Scientific Reports

https://doi.org/10.1038/s41598-026-35795-w

Article in Press

Artificial neural network as a strategy to predict rheological properties in emulgel formulations

Received: 12 June 2025

Accepted: 8 January 2026

Published online: 11 January 2026

Cite this article as: Duarte L.S., Molano L., Jiménez R.A. *et al.* Artificial neural network as a strategy to predict rheological properties in emulgel formulations. *Sci Rep* (2026). https://doi.org/10.1038/s41598-026-35795-w

Laura Sofia Duarte, Laura Molano, Ronald Andrés Jiménez & James Guevara-Pulido

ARTIFICIAL NEURAL NETWORK AS A STRATEGY TO PREDICT RHEOLOGICAL PROPERTIES IN EMULGEL FORMULATIONS

Laura Sofia Duarte,[1] Laura Molano,[1] Ronald Andrés Jiménez,[1] James Guevara-Pulido. *[1]
[1]INQA, Química Farmacéutica, Universidad El Bosque, Bogotá Colombia.
*Corresponding author joguevara@unbosque.edu.co

Abstract
This study presents an innovative approach that combines Quality by Design (QbD) principles with artificial neural networks (ANNs) to predict and optimize the formulation of carbopol-based emulsions. By integrating these two strategies, we enhance our understanding of formulations by linking critical material attributes and process parameters to critical quality attributes, such as viscosity.

The predictive model was refined by selecting key variables: mixing time, mixing speed, and viscosity. These variables were used to estimate carbopol concentration and to capture the nonlinear relationships that influence emulsion behavior. Experimental data were employed to train, validate, and test the ANN model, which was then compared with four commercial formulations to evaluate its predictive accuracy and practical relevance.

Notably, the model demonstrated excellent predictive performance for systems with viscosities exceeding 50,000 mPas, underscoring its applicability to high-viscosity pharmaceutical products. This integrated QbD-ANN framework offers a systematic and effective method for formulation optimization, reducing experimental workloads while improving process understanding.

The findings indicate a strong correlation between predicted and experimental values, confirming the robustness and reliability of the QbD-ANN approach. Integrating the three key variables enables a more in-depth examination of the interactions between process and formulation, providing a comprehensive tool for understanding and controlling emulsion viscosity.

In conclusion, this study establishes a data-driven methodology that facilitates rational pharmaceutical development, ensuring product quality, reproducibility, and innovation in alignment with modern pharmaceutical quality management principles.

Keywords: QbD, ANNs, Carbopol emulsions, pharmaceutical formulation optimization, Predictive modeling, Viscosity and process parameters.

Introduction

Pharmaceutical Quality by Design (QbD) is a systematic approach to development that begins with predefined objectives and emphasizes product and process understanding and control based on sound science and quality risk management [1]. This approach enables identifying and understanding critical quality attributes (CQAs), ensuring the final product is reproducible and robust. Using experimental design and risk analysis, QbD supports evidence-based decision-making, which helps reduce variability and enhance process efficiency. It also allows for establishing a design space that provides regulatory flexibility, demonstrating scientifically that changes within this range do not compromise product quality[1].

Combining QbD approach with artificial neural networks (ANNs) marks a significant advancement in creating emulgel-type pharmaceutical formulations [2-5]. QbD provides a structured way to characterize essential quality attributes and process parameters [6], while ANNs enhance the prediction of physicochemical properties by analyzing nonlinear relationships and uncovering complex interactions [7].

ANNs are now essential in pharmaceutical sciences, as they can analyze vast amounts of data and identify complex patterns that traditional methods struggle to comprehend [8-9]. Their uses include optimizing formulations [10-11], drug stability in formulations [12], viscosity[13], designing new drugs[14-15], predicting stability[16], and personalizing treatments in precision medicine[17]. Specifically, in emulgels, these networks enable modeling the effects of several variables on viscosity, providing a more accurate means to predict and adjust crucial parameters, leading to reproducible and effective formulations [4].

Emulgels represent a sophisticated category of semi-solid systems that merge the characteristics of both gels and emulsions. They offer notable benefits in stability, controlled drug release, and convenient topical application [18]. These systems are extensively employed in dermatological and pharmaceutical products, enhancing the bioavailability of active substances and serving as an effective matrix for incorporating lipophilic compounds [19-20]. Their rheological structure can be accurately adjusted through formulation parameters and more recently using predictive models [21-23], making them a versatile choice for optimizing products, especially in topical and transdermal therapies. Integrating QbD and ANNs presents a groundbreaking method to enhance efficiency and predictability in formulating these systems, creating new opportunities for their use in the pharmaceutical sector.

This study aimed to create a predictive model using artificial intelligence to optimize the formulation of carbopol emulgels by assessing polymer concentration based on key process variables. By integrating QbD principles with ANNs, we can improve the precision of predicting rheological properties. This approach establishes a solid methodological groundwork for employing machine learning in the rational design of pharmaceutical products.

Materials and methods

Materials

Analytical-grade reagents purchased from certified suppliers ((Sigma-Aldrich, USA), (Merck, Germany) (Lubrizol, distributed by Sigma-Aldrich) and were used to prepare the emulgel.

Methods

QbD Elements

Using risk management, the definition of QbD elements, CQAs, and CPPs was performed. Potential Failure Mode and Effects Analysis (FMEA) matrices were constructed to determine criticality in severity, detectability, and occurrence, thereby calculating the probable risk number.

Preparation of emulgels

The oil phase was prepared by dispersing vitamin E in mineral oil, with Tween 80, used as a surfactant to enhance emulsification. This mixture was gradually incorporated into the aqueous phase under moderate stirring, resulting in a stable emulsion. Next, Carbopol 940, was added while the system was continuously stirred until the polymer was uniformly dispersed throughout. The emulgel was then

formed by gradually adding triethanolamine a neutralizing agent to promote gel formation and adjust the system's pH. Finally, the formulation was evaluated for pH, viscosity, physical stability, and homogeneity. Once these assessments were completed, the emulgel was packaged in appropriate containers for storage and future functional evaluation.

Characterization of emulgels

Rheology (Viscosity, Storage-Loss Modulus)

The emulgels' viscoelastic characterization was conducted using an Anton Paar rheometer featuring a standard 5 mm diameter cone-plate geometry. An amplitude sweep was performed to determine the linear viscoelastic region (LVR), examining how the storage modulus (G') and loss modulus (G") varied with the applied strain. The complex viscosity (η) was calculated through an oscillatory sweep at a constant angular frequency of 10 rad/s, covering a shear rate range of 0.1 to 100 s^{1}, with 21 logarithmically spaced measurement points.

Particle size

A LiteSizer Malvern brand DLS (Dynamic Light Scattering) particle size analyzer was used to assess the potential. For both analyses, successive dilutions of the emulgels were prepared using distilled water.

Particle size analysis was performed with an Omega cuvette measuring cell (Mat. No. 225288) at a backscatter angle. The measurements took place at a stable temperature of 25 °C. Six runs, each lasting 10 seconds, were conducted to calculate the average hydrodynamic diameter (z-average) and the polydispersity index (PDI), which helped characterize the droplet size distribution of the formulations.

The zeta potential was measured using the same cell (Omega cuvette, Mat. No. 225288) at 25 °C, with a 1-minute equilibration before each measurement. A voltage of 200.0 V and the Smoluchowski model, incorporating a Henry factor of 1.5, were employed to calculate the zeta potential. Each sample underwent 100 runs.

INQA-ANN Predictive Model Neural Network

Data set preparation

The data used to train the neural network were derived from the experimental design of eleven emulgel formulations. In this design, three process factors were systematically varied: (1) Carbopol® concentration (% w/w), (2) mixing speed (in rotations per second), and (3) mixing time (in minutes). This variation aimed to assess their influence on the viscosity of the system. These variables were selected based on the Failure Modes and Effects Analysis (FMEA).

The experimental viscosity of each formulation was measured in triplicate to ensure reproducibility. Following this, the data underwent statistical analysis, which included calculations of percentage error and standard deviation among the replicates. The average viscosity value for each formulation was then determined and used as representative data for training the predictive model.

Predicted model

The rheological properties of emulgel formulations were predicted utilizing a validated artificial neural network (ANN) previously outlined by Guevara-Pulido et al. (2022) [24]. For training this neural network with a backpropagation algorithm, the most significant factors identified in the Potential Failure Mode and Effects Analysis (FMEA) matrix were selected as input variables: mixing time and speed (critical process parameters, CPP), as well as Carbopol concentration (critical material attribute, CMA). Based on Pearson correlations, a systematic screening was conducted in MATLAB to define these parameters as model inputs, with the viscosity of the emulgels set as the output variable.

During a later optimization stage, modifications to the neural network's structure revealed that the best predictive results were achieved by considering Carbopol concentration (% w/w) as the output variable and using mixing time (min), mixing speed (rpm), and viscosity (Pa · s) as input variables.

Ultimately, the model was validated through leave-one-out cross-validation, and its goodness of fit was assessed using the coefficient of determination (R^2). To improve model efficacy, the nodes in the hidden layer were tested from 50 to 500 in increments (SI) of 50 nodes in each trial to obtain satisfactory R^2 values, ideally near 1, ensuring thorough validation. The selected model parameters were then applied to predict the Carbopol® concentration in the formulated emulgels.

Linear Regression Model

We conducted a multiple linear regression to model and forecast the viscosity of the emulgels, using three independent variables: Carbopol® concentration, mixing time, and mixing speed. The least squares method was employed to fit the model, and the predicted viscosity values were compared with experimental results to assess the model's accuracy and prediction effectiveness.

Results and discussion

A set of nine essential Product Quality Parameters (QTPPs) was outlined, and the key attributes were identified through the Potential Failure Mode and Effects Analysis (FMEA)[25] matrix (SI-1). Consequently, two CPPs were established: mixing speed and time, and one CMA: carbopol concentration. These are recognized as extremely serious due to their effect on viscosity.

It's essential to input experimental values to develop a predictive mathematical model. Thus, we created eleven emulgel formulation trials by varying 1) carbopol concentration (% w/w), 2) mixing speed in revolutions per minute, and 3) mixing time (SI-2). The experimental viscosity for each of the eleven formulations was calculated based on the variables listed in Table 1.

Table 1: An experimental assessment of the viscosity of eleven emulgels influenced by carbopol concentration, mixing speed, and duration of mixing.

Formulation	Carbopol concentration (% w/w)	Mixing speed (rps)	Mixing time (min)	Viscosity (Pa*s)
1	0.20	241.7	0.50	1.80
2	0.40	341.7	0.75	26.1

3	0.60	191.7	1.00	47.3
4	0.80	158.3	1.25	81.7
5	1.00	133.3	1.50	89.4
6	0.20	500.0	1.50	1.60
7	0.40	241.7	1.25	26.3
8	0.60	341.7	1.00	49.7
9	0.80	158.3	0.75	85.0
10	1.00	500.0	0.50	77.6
11	1.20	133.3	2.00	95.8

The initial data analysis employed a multiple linear regression model to create a linear mathematical framework capable of predicting key system variables (SI) based on experimental values. This model exhibited coefficients of determination ranging from good to high (SI-5), indicating a robust correlation among the selected parameters. Nonetheless, some limitations in the model's predictive ability were recognized due to the system's nonlinear characteristics.

To develop a more robust predictive model for predicting key quality parameters or attributes of the material, the architecture of an artificial neural network with back propagation (ANN-INQA) was employed [24]. This approach, created by Guevara et al., aims to classify the inputs based on Pearson correlation values [26]. The analysis revealed a strong link between Carbopol concentration and viscosity (r = 0.9663), while mixing speed shows a moderate correlation (r = -0.4853), along with mixing time (r = 0.2935). See Figure 1.

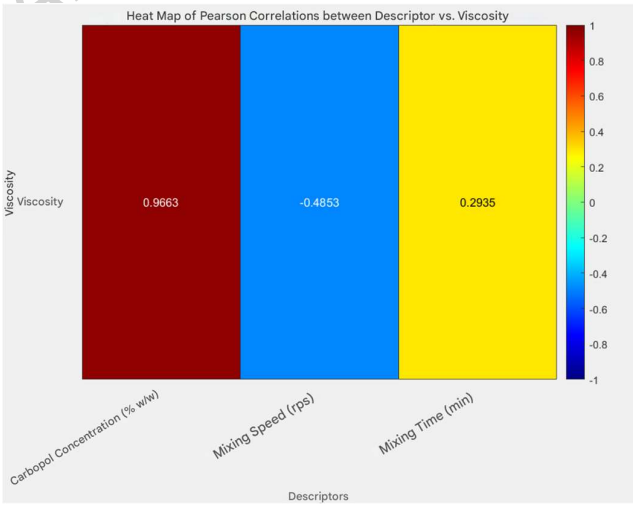

Figure 1: Correlation between key variables measured using Pearson's method

We recognized that carbopol concentration significantly influences viscosity, leading us to predict the process's carbopol concentration as our output value. For this purpose, we initially considered the mixing time and experimental viscosity of the eleven formulations listed in Table 2 as inputs. We also adjusted the number of nodes in the hidden layer, testing values from 50 to 500. We range from 50 to 500, aiming for a correlation coefficient close to one, while prioritizing optimal cross-validation. Ultimately, we developed a predictive mathematical model with 100 nodes, achieving an internal coefficient of determination (R^2) of 0.85 and a cross-validation R^2 of 0.9141 (SI), which confirmed its predictive accuracy (Figure 2)[27].

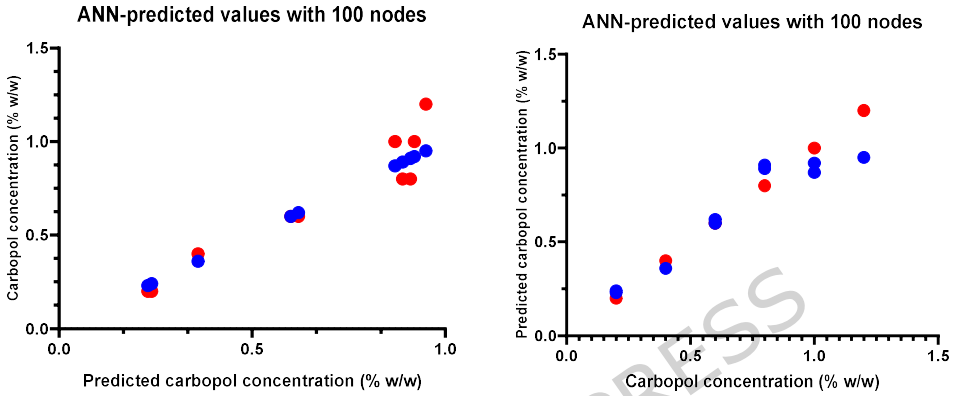

Figure 2: Cross-validation of the INQA-ANN model.

Currently, the coefficient of determination values for linear regression (SI) and ANN model is approximately equal to or greater than 0.8. While relying solely on statistical correlation would lead to the same predictions from both models, it's important to consider the pharmaceutical context. In this case, non-linear relationships in formulations may be better captured by artificial neural network (ANN) models that can predict non-linear variables. In any event, we describe the predicted values using both Linear Regression and ANN in the Supplementary Information (SI).

Table 2: Enter experimental values and ANN-INQA predictions using one hundred nodes.

Experimental values			ANN-predicted values with 100 nodes	
ANN input data (x)		*ANN output data (y)*	Predicted data by the ANN	
Mixing time (min)	Viscosity (Pa·s)	Carbopol concentration (% w/w)	Predicted carbopol concentration (% w/w)	error percentage (%)
0.50	1.80	0.20	0.24	0,17
0.75	26.10	0.40	0.36	0,10
1.00	47.30	0.60	0.60	0,00
1.25	81.70	0.80	0.89	0,11
1.50	89.40	1.00	0.92	0,08
1.50	1.60	0.20	0.23	0,15
1.25	26.30	0.40	0.36	0,10
1.00	49.70	0.60	0.62	0,03

0.75	85.00	0.80	0.91	0,14
0.50	77.60	1.00	0.87	0,13
2.00	95.80	1.20	0.95	0,21
Validation of the ANN with 100 nodes			R^2= 0.85	

Commercial Voltaren emulgel and Artritis Voltaren Emulgel and Artritis gel were utilized to assess the predictive ability of the INQA-ANN-100 model experimentally. The viscosities of both products were measured, and a mixing time was established based on the viscosity values presented in Table 2. This data estimated the carbopol concentration in percent weight-to-weight (w/w) using the INQA-ANN-100 model (predictions shown in Table 3). After mixing the components with the predicted concentrations, a viscosity of 85.005 was achieved for the diclofenac arthritis formulation, demonstrating a 99.9% concordance. The same procedure was applied to the Voltaren emulgel, resulting in a viscosity of 48.852 and a concordance percentage exceeding 94%, as detailed in Table 3.

Table 3: Validation of the INQA-ANN-100 Predictive Model via Experiments.

ANN input data (x)		ANN-predicted data	Experimental data		Reference and experimental viscosities (Pa·s)
Mixing time (min)	Reference viscosity (Pa·s)	Predicted carbopol concentration (% w/w)	Experimental carbopol concentration (% w/w)	Average experimental viscosity (Pa·s)	Error percentage (%)
0.75	46.149	0.608	0.609	48.852	5.856
1.5	85.019	0.940	0.947	85.055	0.042

The model was developed using two inputs, successfully passing statistical parameters and experimental validation with commercial products. We encountered errors ranging from 0.5% to 5%, prompting us to incorporate three additional inputs. This approach integrated the three essential variables—mixing time, mixing speed, and viscosity—to assess potential complex interactions, expanding the range of reference viscosities. Like the previous model, we varied the number of hidden layer nodes, increasing it from 50 to 500 systematically searching for the best correlations (SI-31). Ultimately, we identified a superior model utilizing 400 nodes, which achieved an internal validation R^2 of 0.908 and a cross-validation result that exceeded the statistical benchmarks (Figure 3) and (SI).

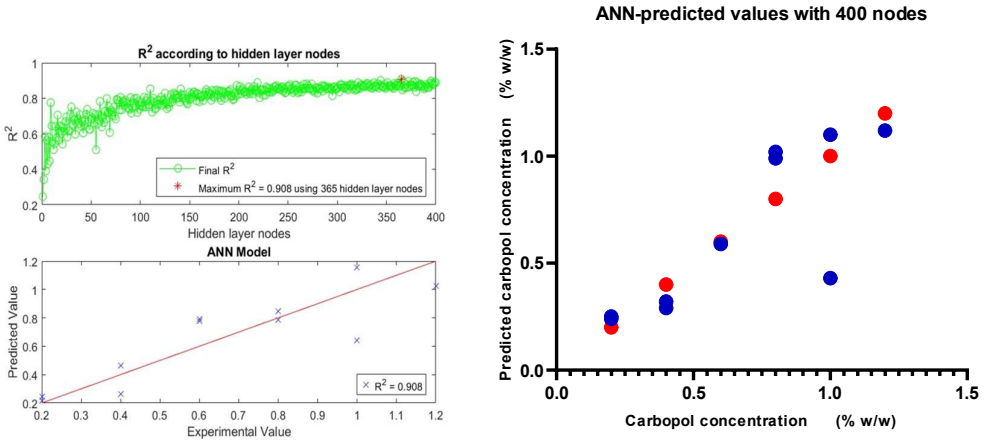

Figure 3: a) according to hidden layer nodes b) Cross-validation of the ANN-INQA-400 model

As the two-input model, we tested its predictive capability with four commercial compounds (Table 4). The results indicated agreement percentages ranging from 96% to 98% (Figure 4), further validating the model's predictive power. However, in the case of low viscosities, as in Formulation A (Table 4), the model's predictive capacity is limited. The short mixing times and high mixing speed needed for low-viscosity fluids may not produce reliable and consistent predictions, as indicated by the data in Table SI-31 for the training set. The reasons for this inconsistency may vary; however, we are focused on developing independent models for low viscosities by identifying new critical parameters.

Table 4: Validation of the INQA-ANN-400 Predictive Model through Experiments.

Formulation	ANN input data			Data predicted by ANN (400 nodes)	Experimental data		Standard deviation of experimental viscosity (Pa.s)	Reference data	Viscosities (Pa.s) Reference and Experimental
	Mixing time (min)	Viscosity (Pa.s)	Mixing speed (rps)	Predicted carbopol concentration (% w/w)	Experimental carbopol concentration (% w/w)	Average experimental viscosity (Pa.s)		Reference viscosity (Pa.s)	Error percentage (%)
A Voltarem	0.75	46.149	341.7	0.360	0.364	13.132	0.558	46.149	71.544
B Aloe vera	1.00	60.538	191.7	0.684	0.684	61.887	6.712	60.538	2.228
C Arnigel	1.25	74.451	158.3	0.929	0.931	77.224	7.014	74.451	3.725
D Arthritis	1.50	85.019	133.3	1.084	1.086	88.507	6.63	85.019	4.103

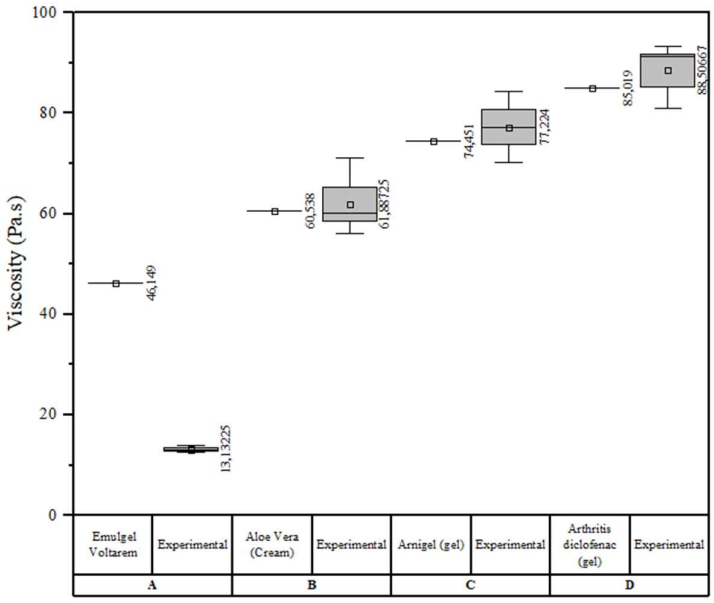

Figure 4: Validation of INQA-ANN-400

This advancement significantly reduces trial and error in the formulation of gels and emulgels [28], representing a notable application of artificial neural network models in the formulation of emulsifiers. Integrating the QbD approach with artificial intelligence tools, particularly ANNs, marks a significant methodological advancement in rational drug development. This study demonstrates that the combined use of QbD and ANNs facilitates modeling the relationship between critical process parameters and quality attributes, leading to a substantial reduction in experimental testing through highly accurate predictions. This strategy optimizes resource utilization, shortens development timelines, and enhances understanding of intricate physicochemical phenomena, such as polymer neutralization-induced gelation and the interactions of dynamic viscosity with time and rate. By designing a robust operating space, this approach demonstrates scientific traceability that surpasses traditional empirical methods, thereby supporting technical decisions based on objective and reproducible data (ICH Q8(R2), Q9, Q10).

This comprehensive approach establishes a foundation for the digital transformation of pharmaceutical development, aligning with the global trend toward smart manufacturing, as highlighted in recent publications. The progressive application of ANN models, utilizing an increasing number of nodes, has demonstrated an enhanced capability to identify nonlinear patterns and multifactorial interactions that are often difficult to capture with traditional linear models. This methodological synergy contributes to more reliable predictions of critical attributes, such as viscosity, even within complex commercial formulations, reinforcing its industrial relevance. Additionally, by integrating cross-validation and testing against reference products, the model's robustness and transferability are effectively ensured. This comprehensive approach lays the groundwork for the digital transformation of pharmaceutical development, in alignment with the global trend toward smart manufacturing, as highlighted in recent publications [29-32].

Advantages, challenges and limitations

The integration of QbD and ANNs represents a significant advancement in the field of pharmaceutical formulation development, offering a robust, data-driven framework for improving process

understanding and optimization. This collaborative approach markedly enhances the predictive accuracy of critical quality attributes, such as viscosity, while also reducing the experimental workload. Consequently, it accelerates formulation screening and the exploration of the design space. However, several challenges remain, particularly concerning quantity and quality data, model interpretability, the sensitivity of the employed instrumental methods, and the complexities of regulatory acceptance. The effectiveness of ANNs is intrinsically linked to the availability of extensive, high-quality datasets. Poor data curation can lead to issues such as overfitting or skewed predictive outcomes [33]. Furthermore, ANNs models are often perceived as "black boxes," making it difficult to clarify predictions on a mechanistic level or justify them in regulatory settings. Finally, integrating these methodologies into regulated environments requires strict adherence to principles of data integrity, traceability, and ongoing risk management [34].

Future Perspectives

Future developments are expected to concentrate on establishing hybrid QbD and Artificial Neural ANNs frameworks that combine data-driven and mechanistic models. This includes integration with physiologically based pharmacokinetic (PBPK) models or population-based models to improve interpretability and transferability across different scales. Progress in digital infrastructure, adherence to the principles of findable, accessible, interoperable, and reusable data, and the use of digital twins for process monitoring will facilitate real-time learning and control within QbD paradigms [35]. From a regulatory standpoint, organizations like FDA and the EMA are actively formulating guidelines to ensure the responsible application of AI and machine learning in pharmaceutical development, focusing on transparency, model lifecycle management, and predefined change control strategies [33]. As these standards evolve, multidisciplinary collaboration among formulation scientists, data engineers, and regulatory experts will be essential to ensure that QbD-ANN methodologies are robust, explainable, and recognized as valuable tools for innovation in the pharmaceutical sector.

Conclusions

This study highlights the feasibility and precision of using ANNs as predictive tools for carbopol emulsion formulations. Model optimization notably enhanced its predictive ability by strategically selecting critical variables, especially in estimating carbopol concentration through viscosity, time, and mixing speed. Experiments demonstrated validation and comparison with a commercial formula, establishing a strong correlation between predicted and actual values, reinforcing the approach's effectiveness.

Additionally, incorporating the three key variables into the model facilitated the evaluation of complex interactions affecting viscosity, thereby broadening its applicability to various formulations. The findings suggest that the model is exceptionally reliable for systems with viscosities exceeding 50,000 mPa ·s·mPa·s, highlighting its capability to enhance formulation processes in the pharmaceutical sector. Consequently, this study lays the groundwork for future research to refine predictive models in gelling systems engineering and optimize manufacturing processes using artificial intelligence tools.

Acknowledgments

We want to thank the INQA group and the pharmaceutical chemistry program for their assistance at VRI Universidad El Bosque in Bogotá, Colombia.

Author contributions

Conceptualization, R.J, J.G-P; methodology, L.D, L.M, R.J, J.G-P software LD, LM, J.G-P validation, L.D, L.M, R.J, J.G-P, formal analysis, L.D, L.M, R.J, J.G-P investigation, L.D, L.M, R.J, J.G-P

resources, R.J, J.G-P data curation, L.D, L.M, R.J, J.G-P writing—original draft, L.D, L.M, R.J, J.G-P preparation, L.D, L.M, R.J, J.G-P.; writing—review and editing, R.J, J.G-P., visualization, R.J, J.G-P.; supervision, R.J J.G-P.; project administration, R.J, J.G-P.; funding acquisition, R.J, J.G-P., All authors have read and agreed to the published version of the manuscript.

Competing interests
The authors declare no competing interests.

Data availability
Data is provided within the manuscript or supplementary information files

Funding
INQA Research Group, Universidad El Bosque

References

1. **Yu, L. X.**, Amidon, G., Khan, M. A., Hoag, S. W., Polli, J., Raju, G. K. & Woodcock, J. Understanding pharmaceutical quality by design. *The AAPS Journal* **16**, 771–783 (2014).
2. **Gholap, A. D.**, Choudhary, S., More, M. P. & Patil, P. O. Advances in artificial intelligence in drug delivery and development: A comprehensive review. *Computers in Biology and Medicine* **165**, 108702 (2024).
3. **Srinivas Murthy, B. R.**, Gupta, V., Sharma, A. & Patel, R. Advances in artificial intelligence for drug delivery and development. *Frontiers in Health Informatics* **13**, 1–15 (2024).
4. **Serrano, D. R.**, García-García, A., O'Connor, A. & Lalatsa, A. Artificial intelligence (AI) applications in drug discovery and delivery: Revolutionizing personalized medicine. *Pharmaceutics* **16**, 1328 (2024).
5. **Tran, T. T. V.**, Tayara, H. & Chong, K. T. Recent studies of artificial intelligence on in silico drug distribution prediction. *International Journal of Molecular Sciences* **24**, 1815 (2023).
6. **Negut, I.** & Bita, B. Exploring the potential of artificial intelligence for hydrogel development—A short review. *Gels* **9**, 845 (2023).
7. **Taskinen, J.** & Yliruusi, J. Prediction of physicochemical properties based on neural network modelling. *Advanced Drug Delivery Reviews* **55**, 1163–1183 (2003).
8. **Lu, M.**, Wang, X., Li, Q., Zhang, Y. & Chen, H. Artificial intelligence in pharmaceutical sciences. *Engineering* **27**, 37–69 (2023).
9. **Paruthi, S.**, Raghav, Y. Y., Kumari, P., Kumar Tipu, R., & Sharma, N. (2025). AI based predictive modelling for compressive strength of metakaolin-based geopolymer concrete incorporated with Nano Titanium. *Journal of Structural Integrity and Maintenance,* 10(3).
10. **Ali, K. A.**, El-Sherbiny, I. M., Salem, H. M. & El-Gizawy, S. A. Influence of artificial intelligence in modern pharmaceutical formulation and drug development. *Future Journal of Pharmaceutical Sciences* **10**, 53 (2024).
11. **Momeni, M., Afkanpour**, M., Rakhshani, S., Mehrabian, A., & Tabesh, H. (2024). A prediction model based on artificial intelligence techniques for disintegration time and hardness of fast disintegrating tablets in pre-formulation tests. *BMC Medical Informatics and Decision Making*, 24(1), 88.
12. **Obaidullah, A. J**., & Mahdi, W. A. (2025). Machine learning-based analysis on pharmaceutical compounds interaction with polymer to estimate drug solubility in formulations. *Scientific Reports*, 15(1), 23683.

13. **Hajibolouri, E.,** Najafi-Silab, R., Daryasafar, A. et al. Using data-driven models to simulate the performance of surfactants in reducing heavy oil viscosity. *Scientific Reports,* 14, 27670 (2024).
14. **Álvarez-Gómez, C.**, Fonseca-Benítez, A. V. & Guevara-Pulido, J. Design, synthesis, and in vitro evaluation of a carbamazepine derivative with antitumor potential in a model of Acute Lymphoblastic Leukemia. *PLoS ONE* **20**, e0319415 (2025).
15. **Huertas, T**., Millán, D., Fonseca-Benítez, A.V. et al. Structure-based design of an antibacterial peptide from the Myotoxin II sequence, evaluating its effectiveness against Gram-negative bacteria and its safety. *Scientific Reports,* 15, 24350 (2025).
16. **Tummala, S. R.** & Gorrepati, N. AI-driven predictive analytics for drug stability studies. *Journal of Pharmaceutical Insights and Research* **2**, 188–198 (2024).
17. **Johnson, K. B.**, Wei, W., Weeraratne, D., Frisse, M. E., Misulis, K., Rhee, K., Zhao, J. & Snowdon, J. L. Precision medicine, AI, and the future of personalized health care. *Clinical and Translational Science* **14**, 86–93 (2021).
18. **Vanpariya, F.**, Shiroya, M. & Malaviya, M. Emulgel: A review. *International Journal of Scientific Research* **10**, 847 (2021).
19. **Huan, S.**, Zhang, L., Huang, Y., Wang, H., Wu, Z. & Ling, J. Two-phase emulgels for direct ink writing of skin-bearing architectures. *Advanced Functional Materials* **29**, 1902990 (2019).
20. **Begum, S. G.**, Reddy, P. N., Rajesh, K. & Kumar, P. A. A review on emulgels—A novel approach for topical drug delivery. *Asian Journal of Pharmaceutical Research and Development* **7**, 70–77 (2019).
21. **Richa, A; Roy C.** (2023). Self-assembled pH-stable gellan/κ-carrageenan bigel: Rheological studies and viscosity prediction by neural network. *International Journal of Biological Macromolecules*.
22. **Mohammad, S**., Akand, R., Cook, K. M., Nilufar, S., & Chowdhury, F. (2024). Leveraging Deep Learning and Generative AI for Predicting Rheological Properties and Material Compositions of 3D Printed Polyacrylamide Hydrogels. *Gels*, 10(10)
23. **Yahoum, M. M.,** Toumi, S., Hentabli, S., Tahraoui, H., Lefnaoui, S., Hadjsadok, A., Amrane, A., Kebir, M., Moula, N., Assadi, A. A., Zhang, J., & Mouni, L. (2023). Experimental Analysis and Neural Network Modeling of the Rheological Behavior of Xanthan Gum and Its Derivatives. *Materials*, 16(7), 2565.
24. **Guevara-Pulido,** J., Jimenez, R. A., Morantes, S. J., Jaramillo, D. N., & Acosta-Guzmán, P. (2022). Design, synthesis, and development of 4-[(7-Chloroquinoline-4-yl) amino] phenol as a potential SARS-CoV-2 mpro inhibitor. *ChemistrySelect*, 7(15), e202200125.
25. **Rana, S.** & Belokar, R. M. Quality improvement using FMEA: A short review. *International Research Journal of Engineering and Technology* **4**, 263–267 (2017).
26. **Schober, P.**, Boer, C. & Schwarte, L. A. Correlation coefficients: Appropriate use and interpretation. *Anesthesia & Analgesia* **126**, 1763–1768 (2018).
27. **Alexander, D. L.**, Tropsha, A. & Winkler, D. A. Beware of R2: Simple, unambiguous assessment of the prediction accuracy of QSAR and QSPR models. *Journal of Chemical Information and Modeling* **55**, 1316–1322 (2015).
28. **Deng, B.**, Lin, Y., Wang, C., Li, F., Liu, Z. & Zhang, X. Predicting rheological properties of HAMA/GelMA hybrid hydrogels via machine learning. *Journal of the Mechanical Behavior of Biomedical Materials* **168**, 107005 (2025).
29. **Li, W.**, Zhang, R., Chen, Y., Zhao, S. & Liu, H. Developing a machine learning model for accurate nucleoside hydrogels prediction based on descriptors. *Nature Communications* **15**, 2603 (2024).

30. **Mohammad, S.**, Alavi, M., Martinez, P. A. & Karim, A. Leveraging deep learning and generative AI for predicting rheological properties and material compositions of 3D printed polyacrylamide hydrogels. *Gels* **10**, 660 (2024).
31. **Aru, P. B.**, Singh, R., Kumar, V. & Sharma, N. Quality by Design (QbD) in Pharmaceutical Development: A Comprehensive Review. *GSC Biological and Pharmaceutical Sciences* **26**, 328–340 (2024).
32. **Suriyaamporn, P.**, Opanasopit, P., Ngawhirunpat, T. & Rojanarata, T. The Artificial Intelligence-Powered New Era in Pharmaceutical Research and Development: A Review. *AAPS PharmSciTech* **25**, 188 (2024).
33. Kapoor, D. U., Sharma, J. B., Gandhi, S. M., Prajapati, B. G., Thanawuth, K., Limmatvapirat, S., & Sriamornsak, P. (2024). AI-driven design and optimization of nanoparticle-based drug delivery systems. Science, *Engineering and Health Studies*.
34. U.S. Food and Drug Administration. (2025). Artificial Intelligence and Machine Learning (AI/ML) in Drug Development: Discussion Paper and Request for Feedback. Silver Spring, MD. Retrieved from https://www.fda.gov/science-research/science-and-research-special-topics/artificial-intelligence-and-machine-learning-aiml-drug-development.
35. Selvaraj, C., Chandra, I., & Singh, S. K. (2022). Artificial intelligence and machine learning approaches for drug design: challenges and opportunities for the pharmaceutical industries. *Molecular diversity,* 26(3), 1893-1913.