceutics and Biopharmaceutics*, 81(3), 463-469. <https://doi.org/10.1016/j.ejpb.2012.04.007>
7. Khan, M. A., Zafaryab, M., Mehdi, S. H., Ahmad, I., & Rizvi, M. M. A. (2016). Characterization and anti-proliferative activity of curcumin loaded chitosan nanoparticles in cervical cancer. *International Journal of Biological Macromolecules*, 93(Pt A), 242-253. <https://doi.org/10.1016/j.ijbiomac.2016.08.050>
8. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). (2009). Q8(R2): Pharmaceutical Development. ICH Harmonised Tripartite Guideline.
9. Rawal, M., Singh, A., & Amiji, M. M. (2019). Quality-by-Design concepts to improve nanotechnology-based drug development. *Pharmaceutical Research*, 36(11), 153. <https://doi.org/10.1007/s11095-019-2692-6>
10. Grangeia, H. B., Silva, C., Simões, S. P., & Reis, M. S. (2020). Quality by design in pharmaceutical manufacturing: A systematic review of current status, challenges and future perspectives. *European Journal of Pharmaceutics and Biopharmaceutics*, 147, 19-37. <https://doi.org/10.1016/j.ejpb.2019.12.007>
11. Takayama, K., Fujikawa, M., & Nagai, T. (1999). Artificial neural network as a novel method to optimize pharmaceutical formulations. *Pharmaceutical Research*, 16(1), 1-6. <https://doi.org/10.1023/A:1011986823850>
12. Hashad, R. A., Ishak, R. A. H., Fahmy, S., Mansour, S., & Geneidi, A. S. (2016). Chitosan-tripolyphosphate nanoparticles: Optimization of formulation parameters for improving process yield at a novel pH using artificial neural networks. *International Journal of Biological Macromolecules*, 86, 50-58. <https://doi.org/10.1016/j.ijbiomac.2016.01.042>
13. Simões, M. F., Silva, G., Pinto, A. C., Fonseca, M., Silva, N. E., Pinto, R. M. A., & Simões, S. (2020). Artificial neural networks applied to quality-by-design: From formulation development to clinical outcome. *European Journal of Pharmaceutics and Biopharmaceutics*, 152, 282-295. <https://doi.org/10.1016/j.ejpb.2020.05.012>
14. Wang, S., Di, J., Wang, D., Dai, X., Hua, Y., Gao, X., Zheng, A., & Gao, J. (2022). State-of-the-art review of artificial neural networks to predict, characterize and optimize pharmaceutical formulation. *Pharmaceutics*, 14(1), 183. <https://doi.org/10.3390/pharmaceutics14010183>