Journal of Controlled Release 376 (2024) 1025–1038

Contents lists available at ScienceDirect

Journal of Controlled Release

journal homepage: www.elsevier.com/locate/jconrel

Leveraging machine learning to streamline the development of liposomal drug delivery systems

Remo Eugster [a], Markus Orsi [a], Giorgio Buttitta [b], Nicola Serafini [c], Mattia Tiboni [c], Luca Casettari [c], Jean-Louis Reymond [a], Simone Aleandri [a], Paola Luciani [a,*]

[a] Department of Chemistry, Biochemistry and Pharmaceutical Sciences, University of Bern, Bern, Switzerland
[b] Department of Chemistry and Technologies of Drugs, Sapienza University of Rome, Rome, Lazio, Italy
[c] Department of Biomolecular Sciences, University of Urbino Carlo Bo, Urbino, PU, Italy

ARTICLE INFO

Keywords:
Artificial intelligence
Machine learning
Drug delivery & development
Liposomes
Microfluidics

ABSTRACT

Drug delivery systems efficiently and safely administer therapeutic agents to specific body sites. Liposomes, spherical vesicles made of phospholipid bilayers, have become a powerful tool in this field, especially with the rise of microfluidic manufacturing during the COVID-19 pandemic. Despite its efficiency, microfluidic liposomal production poses challenges, often requiring laborious, optimization on a case-by-case basis. This is due to a lack of comprehensive understanding and robust methodologies, compounded by limited data on microfluidic production with varying lipids. Artificial intelligence offers promise in predicting lipid behaviour during microfluidic production, with the still unexploited potential of streamlining development. Herein we employ machine learning to predict critical quality attributes and process parameters for microfluidic-based liposome production. Validated models predict liposome formation, size, and production parameters, significantly advancing our understanding of lipid behaviour. Extensive model analysis enhanced interpretability and investigated underlying mechanisms, supporting the transition to microfluidic production. Unlocking the potential of machine learning in drug development can accelerate pharmaceutical innovation, making drug delivery systems more adaptable and accessible.

1. Introduction

Liposomes, have revolutionized the field of drug delivery (Fig. 1a). [1] Due to their versatile applications and the pandemic-driven push to standardize lipid-based delivery methods, these vesicles have garnered increasing attention. [2] However, while their potential is vast, challenges such as scalability, and cost-effectiveness still need to be addressed to fully realize their widespread application in clinical settings. [3] Nevertheless, over the last three decades, more than 14 liposome-based drug products have been approved, with applications ranging from cancer treatments to vaccines. [2] Liposomal carriers can encapsulate both hydrophilic and hydrophobic drugs, offering a protective environment and enhancing their solubility, stability, and bioavailability. [1,4] Further, the liposomal surface can be engineered to circulate longer in the bloodstream, improving the pharmacokinetics and biodistribution. [5,6]

Systemic drug delivery is profoundly impacted by liposome size, affecting physiological processes such as hepatic uptake, tissue diffusion, extravasation, and renal clearance. [7–9] The size range of 50-200 nm is considered optimal for drug nanocarriers in systemic parenteral administration, balancing tissue and capillary pore size limitations. [7,10,11] Additionally, smaller liposomes (<150 nm) demonstrate enhanced lymphatic uptake and transport, crucial for effective drug delivery. [10,12] Further, larger liposomes might escape clearance and act as long-acting depot systems. [13,14] Liposomes mainly consist of phosphatidylcholine (PC) lipids, essential for drug delivery due to their biocompatibility and ability to form stable bilayers. [15]

Traditional methods of liposome production, such as thin-film hydration and ethanol injection, present challenges in scalability, reproducibility, and control over liposome size and polydispersity (Fig. 1b). [16,17] In contrast, microfluidic technology offers a promising alternative by enabling precise control over the production process. With the increasing accessibility of microfluidic devices for lipid nanoparticle (LNP) production, this technology is becoming a reality for small and medium enterprises (SMEs), encouraging formulation scientists to explore more liposome-based drug delivery systems. [16–20] This

* Corresponding author.
E-mail address: paola.luciani@unibe.ch (P. Luciani).

https://doi.org/10.1016/j.jconrel.2024.10.065
Received 12 July 2024; Received in revised form 3 October 2024; Accepted 29 October 2024
Available online 8 November 2024

technology involves manipulating fluids within microchannels, allowing for controlled mixing and assembly of liposomal components (Fig. 1b). Microfluidic production enhances reproducibility and scalability, generating liposomes with uniform size distributions and tailored properties, fundamental for consistent therapeutic performance. [7,16,18,21] Additionally, microfluidic production is more sustainable because it does not rely on the use of chlorinated solvents like chloroform compared to conventional methods such as the thin film hydration, resulting in less toxic waste and ultimately contributing to a reduced environmental footprint. [16] Further, this production offers fine control over critical process parameters (CPPs) such as flow rates, pressure, and temperature, which are crucial for optimizing liposome size, lamellarity, encapsulation efficiency, and stability. [22–24] Despite these benefits, however, several challenges hinder clinical and industrial applications. The primary obstacle is translating bench-scale liposomal production to larger scales, as traditional methods often do not scale up seamlessly. [24] That said, transitioning to more scalable microfluidic production is not trivial as different liposomal formulations respond uniquely to microfluidic conditions, requiring extensive optimization on a case-by-case basis. [24] Variations in microfluidic chip design further complicate standardization. Advances in 3D printing could help by enabling robust, standardized chip fabrication (Fig. 1c). [25,26] While extrusion can control liposome size, achieving this precision in microfluidics is more complex due to the interplay of various factors, often requiring extensive experimentation to adjust the flow rate ratio (FRR; aqueous: organic flow rate) at a sufficient total flow rate. [21,24,27]

Mathematical models like Design of Experiments (DoE) have guided the development of microfluidic liposome production processes, elucidating the relationships between process parameters and the resulting liposome characteristics. [23,25,28] However, their applicability is often limited to low-dimensional design spaces, and their scalability and generalization are restricted. In recent years, the application of machine learning (ML) and artificial intelligence (AI) in pharmaceutical sciences has gained traction due to their ability to model multi-dimensional problems. [22,23,29–32] ML is valuable throughout the entire drug discovery and development pipeline, [22,33] from exploring chemical space to identify new hit compounds [33–37], to predicting synthetic pathways, [38–40] Absorption-Distribution-Metabolism-Excretion Toxicity (ADMET) properties [41,42] and target activity. [43] In complex processes influenced by multiple factors, conventional mathematical models may struggle to capture underlying patterns, whereas AI can effectively reveal and support these mechanisms. [44] However, many methods rely on external datasets often lacking quality. [45,46] Literature [47–51] and previous work using in-house data have demonstrated that ML can accurately learn and predict experimental outcomes with clean data, even with smaller datasets. [52] However, ML itself can generally not overcome the data sparsity problems which are primary limitations for most applications so far as comprehensive and high-quality data are often lacking. However, this does not apply to microfluidic liposome production, as a fair amount of in-house data can readily be produced. Therefore, ML holds significant promise for predicting lipid behaviour during microfluidic production with high accuracy. [23,31,32,53,54]

Despite their ability to model complex multi-dimensional data, though, ML models often lack explainability, making them less suitable for gaining deeper insights. [55] Explainable AI (XAI) techniques, allow researchers to extract information from black-box models and provide more transparency to the predictions. [55] Combined, ML and XAI can potentially reduce reliance on empirical methods and provide increased transparency, accelerating the optimization of the liposomal production process. [31,32,53,54,56]

Herein we report on training ML models, with an in-house generated data set to accurately predict critical quality attributes (CQAs), including the formation and size of liposomes, as well as process parameters for their microfluidic production. We applied the principles of XAI and gained deeper insights into the complex interactions between formulation- and process-parameters, improving the prediction of lipid behaviour during microfluidic production. By training these models, we aim to establish a reliable framework that enables transitioning from traditional liposome production methods to microfluidic techniques (Fig. 1d). This capability can streamline the development of liposomal drug delivery systems, reducing reliance on empirical methods and

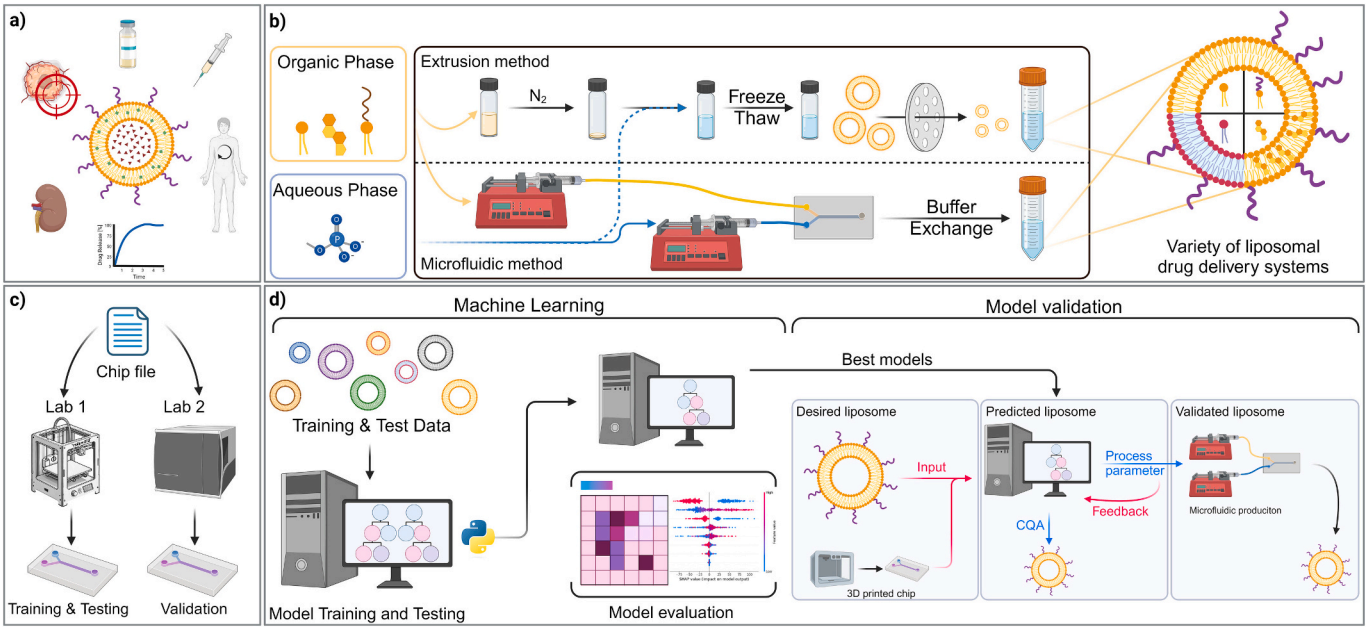

Fig. 1. Schematic demonstration of traditional and data-driven formulation development for liposomal drug delivery systems. a) Liposomes serve as effective drug delivery systems, enhancing drug stability, enabling parenteral administration for systemic effects, and offering controlled drug release. PEGylation and precise size adjustments help reducing clearance and increase the retention of drug carriers in the body. Additionally, liposome size influences physiological processes such as hepatic uptake and tissue diffusion. b) Traditional extrusion-based versus microfluidic production method of liposomes. c) 3D printing of microfluidic chips with specific channel geometries. d) Machine learning (ML) model development, testing model selection based on model interpretation and validation in the wet lab.

accelerating the optimization process. [31,32,53,54] Our ultimate goal is to develop a robust, scalable, and standardized process widely adoptable in the pharmaceutical sciences, to accelerate the delivery of therapeutic agents to clinics and patients.

2. Results and discussion

2.1. Liposomal formulation screening

A comprehensive range of lipid-based formulations has been screened following multiple full factorial DoEs and using a microfluidic process, varying CPPs to optimize liposome characteristics. The primary CQA obtained in response to these variations was liposome size, a key determinant of their therapeutic efficacy and stability. [7] The full factorial design allowed us to screen a vast portion of the experimental space, distributing the experimental points among the investigated CPPS' s range. The lipid excipients selected for this study include dimyristoylphosphatidylcholine (DMPC), dioleoylphosphatidylcholine (DOPC), hydrogenated soy phosphatidylcholine (HSPC), dipalmitoylphosphatidylcholine (DPPC), and palmitoyloleoylphosphatidylcholine (POPC). This selection covers a broad spectrum of lipids commonly used in the development of liposomal drug delivery systems, providing a diverse basis for the analysis. [15] The chosen lipids are pivotal for subcutaneous (s.c.) and parenteral drug delivery, given their distinct biophysical properties and interactions within liposomal structures. [15] Cholesterol is included as a crucial component due to its role in enhancing membrane stability and reducing permeability, thus improving drug retention. [15] Additionally, PEGylated lipids are incorporated to prolong circulation time by providing a steric barrier against opsonization and subsequent clearance by the mononuclear phagocyte system. [15] Further it assures stability of the liposomes by avoiding the coalescence and fusion of the vesicles. [57]

Liposomes were manufactured following a microfluidic technology as outlined in Fig. 1b. Chips with different channel geometry were 3D printed and utilized in the production offering flexibility and independence at low cost (Fig. S1). [25] Immediately following liposome production, the sample underwent analysis via dynamic light scattering (DLS) and the hydrodynamic diameter and the polydispersity index (PDI) of the liposomes were chosen as the primary response for optimization in this study. [7]

The data set obtained from the lipid screening was analyzed using the Tree map (TMAP) dimensionality reduction method (Fig. 2a-c) and an alluvial diagram (Fig. 2d). There are currently limited algorithms to visualize multi-dimensional data sets containing chemical structures and associated properties while preserving both global and local features with a sufficient level of detail to allow for human inspection and interpretation. [58] TMAPs offer a solution to this problem with a new data visualization method, capable of representing data sets of arbitrary multidimensionality as a two-dimensional tree. [58] Hence, TMAPs are a tool to inspect multi-dimensional datasets in 2-dimensional space while preserving nearest neighbour relationships. [58] Visualizations based on TMAP are better suited than t-SNE or PCA for the exploration

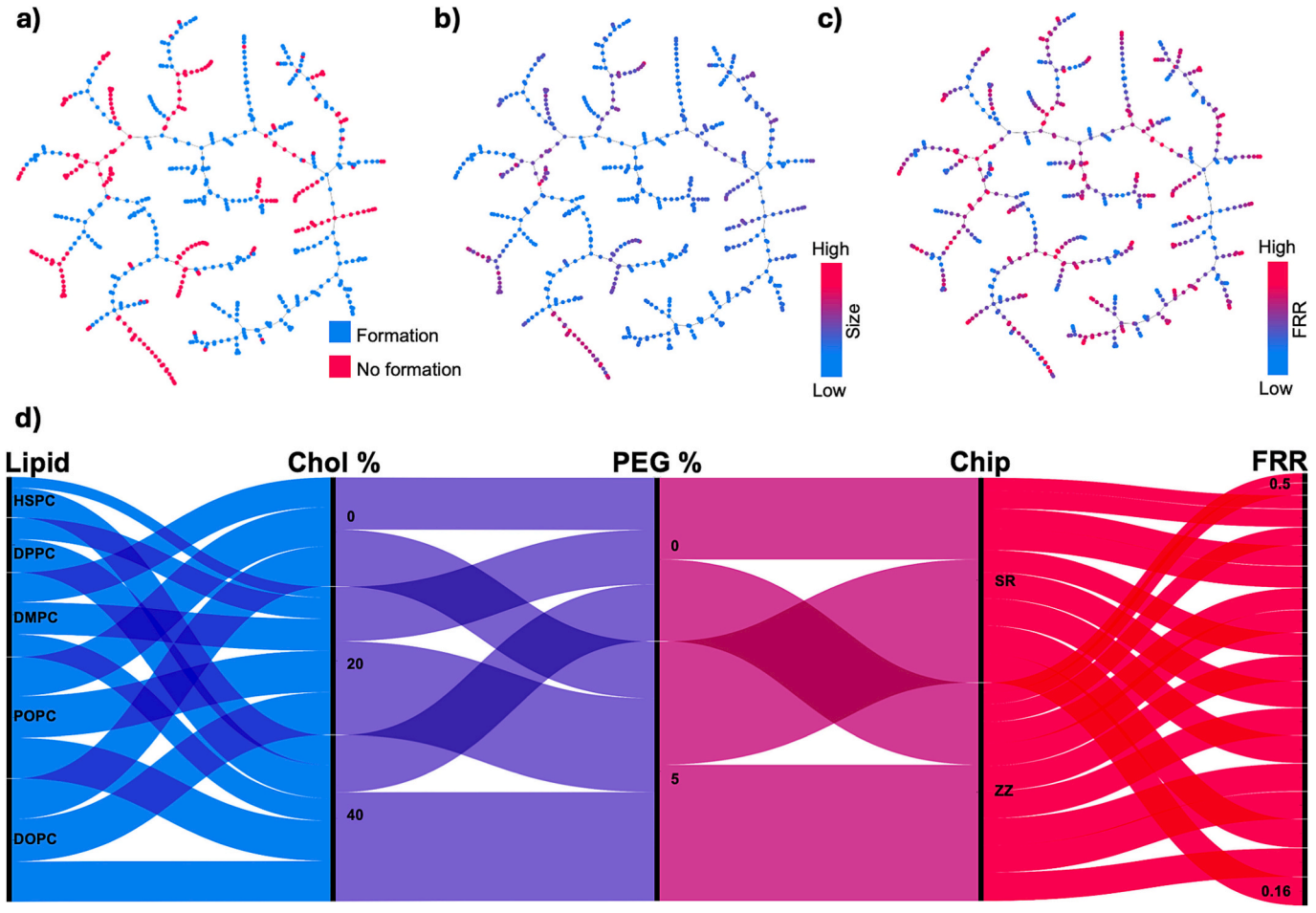

Fig. 2. Visualization of liposome data set. a-c) TMAP visualization of the data set for three responses: formation of liposomes (a), size (b), and (c) flow rate ratio (FRR). The features include formulation parameters such as PEGylated lipids (PEG %), cholesterol (Chol %), and simplified lipid descriptors, and chip geometry (Chip) as well as flow rate ratio (FRR). d) Alluvial plot depicting trends and patterns in the formation of liposomes based on formulation and production parameters. The width of the flows corresponds to the proportion of liposomes in the data set, illustrating the distribution and significance of different parameters.

and interpretation of large data sets due to their tree-like nature, increased local and global neighborhood and structure preservation, and the transparency of the methods the algorithm is based on, as reported elsewhere. [58] TMAPs avoid overlaps of data and use the entire space which makes visualization and interpretation easier. Lastly the maps are interactable allowing close interpretation of each data point (Supporting Information). A direct comparison with PCA and tSNE can be seen in Fig. S2. Each point in the TMAP plot corresponds to an instance in the dataset where its location is determined by the underlying formulation parameters such as PEGylated lipids (PEG), cholesterol (Chol), simplified lipid descriptors, and chip geometry (Chip). Further, the colour of each point represents the response value (e.g. Size) of the according data instance. Visualizing this multi-dimensional data using TMAP analysis reveals clear clusters corresponding to liposome formation, size, and FRR, indicating that certain combinations of CPPs and lipid formulations consistently produce optimal characteristics. A closer inspection of the clusters reveals that saturated lipids are less likely to form liposomes and if they do tend to form larger particles. This may be due to the fact that lipids with saturated fatty acid chains tend to form rigid bilayers with high phase transition temperatures (T_m), making them less likely to form stable liposomes. [59] Additionally, their rigidity presents challenges in achieving the narrow curvature required, which could explain the tendency to form larger particles. [60] Efficient clustering indicates that the combination and interplay of underlying descriptors can effectively differentiate the data and group it accordingly. This capability serves as a strong foundation for subsequent modelling efforts. However, it must be understood that this alone does not justify a machine-learning approach.

The alluvial plot effectively illustrates the influence of specific CPPs and lipid formulations on liposome formation (Fig. 2d and Fig. S3). By depicting data flows that represent the number of formed liposomes across different categories, it visually demonstrates the interactions between various variables. For instance, cholesterol content noticeably affects liposome formation, as evidenced by the varying diameters of flows between lipid types and cholesterol percentages. Particularly, liposomes based on HSPC and DPPC appear to be significantly impacted by the absence of cholesterol. This can be attributed to the structural characteristics of the lipid chains. HSPC and DPPC, which both contain two saturated fatty acid chains, typically form rigid bilayers with high T_m, making them less prone to forming stable liposomes, supporting the previous observations in TMAPs. [59] The addition of cholesterol disrupts this tight packing and introduces disorder into the bilayer, facilitating the curvature required to form closed vesicles, especially in a microfluidic environment. [60]

In contrast, liposomes based on lipids with unsaturated fatty acid chains, such as DOPC and POPC, form consistently across all cholesterol ranges. These lipids have lower T_m and inherently higher fluidity, allowing for liposome formation without the need for added cholesterol. A similar trend is observed with the addition of PEG % (Fig. S2). The presence of PEG in liposomes prevents surface interactions and stabilizes the formed liposomes. [61,62] Interestingly, the presence of both cholesterol and PEG % appears to slightly influence each other symbiotically, with their combination being the most favorable for liposome formation. This can be explained by the combinational effect of steric repulsion and introduction of additional asymmetry and space into the bilayer, which facilitates liposome formation. [60]

Lastly, the relationship between Chip and FRR (Flow Rate Ratio) shows that the data flows between different chips are of similar size, indicating similar behaviour across chips. However, liposome formation is more prevalent at lower FRR, which favors the aqueous phase. This is likely due to the high presence of ethanol at higher FRR, making it thermodynamically less favorable for lipids to form vesicles. [63]

However, the plot illustrates the dataset's complexity, making direct extrapolation impractical and suggesting the use of an ML approach, which can handle large amounts of multi-dimensional data and identify patterns not easily seen by traditional methods (Table S1).

2.2. Feature selection

A first task in ML is feature selection, which involves identifying and using the most relevant variables (features) in a dataset, associated with microfluidic liposome production. This process reduces overfitting, avoids correlated features that confound the model and improves generalization. The alluvial plot in Fig. 2d supports initial feature selection for further analysis. The selected features include formulation parameters, such as lipid molecular descriptors, cholesterol content, and PEG content, as well as production parameters like chip geometry and FRR. These variables have relevant influences on the process and uniquely contribute to the formation and characteristics of liposomes. [23,25,64,65] Lipid molecular descriptors provide insights into the physicochemical properties of the lipids, influencing their behaviour in the microfluidic process. While formulation parameters have a direct impact on circulation and drug retention, chip geometry affects the mixing and assembly of lipids, influencing liposome size and uniformity. [25] The FRR determines the relative velocities of the aqueous and lipid phases, impacting liposome size and encapsulation efficiency. [23,25] Previous experiments have shown that total flow rate (TFR) becomes less important above a certain high turbulent flow, [28] where its influence on liposome formation is minimal which is supported by simulations elsewhere. [28,64] Therefore, the TFR was excluded from the initial feature vector after ensuring to be above the threshold of the used chips (>8 mL/min). A well-constructed feature vector (Fig. 3a) avoids confounding and overfitting, while improving generalization and thereby promises to model the complex underlying data, providing predictive insights and optimizing the microfluidic production process.

The Spearman correlation matrix provides insights into the correlation between different features in the dataset, independent of the predictive modelling process (Fig. 3b). Each cell in the matrix represents the Spearman correlation coefficient, indicating the strength and direction of the relationship between two features. Positive correlation coefficients indicate a positive linear relationship between features, while negative coefficients signify a negative linear relationship. For example, enhanced correlation coefficients between pairs of features, such as between the lipid descriptors, may indicate multicollinearity, where one feature can be linearly predicted from another which potentially leads to overfitting. However, in our study, we included these correlated lipid features because, despite their correlation, as literature suggests that multiple lipid characteristics, as captured in our features, are crucial for capturing the underlying microfluidic process. [66] Additionally, tree-based models are considered favorable because they are less sensitive to multicollinearity, as they do not assume that predictor variables are independent. [67] Tree-based ensemble methods automatically identify and prioritize the most important features in a dataset during the model training process. [68] Unlike some other machine learning algorithms tree-based models naturally evaluate each feature's importance by how effectively it splits the data to reduce impurity. Features that provide the most significant reduction in impurity are used at the higher levels of the tree, while less important features may not be used at all or only in the deeper parts of the tree. This further reduces the risk of including multicorrelated features as this "autonomous feature selection" itself helps in selecting relevant features requiring less manual preprocessing, which can simplify the model-building process and often leads to robust models that generalize well to unseen data. [68] However, while tree based models are less affected by multicollinearity, they can still impact the interpretability and importance scores of features. Hence, Lasso Regularization techniques should be employed within the models to mitigate any potential issues with multicollinearity, ensuring that the model remains robust and interpretable.

Lastly, besides lipid descriptors, the correlations between other features are low, indicating their independence, avoiding overfitting. For example, if two features in the production process model, such as FRR and CHIP, have low correlations it means that changes in CHIP is most likely not affecting the FRR. Yet, this interpretation needs to be taken

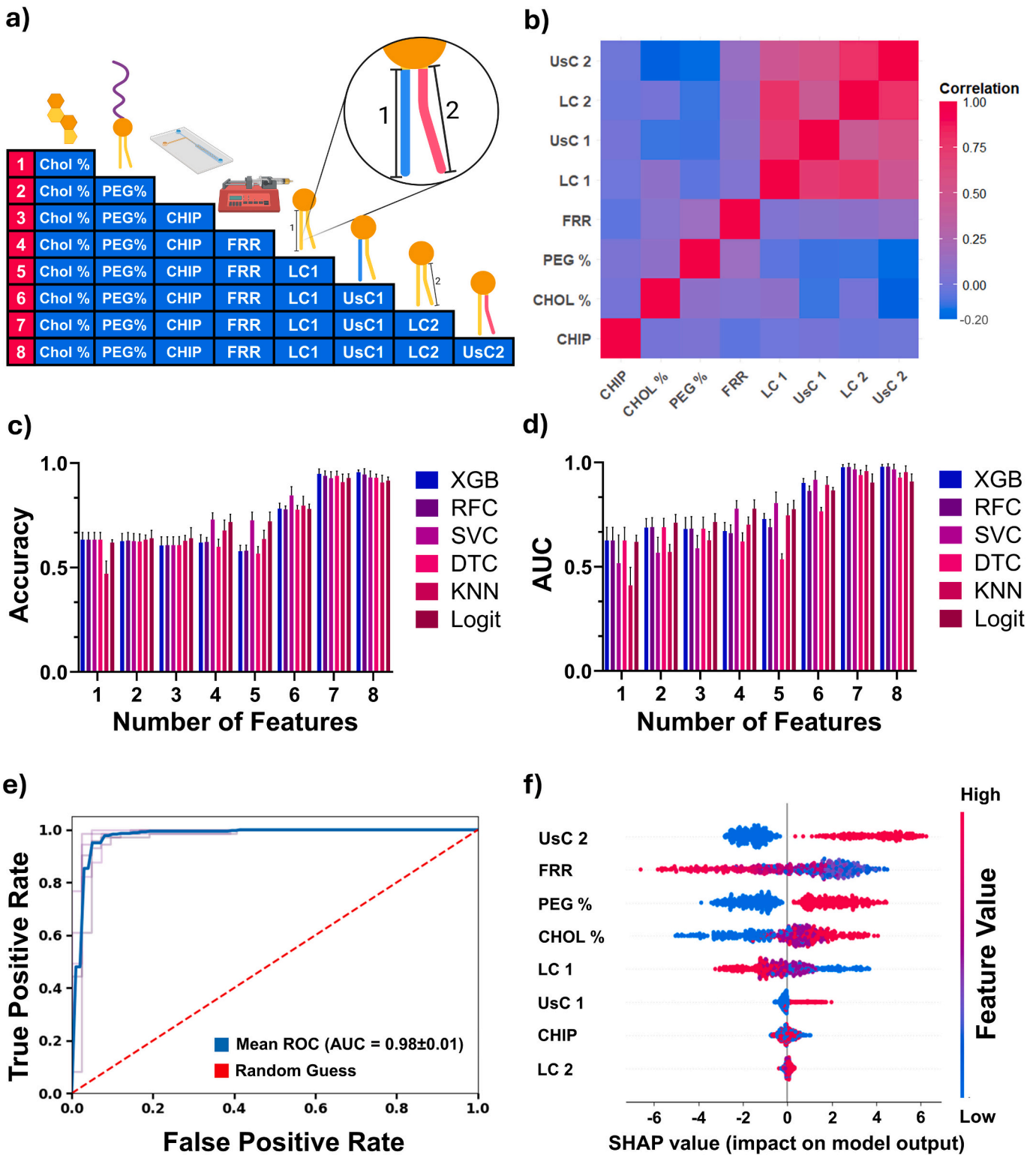

Fig. 3. Feature selection and predicting liposome formation: a) Schematic presentation of features used for modelling, including Cholesterol (CHOL), DSPE-PEG percentage (PEG%), Chip geometry, Flow rate ratio (FRR), Length of lipid chains 1 and 2 (LC1/2), and Saturation of lipid chains 1 and 2 (UsC1/2) indicated in blue and magenta respectively. b) Correlation heatmap of all input features. A magenta colour indicates an absolute correlation (=1) and a blue colour indicates a negative correlation. c) accuracy, and d) Area under the Curve (AUC) of different models for predicting liposome formation, averaged over 5 repeated tests with different data splits (5-fold cross-validation). e) Receiver operator curve (ROC) of 5-fold cross-validation for the XGB model predicting the formation of liposomes. f) SHAP analysis for the 8-feature XGB model illustrating the impact of each feature on the model's output "liposome formation", using a swarm plot of SHAP values. The dot colour represents the feature value (magenta for high, blue for low), while horizontal positioning shows the positive or negative contribution of each feature in each prediction instance. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

with caution as Spearman's correlation quantifies monotone dependences between two features. However, the independence allows the model to understand and predict the effects FRR and Chip have on the outcome without interference. It ensures that each feature contributes uniquely to the model, leading to more reliable and interpretable predictions. In brief, a feature vector of 8 process and formulation parameters, relevant to microfluidic liposome production, was chosen for further modelling (Fig. 3a).

2.3. Predicting formation and size of liposomes

2.3.1. Formation of liposomes

The formation of liposomes during microfluidic production is not always a guaranteed outcome. [69] It hinges on a complex interplay of formulation and process parameters. [69] Accurately predicting whether liposomes will form under specific conditions is therefore critical to enriching the pool of successful formulations. To address this challenge, six predictive models based on different underlaying architectures were trained on our dataset generated in-house. The model selection consisted of tree-based models such as decision trees (DT), random forest (RF), and extreme gradient boosting (XGB). Further, models based on support vector classifier (SVC), logistic regression classifier (Logit), and k-nearest neighbour (KNN) were trained. Investigating different models allows for a thorough evaluation of their performance and generalization capabilities, facilitating the selection of the most suitable model. [29,30] In the training process, the feature's impact on model accuracy, and Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) curves were evaluated as seen in Fig. 3c-d. ROC curves are graphical representations that show the performance of a classification model, plotting the true positive rate of a test data set against the false positive rate. [70] For example, if a model is predicting the formations of liposomes correctly, the false positive rate is low while the true positive rate approaches 1, leading to an AUC of 1, as seen in Fig. 3e. In short, the AUC measures how well a classifier distinguishes between classes, with higher values indicating better performance. [70]

Features were added incrementally (Fig. 3a) in the training process and for each addition, a 5-fold cross-validation was performed. The order of the feature addition was based on reported, [23,25,28,32,53,65] and experimentally observed influences of certain features during microfluidic production. In short, the feature vector was constructed of formulation features, followed by process features, and finalized with lipid descriptors.

Among the models tested, the XGB model, trained on all 8 features, consistently exhibited the highest accuracy, and AUC (Fig. 3e), along with minimal variation between cross-validation sets (Fig. 3c-d). Although the simpler DT model showed only slightly lower performance, the XGB model was deemed better suited for addressing subsequent regression problems, ensuring consistency in the modelling architecture throughout the study. This made XGB the preferred model architecture for the subsequent tasks. Furthermore, tree-based models as XGB are very amenable to model analysis due to their speed which makes XGB an attractive choice for XAI. [30]

The detailed ROC plot for the XGB model, shown in Fig. 3e, illustrates the model's ability to predict liposome formation with high discrimination power. This is further evidenced by the high AUC scores (Fig. 3e, blue line) consistently achieved by the XGB model across all 5 cross validations. However, as class imbalances in the data set could bias the predictions, the target data was scrambled, and the model reassessed on the scrambled data. Fig. S4 shows a pure random model indicating that class imbalances have not biased the original model.

While a functional model is fundamental, it is crucial to learn from the model and derive the findings on the development process. [55,71,72] For example, it can be evaluated what formulation and process features were key contributing factors to the outcome of the model. [71,72] Such an approach can justify the use of a reduced experimental design and therefore resulting in significant savings in time and cost. [49]

Tree-based models clearly describe the decision-making process, however direct tree interpretation is challenging. In practice, trees can reach enormous depth and size, and as a result, the collection of rules is vast and complicated (see Supplementary Material). [73] The importance of input features in predicting liposome formation was therefore determined using Shapley additive explanations (SHAP) [56,74] analysis for the 8-feature XGB model. SHAP analysis provides a visual interpretation to gain insights into feature importance and their influence on the model output.

The plot highlights the influence of various input features on the model's predictions and, consequently, potential correlations of these features on liposome formation. For each feature, datapoints are plotted in a dot plot based on their impact on the model prediction (SHAP value) and coloured by their feature value. The plot illustrates the relationship between each feature's value and its contribution to the model's prediction, allowing us to understand how changes in the feature value affect the model output and identify patterns such as whether higher or lower values of the feature increase or decrease the prediction. For instance, a high FRR results in a negative contribution to liposome formation, while high PEG content has a positive contribution. Hereby, the SHAP analysis offers insights into potential key drivers influencing liposome formation, such as the degree of unsaturation in the second lipid chain, PEGylation content, and cholesterol content. The degree of unsaturation in the second lipid chain, for example, is highlighted as a key driver of liposome formation. This can be explained as lipids with higher degrees of unsaturation are more likely to introduce fluidity into the bilayer, facilitating the formation of stable liposomes. [59] The analysis further revealed that a high FRR tends to negatively impact liposome formation. This is likely due to the enhanced amount of ethanol in the mixture which will prevent the liposomes from forming since it is not necessarily thermodynamically favorable. [63] This is well in line with the recent observations that unilamellar vesicles change their bilayer composition within milliseconds when exposed to certain ethanol concentrations. [63] However, it appears that lipids with high T_m are more likely affected. [63] Further, a high PEG content shows a positive contribution to liposome formation. PEGylation is known to enhance the stability of liposomes by providing steric protection, reducing aggregation, and improving circulation time in biological systems. [6] Cholesterol content, another important factor, plays a dual role by modulating membrane fluidity and stability, though its impact can vary depending on concentration. [60] On the other hand, certain factors negatively impact liposome formation. Saturated lipid chains tend to form more rigid bilayers, which can hinder the formation of stable liposomes, especially when combined with lower PEG content or reduced cholesterol levels. [59] The chip geometry, as often reported in the literature, seems to have a minimal effect on liposome formation. This could be due to the fact that we have operated with total flow rates that ensure consistent mixing. [28,64]

Hence, SHAP was used to identify and rank features impacting microfluidic liposome production which can each be explained by literature. It should be noted however, that the SHAP analysis shows the effect of each individual feature on model predictions independently, without considering potential synergies or interactions between input features.

In this section, we validated the XGB model as a robust tool for predicting liposome formation using our in-house dataset. The XGB model exhibited high discriminative power, achieving a mean AUC of 0.98 ± 0.01 across 5-fold cross-validations. Furthermore, the XGB model was suitable for SHAP analysis, an XAI technique that facilitates the analysis of feature importance and their impacts on model predictions. Our analysis highlighted significant influences on the model predictions from factors such as unsaturation of the second lipid chain, PEGylation content, and cholesterol content. These findings suggest that these parameters may play critical roles in liposome formation.

2.3.2. Predicting size

Having established a reliable model for predicting the formation of liposomes, the next step is to predict their size, a parameter of paramount importance in parenteral drug delivery. [7]

Given the significance of size in determining the therapeutic application, efficacy, and safety of liposomes, developing a model to predict liposome size during microfluidic production is essential. ML can be a powerful tool to predict CQAs such as size in therapeutic nanoparticle production. [50] Accurate size prediction facilitates the production of liposomes with tailored properties, enhancing their performance in clinical applications. [7,10,75]

To this end, the previously validated XGB model's architecture was adapted to a regression model to predict the size of liposomes. By incorporating key formulation and process parameters identified during the formation prediction, the model can be trained to predict liposome size. The selected parameters include lipid structure descriptors, Chol content, PEG content, Chip, and FRR.

To ensure a robust performance assessment, the model was evaluated in a 10-fold cross-validation. In this process, the dataset is split into ten equal parts (folds), with the model iteratively trained on nine folds and tested on the remaining fold. Following cross-validation, the coefficient of determination (R^2 value) was calculated to quantify the goodness of fit between the predicted and actual particle sizes. An R^2 value of 0.74 ± 0.08 indicates that approximately 75 % of the variance in the particle sizes can be explained by the model. This level of performance can be considered quite reliable in predictive modelling contexts, as it suggests that the model captures a substantial portion of the variability in the target variable. This evaluation framework offers valuable insights into the model's accuracy, robustness, and generalization ability, supporting informed decision-making in particle size prediction tasks.

Further, the relationship between experimentally measured particle size and predicted values is visualized through a scatter plot with a linear regression curve in Fig. 4a. This plot provides insights into the model's strong predictive capability, with most points clustering around the regression line. Some deviations and outliers, particularly at higher sizes, highlight areas of lesser accuracy. However, these areas are outside the typical size range for therapeutic applications, which is generally between 30 and 120 nm. In this range, predictive capability was found to be higher (R^2: 0.78). Also, residual plots for the regression tasks, as shown in Fig. S7, provide valuable insights into model performance. In Fig. S7a, residuals for predictions within the typical liposome size range (30-180 nm) are closely clustered around the zero line, indicating strong predictive accuracy. However, predictions for sizes above 180 nm show greater deviations, likely due to limited data for larger liposomes. Additionally, there is a tendency for the model to overestimate outliers in the lower size range.

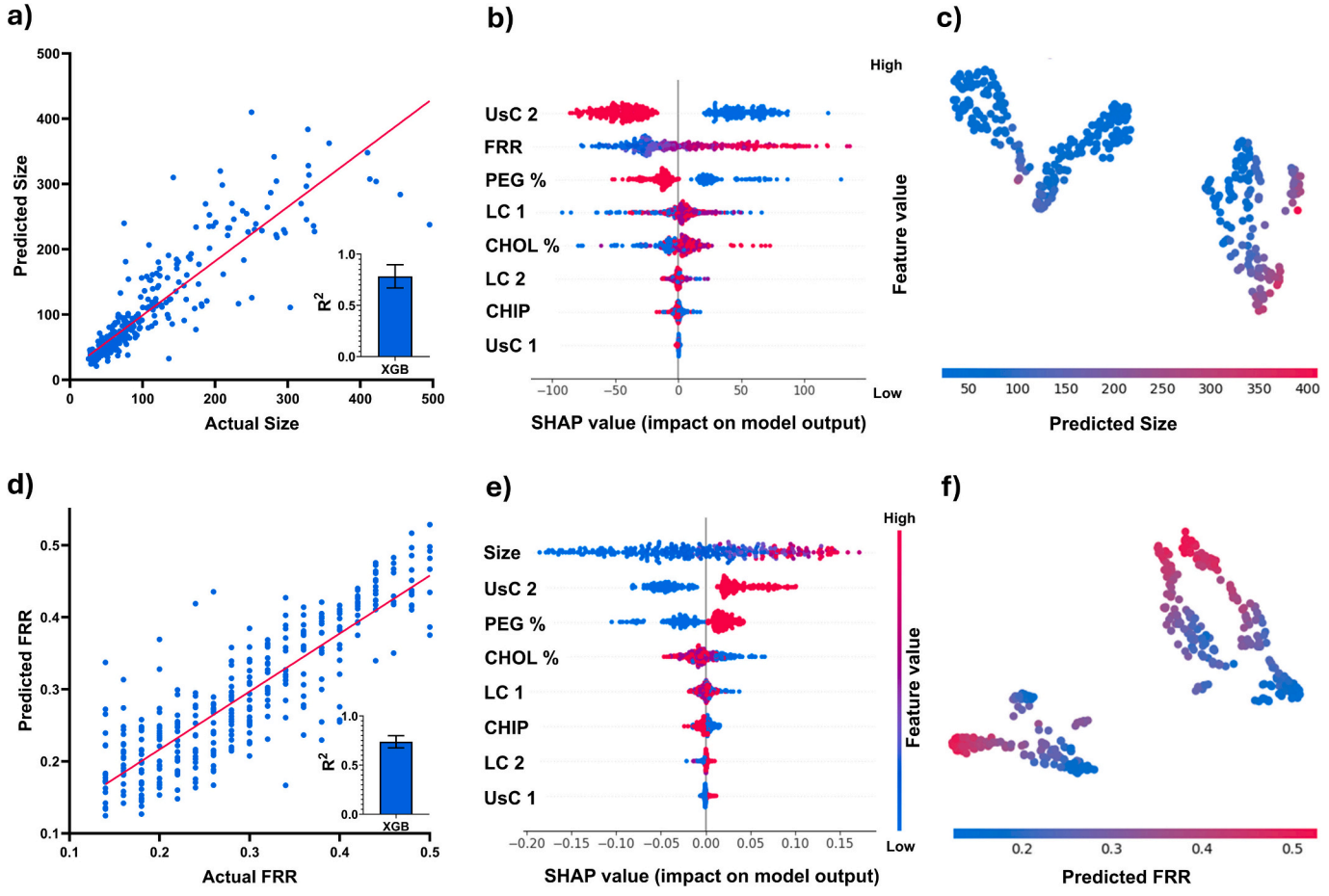

Fig. 4. Prediction of the liposome's size and the flow rate ratio (FRR). Model performance depicted as relationship between the experimental a) particle size, or d) FRR, and predicted values. R^2 value over 10-fold cross-validation quantifies the goodness of fit between the predicted and actual particle sizes or FRR respectively. SHAP analysis for the 8-feature XGB regression model illustrates the impact of each feature on the model's output b) "liposome size" or e) "FRR", using a swarm plot of SHAP values. Dot colour represents the feature value (magenta for high, blue for low), while horizontal positioning shows the positive or negative contribution of each feature in each prediction instance. Two-dimensional visualization of the SHAP values calculated for the input features of the XGB model for particle c) size or f) FRR. The SHAP values for the 8 input features were condensed into two principal components using principal component analysis (PCA) and then grouped together using t-distributed Stochastic Neighbour Embedding (t-SNE). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The importance of each input feature for liposome size prediction was investigated via SHAP analysis and visualized in a beeswarm plot (Fig. 4b) which shows the relationship between individual feature values, SHAP values (a measure of feature impact), and predicted sizes. Each point represents an instance in the dataset, with the x-axis displaying the SHAP value for the predicted size. The features are arranged from top to bottom in order of their decreasing impact on the model's output. This plot reveals how changes in feature values affect predictions, highlighting the importance and influence of different features. Features with a larger spread over the x-axis, such as Saturation of lipid chain 2 (UsC2), Flow rate ratio (FRR), DSPE-PEG percentage (PEG%), Length of lipid chain 1 (LC1), and Cholesterol content (CHOL%), may have a more significant impact on the model's predictions. These are potential key formulation and production factors that determine the size of a liposome. For instance, a high (magenta) degree of unsaturation in the lipid chain 2 and a high PEG content seem to have a negative contribution to the size output. This suggests that lipids with a higher degree of unsaturation and higher PEG content may lead to the formation of smaller liposomes. The unsaturated chains likely introduce more fluidity and flexibility into the lipid bilayer, facilitating more narrow bilayer curvatures and hence the formation of smaller, more compact liposomal structures. [76] Similarly, PEGylation is known to influence liposome size by sterically stabilizing the liposomes and potentially limiting their growth during the production process. [77] This trend might change with PEG >8 % where its composition switches from mushroom to brush. [77] Hence, the presented model is likely suited to predict liposomes containing PEG <5 %. Conversely, factors such as a high flow rate ratio (FRR), the length of the first lipid chain, and cholesterol content appear to contribute positively to liposome size. A high FRR might promote larger liposome formation by reducing polarity and thereby the thermodynamic pressure encouraging the aggregation of lipids into larger vesicles. [63] The length of the first lipid chain could influence the overall membrane structure, with longer chains possibly leading to larger liposomes due to increased hydrophobic interactions within the bilayer. [63] Cholesterol, a well-known modulator of membrane fluidity and stability, might also contribute to larger liposome sizes by ordering the lipid bilayer and reducing membrane curvature, which could promote the formation of larger vesicles. [60,63]

The presence of clusters or patterns in the SHAP plot, particularly those related to UsC2 and PEG%, indicates that the model has identified groups of instances where these features consistently influence liposome size in a similar manner. These clusters suggest that the model not only recognizes individual feature contributions but also discerns underlying patterns and relationships between features that are critical for accurate size prediction.

This analysis underscores the complex and multifactorial nature of liposome size determination. While certain formulation and process parameters, such as lipid chain saturation, PEG content, FRR, and cholesterol, play significant roles in determining liposome size, it is the interaction between these factors that ultimately dictates the outcome. The observed patterns align with existing literature, which consistently highlights the interdependent nature of these variables in liposome formation and size regulation. [17,23,24] The insights provided by the SHAP analysis emphasize the importance of considering both individual and combined effects of formulation and production parameters to fully understand and predict liposome size in microfluidic production systems.

For enhanced interpretability, SHAP values of all 8 input features were reduced using Principal Component Analysis (PCA) followed by t-distributed Stochastic Neighbour Embedding (t-SNE) which is an approach in cases where the data is initially multi-dimensional. PCA serves to denoise and compress the data to 5 dimensions, making the t-SNE step more computationally efficient and effective at revealing meaningful structures in the data. This approach allows us to understand how different features contribute to the model's predictions, revealing patterns that might not be visible with simpler visualizations as shown in Fig. 4c. The SHAP values for the 8 input features are condensed into two principal components and visualized in a scatter plot. Each point represents an instance in the dataset, positioned by the t-SNE algorithm, and coloured according to the predicted size. Clusters of nearby points indicate instances with similar feature contributions, revealing patterns within the data. The Colour coding highlights regions with similar predicted sizes, providing further insights into prediction patterns. Isolated or distant points may indicate anomalies or unique feature contributions. Larger sizes (magenta) are clearly clustered, indicating that the model effectively distinguishes between particle sizes. To ensure that the projection was faithful to the structure of the original data we have calculated the Kullback-Leibler (KL) divergence value. The low KL divergence value (0.16) that the dimensionality reduction technique embedding is a good representation of the original data, meaning that the clusters or patterns observed in Fig. 4c are likely reflective of true structures in the multi-dimensional space.

Overall, we demonstrate a robust model predicting liposome size with high predictive capability (R^2: 0.78). SHAP analysis yields practical insights, such as identifying influential formulation and process features that clarify the model's decision-making process and key factors driving predicted sizes. Evaluating prediction variations across segments validates the model's performance and its ability to capture underlying data patterns.

2.3.3. Models to predict microfluidic process parameters

Translating liposome production from the laboratory to clinical and industrial applications poses significant challenges. [18,78,79] Scaling up production while maintaining consistency and reproducibility is a primary concern. [18] With traditional liposome production methods often failing to scale effectively, prompting a shift towards microfluidic techniques due to their sustainability and applicability in industry. [16] However, this transition is not straightforward. Each liposomal formulation can respond differently to microfluidic processes, necessitating extensive trial and error and laborious optimization on a case-to-case basis. [16] Furthermore, variations in microfluidic chip design complicate the standardization of production processes. [64] Robustly 3D-printing these chips might offer a pathway to standardization, facilitating broader application in pharmaceutical manufacturing. [25] Additionally, microfluidic production depends on many controllable parameters (Formulation, Chip, FRR, etc.), but this makes the design space multi-dimensional and impractical to empirical testing. ML can help to efficiently navigate the multi-dimensional design space and help us find the correct parameters to produce liposomes with desired properties in a microfluidics production. Predicting CPPs and understanding their influence on CQAs are of profound interest in different branches of pharmaceutical sciences. [49,51] A predictive model that takes formulation parameters and chip geometry into account to determine the optimal FRR for microfluidic production, targeting desired CQA like size, can significantly ease the transition to microfluidics and scale-up. By accurately predicting FRR, such a model can streamline development, reducing the need for extensive empirical optimization. Building on previous analyses, key features influencing liposome formation and size have been identified. Namely, formulation parameters such as unsaturation and length of the lipid chains, cholesterol and PEGylation content along with the liposome's desired size and the used chip geometry. These features serve as the foundation for developing an XGB model to predict the most critical process parameter, FRR, ensuring comprehensive optimization of the microfluidic process. However, it is important to note that FRR is, by the nature of the process, an independent variable typically used as a predictor rather than a target. To predict process parameters such as FRR, an inverse prediction approach is required, where the independent variable is treated as the target. Inverse prediction is an emerging field in the prediction of process parameters within engineering and pharmaceutical sciences. [80,81] It has for example been demonstrated by Rouco et al. that AI can be leveraged as a powerful tool in performing inverse prediction of CPPs

for the production of lipid nanocarriers produced by shear homogenization. [81]

The relationship between the experimental and predicted FRR values for a given liposome batch is visualized in a scatter plot (Fig. 4d), showing a linear trend, implying that the predicted FRR for a certain liposome batch is in accordance with the experimental. The model's performance is further evaluated using 10-fold cross-validation, involving 10 repeated tests with different data splits. Following cross-validation, the R^2 value of 0.74 ± 0.06 indicates that the model explains approximately 75 % of the variance in FRR, demonstrating reliable performance in predicting the FRR. A beeswarm plot (Fig. 4e) of calculated SHAP values reveals the impact of the feature "liposome size" as the main driver for FRR prediction before the unsaturation of lipid chain 2 (UsC2). It is evident that larger liposomes, high UsC2, and PEG% along with a low Chol % have a positive contribution to the model's output. This might imply that both, chemical formulation and morphological factors play a major role in selecting a FRR to obtain a desired liposome in a microfluidic process. Further, multi-dimensional SHAP values are reduced using PCA and t-SNE (Fig. 4f) and visualized to reveal data patterns and segments. It is visible that clear gradients of FRR ratio span through the clusters, implying the model's ability to making decision and thereby further justifies the feature vector.

This promising predictive model for FRR has the potential to ease the transition to microfluidic production of liposomes, ensuring consistent and scalable manufacturing processes. By predicting FRR, this model reduces dependency on empirical optimization, accelerating the transition from bench-scale experiments to industrial production. However, it needs to be pointed out that an inverse prediction needs wet lab validation to ensure its robustness. Nevertheless, this advancement holds immense potential to streamline the development of liposomal

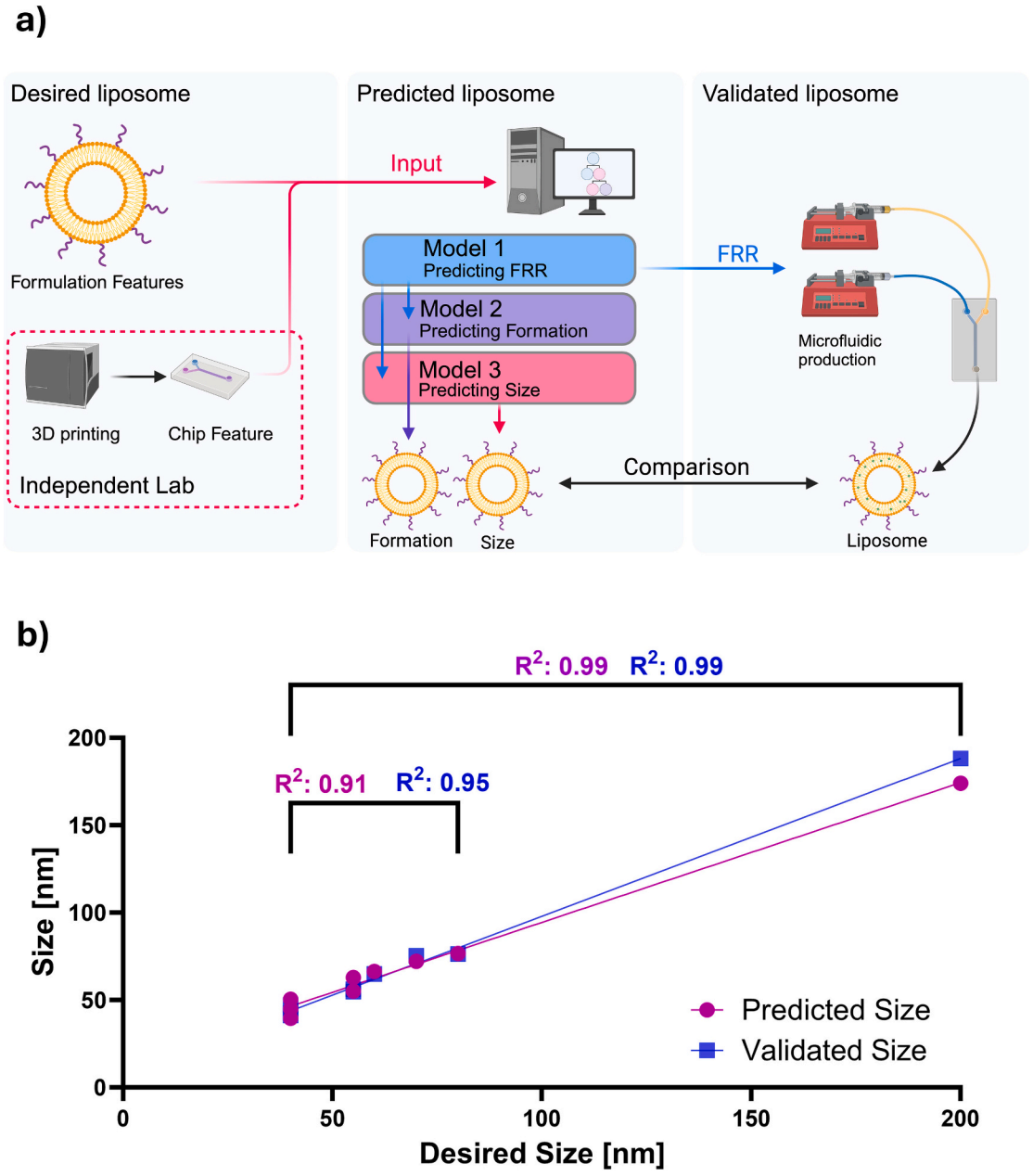

Fig. 5. Transition workflow for validation experiments. a) The desired liposome and lab setup were fed to model to predict process parameters which were fed to consecutive models to predict CQAs such as the formation and size of liposomes that were compared to liposomes from microfluidic production. b) Wet lab validation for 10 liposome formulations with desired input composition and size, predicted size, and size produced during the validation are plotted against the desired input size.

drug delivery systems, facilitating their journey from the laboratory to clinical and industrial settings. It represents a significant step forward in optimizing these systems, promising improved therapeutic outcomes and broader accessibility.

2.3.4. Wet lab validation

The last step involved validating our XGB models by assessing how effectively they can support the transition of liposome production from traditional production methods to microfluidics. Hence, to validate the models, a transition workflow from traditional to microfluidic liposome production was executed in the wet lab, employing the predictive models through a series of experiments as illustrated in Fig. 5a. The validation process involved randomly selecting 10 distinct liposome formulations, each assigned a specific desired size. Additionally, a chip geometry (i.e. ZZ or SR) [25] was chosen, and printed in an independent laboratory. This further allowed to account for potential variability due to 3D printing.

The workflow begins with a model that predicts the FRR required for each formulation, based on the input data (Table S2). Following this, a second model evaluates whether liposomes will form under the specified process conditions for the given lipid formulation. Finally, a third model predicts the resulting liposome size. Once the predictive accuracy of the model was confirmed - illustrated in Fig. 5b (magenta) as the relationship between desired and predicted sizes resulting in a good fit - the liposomes were produced using the predicted process parameters (Table S2). The actual sizes of the produced liposomes were measured using DLS and compared to both the predicted and desired sizes. The results in Fig. 5b (blue) show the relationship between the obtained sizes from the wet lab and the desired sizes, demonstrating that the predictions and validations closely align with the desired sizes, with deviations of only a few nanometres (Table S2). This indicates that the model's predictive power is strong. Hence, this validation process underscores the model's robustness in predicting critical parameters for liposome formation and size, ensuring a seamless transition from traditional to microfluidic production methods. Furthermore, the model proves highly valuable in adapting existing liposomal formulations to microfluidic processes, significantly streamlining the development of advanced liposomal drug delivery systems.

3. Conclusion

In this study, we demonstrated that the integration of predictive modelling with wet lab validation underscores the potential of ML for precise and efficient optimization of microfluidic liposome production, capable of enhancing the development effort, scalability and transitioning to microfluidic technology in pharmaceutical sciences. This approach not only streamlines the production process but also paves the way for similar models in drug encapsulation within liposomes and the expansion to commercial microfluidic chips. Future research could explore the application of such models in also optimizing drug release profiles, ensuring that liposomal formulations deliver therapeutic agents effectively and consistently. The advancements in ML-driven microfluidic production hold promise to accelerate the development of drug delivery systems, making them more accessible and adaptable to a wide range of pharmaceutical applications.

4. Materials and methods

4.1. Materials

The phospholipids DMPC (1,2-dimyristoyl-*sn*-glycero-3-phosphatidylcholine), DOPC (1,2-dioleoylglycerol-3-phosphorylcholine), DPPC (1,2-dipalmitoyl-*sn*-glycero-3-phosphatidylcholine), HSPC (hydrogenated soy phosphatidylcholine), DSPE-PEG2000 (1,2-distearoyl-*sn*-glycero-3-phosphoethanolamine-N-[methoxy(polyethylene glycol) monosodium salt), POPC (1-palmitoyl-2-oleoyl-glycero-3-phosphocholine) were purchased from Lipoid, Germany. Cholesterol was purchased from Sigma Aldrich, USA. Purified water was obtained using a Barnstead Smart2Pure system from Thermo Fisher Scientific Inc., Germany. PBS (phosphate buffered saline, pH 7.4) was obtained using KCl (potassium chloride), KH_2PO_4 (Potassium Dihydrogen Phosphate), NaCl (sodium chloride), NaH_2PO_4 (Sodium Phosphate Monobasic), Na_2HPO_4 (Sodium Phosphate Dibasic), NaOH (sodium hydroxide) from Carl Roth LLC, Germany. NaOH (sodium hydroxide in pellet form) from Haenseler, Switzerland. Absolute Ethanol (EtOH) was purchased from VWR International SAS, France. Ethanol denatured with ketone (EtOH 94 %) was purchased from Dr. Grogg Chemie AG, Switzerland. Methanol (MeOH) was purchased from Fisher Chemical, Belgium.

4.2. Methods

4.2.1. Data collection

For the liposome screening, formulations were collected following a Design of Experiment (DoE) ensuring evenly distributed data points. A multi-level full factorial design plan was created on Minitab 2018 (Minitab Inc., USA) for each considered lipid (DMPC, DOPC, DPPC, HSPC, POPC). The design is characterized by a categorical factor, the chip design coded as ZZ (zigzag bas-relief) or SR (split and recombine channels), and by four numerical factors, FRR, TFR, cholesterol and PEG content as illustrated in Table 1. A second dataset was created to cover a broader FRR spectrum holding the TFR constant at 12 mL/min, ensuring constant mixing, and varying the FRR over 20 levels. Every point within the design space was characterized by one measurement.

For the screening, lipid masses were calculated, weighed, and dissolved in ethanol (≥99.8 % *v*/v) at a total lipid concentration of 5 mg/mL. The aqueous buffer used was PBS with a pH of 7.4. For the preparation of liposomes via microfluidics, the chip, positioned in a 55 °C water bath, was connected to two syringes, which were mounted on two NE-1010 Higher Pressure Syringe Pump, KF TECHNOLOGY SRL, Italy. The connection was established through polyethylene tubing. A specific amount of lipid was delivered against PBS at controlled flow rates. Various amounts of DSPE-PEG 2000, ranging from 0 % to 5 %, and Cholesterol, ranging from 0 %-40 %, were added to the lipids and were included in the screening. The two syringes containing the respective phases were preheated to a temperature of 62 °C to remain above the phase transition temperature. The microfluidic chips from the University of Urbino, Italy, were manufactured using ultrafuse polypropylene (PP) with an Ultimaker 3D printer from Ultimaker, The Netherlands (Fig. S1). After discarding the initial 1 mL from the outlet of the chip, the resulting samples were collected.

The chip cleaning process involved sequential flushing steps, each carried out consecutively (10 mL of NaOH 1 M, 10 mL of NaOH 0.1 M, 20 mL of MQ Water, 10 mL of EtOH 94 %). Upon the conclusion of the procedure, the chip was dried and flushed using 6 mL of methanol. 1 mL of methanol was collected and analyzed via HPLC CAD [82] to validate the process and confirm the chip's cleanliness with an UltiMate 3000 HPLC from Thermo Fisher Scientific Inc., Germany (Fig. S4). To conclude the cleaning process, the chip was washed with 10 mL of EtOH 94 % and dried.

Immediately following liposome production, the sample underwent

Table 1
Factor levels for DoE used to ensure balanced data collection. *Used for secondary screening to give more emphasis to FRR.

	CHIP	FRR	TFR [mL/min]	Col n %	PEG n %
Levels	N/A	0.2, 0.33, 0.5 (0.14-0.5, in 0.02 steps)*	8, 10, 12 (12)*	0, 20, 40	0, 5
Categorical levels	ZZ, SR	N/A	N/A	N/A	N/A

analysis via Dynamic Light Scattering (DLS). DLS was utilized to evaluate the population, size, and polydispersity index (PDI) of liposomes using the Litesizer 500 (Anton Paar GmbH, Austria) with a 173° backscatter angle and a 633 nm helium–neon-laser. Particle size measurements were performed in PBS (Refractive index: 1.335) at a pH of 7.4 and a temperature of 25 °C. The measurements were corrected to the viscosity of the according sample. After equilibrating the sample, the measurement (10 runs × 10 s) was performed. If the intensity size distribution of the liposome was unimodal, the autocorrelation function was analyzed according to the cumulant method by the Kalliope™-software. For our model we considered samples below 500 nm and a monodisperse size distribution as forming liposomes.

4.2.2. Data curation

In an initial screening, 811 data points were collected, additional 571 data points were collected to give more emphasis on the variation of FRR in order to create more meaningful predictions which were then used for modelling. Each record is characterized by at least eight features. The features include both categorical variables (e.g. CHIP), continuous variables (e.g. CHOL %, PEG %, FRR %) and simple molecular lipid descriptors (e.g. lipid chain length). Target variable in the classification task was Population, representing the formation of uniform liposomes. Target variable in the regression task was Size, representing the measured size of liposomes in nanometres. Further, FRR was used as target variable for an inverse prediction, representing the process parameter flow rate ratio.

The data is organized into a structured format, with each row representing an individual sample (i.e., a single experimental run or observation), and each column representing a specific feature or the target variable. The categorical feature 'CHIP' indicates the type of chip used in the experiment, which was one-hot encoded during preprocessing to ensure compatibility with the machine learning models.

The dataset used for ML model development underwent a preprocessing step to ensure data quality and consistency. Initially, any observations with the 'Size' variable exceeding 500 nm were filtered out to focus the analysis on liposomes within a specific size range. Also, samples with PDIs above 0.3 were considered as "no formation"as the consensus in parenteral drug delivery. The remaining data was then prepared for model training and evaluation. Additionally, observations associated with a multidisperse size distribution (Population > 1) were excluded from the dataset for size and FRR modelling (prescreening:575 and FRR Data:357). Potential class imbalances were addressed by scrambling the target variable for all data points and retraining the classification problem as indicated in Fig. S4.

4.2.3. TMAP

For each data point, the approximate nearest neighbours in the feature space, consisting of eight features (i.e. Cholesterol amount (CHOL), DSPE-PEG amount (PEG%), Chip geometry, Flow rate ratio (FRR), Length of lipid chains 1 and 2 (LC1/2), and Saturation of lipid chains 1 and 2 (UsC1/2)) were extracted and used to calculate the TMAP layout. The feature FRR was excluded for the response of FRR and replaced by the feature Size. The Dataset used consisted of 571 unique data points. The resulting layout was displayed in a TMAP using the open-source Faerun package (http://tmap.gdb.tools). [58] Thorough documentation and Validation of the TMAP method using benchmark data sets can be found elsewhere. [58]

4.2.4. ML model development, evaluation and interpretation

The ML models were given the input features comprising of 'CHOL %', 'DSPE-PEG %', 'FRR', and parameters related to chain length and unsaturation. The predictive process initiates with the transformation and standardization of feature data, ensuring its readiness for modelling. Categorical variables, such as the 'CHIP' parameter, were one-hot encoded to convert them into a suitable format for model training. For the models trained to predict the FRR the feature 'FRR' was replaced by 'Size'. Python module scikit-learn and xgboost were mainly used for modelling. L1 Regularization (Lasso Regularization) technique was used for XGB models helping to mitigate collinearity. L1 regularization adds a penalty equal to the absolute value of the magnitude of the coefficients. It can reduce some coefficients to zero, effectively selecting a subset of features and thus handling multicollinearity. Alpha was set to 0.1 after preliminary tests. This was chosen to prevent overfitting by penalizing large weights, promoting sparsity in the model. The epochs for boosting were set to 100.

The performance of the developed machine learning (ML) models was assessed using a stratified K-fold cross-validation approach implemented through the scikit-learn library in Python. This approach was employed differently for classification and regression models, based on the nature and size of the respective datasets.

For the classification task, we trained six models simultaneously on a larger dataset while incrementally adding features. Hence, we opted for a 5-fold cross-validation strategy. This choice allowed to perform fewer iterations, thus reducing computational time while accepting a slightly higher bias in exchange for lower variance. The performance of each model was evaluated using metrics such as accuracy, precision, and the area under the curve (AUC) of the receiver operating characteristic (ROC) curve for each fold.

In contrast, the regression task involved a smaller dataset, as it was limited to liposomes that successfully formed, whereas the classifier utilized data from both forming and non-forming liposomes. To better estimate the model's performance with reduced bias, we employed a 10-fold cross-validation approach. This method allowed the regression models to train on a larger portion of the data in each fold (90 % for training, 10 % for testing), which is particularly beneficial for smaller datasets. The performance of the regression models was evaluated primarily through their R-squared (R^2) values.

Additionally, SHAP [56] (SHapley Additive exPlanations) analysis was conducted to interpret the model's predictions and understand the relative importance of each input feature. SHAP values were computed to quantify the contribution of each feature to the predicted output. Visualization techniques, including SHAP summary plots and scatter plots, were employed to visualize the relationships between predicted and actual liposome sizes, providing insights into the model's performance and feature importance. Furthermore, Spearman's correlation heatmap were generated to explore the correlation between input features and liposome size.

4.2.5. Simple supervised clustering

Simple supervised clustering was done using the SHAP toolbox for clustering (https://shap.readthedocs.io/). In brief, SHAP values of all features in the model were computed to provide a comprehensive view of feature importance across all inputs. Principal Component Analysis (PCA) was used to reduce the dimensionality of all input feature's SHAP values to 5 components. PCA was chosen because it effectively reduces noise and compresses the data while preserving as much variance as possible. Following PCA, t-SNE was applied to further reduce the data to 2 dimensions. t-SNE was chosen because it is particularly effective at preserving local structure and revealing clusters in the data. Even though PCA had already reduced the dimensionality, t-SNE helps to refine this reduction by focusing on the relationships between nearby points, which can uncover meaningful clusters that PCA might not fully reveal. Kullback-Leibler divergence was determined to ensure a faithful projection.

4.2.6. Validation

The validation process involved selecting ten distinct liposome formulations, each assigned a specific desired size all within the ranges specified in Table 2. These ranges were chosen based on typical excipient use for parenteral liposomes. Although the compositions were generally randomized, factors such as lipid types, CHOL %, and PEG % were strategically varied to ensure diversity in the validation test set.

Table 2
Input ranges for wet lab validation.

Input					
	Lipid	CHOL %	PEG %	Size [nm]	CHIP
Upper limit	N/A	40	5	30	N/A
Lower limit	N/A	0	0	200	N/A
Categorical	DPPC, DMPC, DOPC, POPC, HSPC	N/A	N/A	N/A	SR, ZZ

The size range was limited to 200 nm, aligning with the typical size range for parenteral liposomes. [83]

Each formulation was also assigned a specific chip geometry (i.e., SR or ZZ), which was then 3D printed. The microfluidic chips were produced in a separate and independent lab using a Stratasys Eden260VS 3D printer (Stratasys, Germany) and the photopolymer IORA Model White (iSquared2, Switzerland). Predictive models were employed to forecast the FRR, assess the likelihood of liposome formation, and predict the resulting liposome size. Liposomes were subsequently produced using these predicted process parameters. The size of the produced liposomes was measured via DLS, as previously described, and compared to both the predicted and desired sizes. Linear regression was used to quantify the fit of the relationship between predicted, actual and desired size using GraphPad Prism 10.3.1 (GraphPad Software, USA).

5. Statistics and reproducibility

All statistical analyses performed are available for each experiment in the results & discussion section of the manuscript. No data were excluded from the analysis. GraphPad Prism 10.2.3 (GraphPad Software, USA) and R 4.4.0 (The R Foundation, Germany) were used for statistical analysis.

Code availability

The source codes that support the findings of this study are publicly available on GitHub (https://github.com/Luciani-Group/micro_fluidic_Liposome). Code was written and executed in Python 3.11.9 in an Anaconda 24.3.0 environment. Python packages and libraries used are disclosed on GitHub.

CRediT authorship contribution statement

Remo Eugster: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Markus Orsi:** Writing – review & editing, Visualization, Methodology, Investigation, Formal analysis. **Giorgio Buttitta:** Writing – review & editing, Validation, Methodology, Investigation, Conceptualization. **Nicola Serafini:** Validation, Investigation. **Mattia Tiboni:** Writing – review & editing, Methodology. **Luca Casettari:** Writing – review & editing, Methodology, Funding acquisition. **Jean-Louis Reymond:** Writing – review & editing, Methodology, Funding acquisition. **Simone Aleandri:** Writing – review & editing, Supervision, Methodology, Data curation, Conceptualization. **Paola Luciani:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

No private study sponsors had any involvement in the study design, data collection, or interpretation of data presented in this manuscript. P. L. declares the following competing interests: has consulted and received research grants from Lipoid GmbH, Sanofi-Aventis Deutschland and DSM Nutritional Products Ltd.; received research grants from PPM Services S.A.

Acknowledgement

Prof. Dr. Manuela Eugster from the ARTORG Center for Biomedical Engineering Research at the University of Bern, for printing the chips used for the wet lab validation.

Figs. 1 & 5 were created with Biorender.com.

This work has been partially funded by the European Union – NextGenerationEU - under the Italian Ministry of University and Research (MUR) National Innovation Ecosystem grant ECS00000041 – VITALITY - CUP [H33C22000430006].

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jconrel.2024.10.065.

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information file. The complete datasets used for training and testing the ML models are publicly available as .csv files on GitHub (https://github.com/Luciani-Group/micro_fluidic_Liposome).

References

[1] Y. Barenholz, Doxil® - the first FDA-approved nano-drug: lessons learned, J. Control. Release 160 (2012) 117–134, https://doi.org/10.1016/j.jconrel.2012.03.020.
[2] S. Giordani, V. Marassi, A. Zattoni, B. Roda, P. Reschiglian, Liposomes characterization for market approval as pharmaceutical products: analytical methods, guidelines and standardized protocols, J. Pharm. Biomed. Anal. 236 (2023) 115751, https://doi.org/10.1016/j.jpba.2023.115751.
[3] G. Buttitta, S. Bonacorsi, C. Barbarito, M. Moliterno, S. Pompei, G. Saito, I. Oddone, G. Verdone, D. Secci, S. Raimondi, Scalable microfluidic method for tunable liposomal production by a design of experiment approach, Int. J. Pharm. 662 (2024) 124460, https://doi.org/10.1016/j.ijpharm.2024.124460.
[4] R. Eugster, A.A. Ganguin, A. Seidi, S. Aleandri, P. Luciani, 3D printing injectable microbeads using a composite liposomal ink for local treatment of peritoneal diseases, drug Deliv, Transl. Res. 14 (2023) 1567–1581, https://doi.org/10.1007/s13346-023-01472-y.
[5] A. Dolor, P. Kierstead, Z. Dai, F.C. Szoka, Sterol-modified PEG lipids: alteration of the bilayer anchoring moiety has an unexpected effect on liposome circulation, Chem. Commun. 54 (2018) 11949–11952, https://doi.org/10.1039/c8cc05011b.
[6] M. Petrini, W.J.M. Lokerse, A. Mach, M. Hossann, O.M. Merkel, L.H. Lindner, Effects of surface charge, pegylation and functionalization with dipalmitoylphosphatidyldiglycerol on liposome– cell interactions and local drug delivery to solid tumors via thermosensitive liposomes, Int. J. Nanomedicine 16 (2021) 4045–4061, https://doi.org/10.2147/IJN.S305106.
[7] M. Danaei, M. Dehghankhold, S. Ataei, F. Hasanzadeh Davarani, R. Javanmard, A. Dokhani, S. Khorasani, M.R. Mozafari, Impact of particle size and polydispersity index on the clinical applications of lipidic nanocarrier systems, Pharmaceutics 10 (2018) 1–17, https://doi.org/10.3390/pharmaceutics10020057.
[8] G. Lou, G. Anderluzzi, S. Woods, C.W. Roberts, Y. Perrie, A novel microfluidic-based approach to formulate size-tuneable large unilamellar cationic liposomes: formulation, cellular uptake and biodistribution investigations, Eur. J. Pharm. Biopharm. 143 (2019) 51–60, https://doi.org/10.1016/j.ejpb.2019.08.013.
[9] J. Tang, M. Rakshit, H.M. Chua, A. Darwitan, L.T.H. Nguyen, A. Muktabar, S. Venkatraman, K.W. Ng, Liposome interaction with macrophages and foam cells for atherosclerosis treatment: effects of size, surface charge and lipid composition, Nanotechnology 32 (2021) 505105, https://doi.org/10.1088/1361-6528/ac2810.
[10] C. Oussoren, J. Zuidema, D.J.A. Crommelin, G. Storm, Lymphatic uptake and biodistribution of liposomes after subcutaneous injection. II. Influence of liposomal size, lipid composition and lipid dose, Biochim. Biophys. Acta 1328 (1997) 261–272, https://doi.org/10.1016/s0005-2736(97)00122-3.
[11] T. Lammers, Nanomedicine tumor targeting, Adv. Mater. (2024), https://doi.org/10.1002/adma.202312169.
[12] T. Duong, S.T. Proulx, P. Luciani, J.C. Leroux, M. Detmar, P. Koopman, M. Francois, Genetic ablation of SOX18 function suppresses tumor lymphangiogenesis and metastasis of melanoma in mice, Cancer Res. 72 (2012) 3105–3114, https://doi.org/10.1158/0008-5472.CAN-11-4026.
[13] S. Aleandri, L. Rahnfeld, D. Chatzikleanthous, A. Bergadano, C. Bühr, C. Detotto, S. Fuochi, K. Weber-Wilk, S. Schürch, P. van Hoogevest, P. Luciani, Development and in vivo validation of phospholipid-based depots for the sustained release of bupivacaine, Eur. J. Pharm. Biopharm. 181 (2022) 300–309, https://doi.org/10.1016/j.ejpb.2022.11.019.

[14] V. Forster, R.D. Signorell, M. Roveri, J.-C. Leroux, Liposome-supported peritoneal dialysis for detoxification of drugs and endogenous metabolites, Sci. Transl. Med. 6 (2014) 141–151, https://doi.org/10.1126/scitranslmed.300913.
[15] D.E. Large, R.G. Abdelmessih, E.A. Fink, D.T. Auguste, Liposome composition in drug delivery design, synthesis, characterization, and clinical application, Adv. Drug Deliv. Rev. 176 (2021) 113851, https://doi.org/10.1016/j.addr.2021.113851.
[16] S. Shah, V. Dhawan, R. Holm, M.S. Nagarsenker, Y. Perrie, Liposomes: advancements and innovation in the manufacturing process, Adv. Drug Deliv. Rev. 154 (2020) 102–122, https://doi.org/10.1016/j.addr.2020.07.002.
[17] C.B. Roces, G. Lou, N. Jain, S. Abraham, A. Thomas, G.W. Halbert, Y. Perrie, Manufacturing considerations for the development of lipid nanoparticles using microfluidics, Pharmaceutics 12 (2020) 1095, https://doi.org/10.3390/pharmaceutics12111095.
[18] C. Webb, N. Forbes, C.B. Roces, G. Anderluzzi, G. Lou, S. Abraham, L. Ingalls, K. Marshall, T.J. Leaver, J.A. Watts, J.W. Aylott, Y. Perrie, Using microfluidics for scalable manufacturing of nanomedicines from bench to GMP: a case study using protein-loaded liposomes, Int. J. Pharm. 582 (2020) 119266, https://doi.org/10.1016/j.ijpharm.2020.119266.
[19] C.B. Roces, E.C. Port, N.N. Daskalakis, J.A. Watts, J.W. Aylott, G.W. Halbert, Y. Perrie, Rapid scale-up and production of active-loaded PEGylated liposomes, Int. J. Pharm. 586 (2020) 119566, https://doi.org/10.1016/j.ijpharm.2020.119566.
[20] S. Joshi, M.T. Hussain, C.B. Roces, G. Anderluzzi, E. Kastner, S. Salmaso, D. J. Kirby, Y. Perrie, Microfluidics based manufacture of liposomes simultaneously entrapping hydrophilic and lipophilic drugs, Int. J. Pharm. 514 (2016) 160–168, https://doi.org/10.1016/j.ijpharm.2016.09.027.
[21] N. Forbes, M.T. Hussain, M.L. Briuglia, D.P. Edwards, J.H. ter Horst, N. Szita, Y. Perrie, Rapid and scale-independent microfluidic manufacture of liposomes entrapping protein incorporating in-line purification and at-line size monitoring, Int. J. Pharm. 556 (2019) 68–81, https://doi.org/10.1016/j.ijpharm.2018.11.060.
[22] A. Ortiz-Perez, D. van Tilborg, R. van der Meel, F. Grisoni, L. Albertazzi, Machine learning-guided high throughput nanoparticle design, Dig. Dis. 1 (2024) 1280–1291, https://doi.org/10.1039/D4DD00104D.
[23] R. Rebollo, F. Oyoun, Y. Corvis, M.M. El-Hammadi, B. Saubamea, K. Andrieux, N. Mignet, K. Alhareth, Microfluidic manufacturing of liposomes: development and optimization by Design of Experiment and Machine Learning, ACS Appl. Mater. Interfaces 14 (2022) 39736–39745, https://doi.org/10.1021/acsami.2c06627.
[24] D. Carugo, E. Bottaro, J. Owen, E. Stride, C. Nastruzzi, Liposome production by microfluidics: potential and limiting factors, Sci. Rep. 6 (2016) 25876, https://doi.org/10.1038/srep25876.
[25] M. Tiboni, M. Tiboni, A. Pierro, M. Del Papa, S. Sparaventi, M. Cespi, L. Casettari, Microfluidics for nanomedicines manufacturing: an affordable and low-cost 3D printing approach, Int. J. Pharm. 599 (2021) 120464, https://doi.org/10.1016/j.ijpharm.2021.120464.
[26] M. Tiboni, S. Benedetti, A. Skouras, G. Curzi, D.R. Perinelli, G.F. Palmieri, L. Casettari, 3D-printed microfluidic chip for the preparation of glycyrrhetinic acid-loaded ethanolic liposomes, Int. J. Pharm. 584 (2020) 119436, https://doi.org/10.1016/j.ijpharm.2020.119436.
[27] J.R. Shah, T. Dong, A.T. Phung, T. Reid, C. Larson, A.B. Sanchez, B. Oronsky, S. L. Blair, O. Aisagbonhi, W.C. Trogler, A.C. Kummel, Development of adenovirus containing liposomes produced by extrusion vs. homogenization: a comparison for scale-up purposes, Bioengineering 9 (2022) 620, https://doi.org/10.3390/bioengineering9110620.
[28] M. Sedighi, S. Sieber, F. Rahimi, M.A. Shahbazi, A.H. Rezayan, J. Huwyler, D. Witzigmann, Rapid optimization of liposome characteristics using a combined microfluidics and design-of-experiment approach, drug Deliv, Transl. Res. 9 (2019) 404–413, https://doi.org/10.1007/s13346-018-0587-4.
[29] S. Rezvantalab, S. Mihandoost, M. Rezaiee, Machine learning assisted exploration of the influential parameters on the PLGA nanoparticles, Sci. Rep. 14 (2024) 1114, https://doi.org/10.1038/s41598-023-50876-w.
[30] P. Bannigan, Z. Bao, R.J. Hickman, M. Aldeghi, F. Häse, A. Aspuru-Guzik, C. Allen, Machine learning models to accelerate the design of polymeric long-acting injectables, Nat. Commun. 14 (2023) 35, https://doi.org/10.1038/s41467-022-35343-w.
[31] V. Di Francesco, D.P. Boso, T.L. Moore, B.A. Schrefler, P. Decuzzi, Machine learning instructed microfluidic synthesis of curcumin-loaded liposomes, Biomed. Microdevices 25 (2023) 29, https://doi.org/10.1007/s10544-023-00671-1.
[32] R. Han, Z. Ye, Y. Zhang, Y. Cheng, Y. Zheng, D. Ouyang, Predicting liposome formulations by the integrated machine learning and molecular modeling approaches, Asian J. Pharm. Sci. 18 (2023) 100811, https://doi.org/10.1016/j.ajps.2023.100811.
[33] A. Volkamer, S. Riniker, E. Nittinger, J. Lanini, F. Grisoni, E. Evertsson, R. Rodríguez-Pérez, N. Schneider, Machine learning for small molecule drug discovery in academia and industry: ML for small molecules drug discovery, Artif. Intell. Life Sci. 3 (2023) 100056, https://doi.org/10.1016/j.ailsci.2022.100056.
[34] K. Swanson, G. Liu, D.B. Catacutan, A. Arnold, J. Zou, J.M. Stokes, Generative AI for designing and validating easily synthesizable and structurally novel antibiotics, Nat. Mach. Intell. 6 (2024) 338–353, https://doi.org/10.1038/s42256-024-00809-7.
[35] E. Zakharova, M. Orsi, A. Capecchi, J.L. Reymond, Machine learning guided discovery of non-hemolytic membrane disruptive anticancer peptides, Chem. MedChem 17 (2022) e202200291, https://doi.org/10.1002/cmdc.202200291.
[36] A. Capecchi, X. Cai, H. Personne, T. Köhler, C. van Delden, J.L. Reymond, Machine learning designs non-hemolytic antimicrobial peptides, Chem. Sci. 12 (2021) 9221–9232, https://doi.org/10.1039/d1sc01713f.
[37] H. Öztürk, A. Özgür, P. Schwaller, T. Laino, E. Ozkirimli, Exploring chemical space using natural language processing methodologies for drug discovery, Drug Discov. Today 25 (2020) 689–705, https://doi.org/10.1016/j.drudis.2020.01.020.
[38] P. Schwaller, R. Petraglia, V. Zullo, V.H. Nair, R.A. Haeuselmann, R. Pisoni, C. Bekas, A. Iuliano, T. Laino, Predicting retrosynthetic pathways using transformer-based models and a hyper-graph exploration strategy, Chem. Sci. 11 (2020) 3316–3325, https://doi.org/10.1039/c9sc05704h.
[39] D. Kreutter, P. Schwaller, J.L. Reymond, Predicting enzymatic reactions with a molecular transformer, Chem. Sci. 12 (2021) 8648–8659, https://doi.org/10.1039/d1sc02362d.
[40] P. Schwaller, T. Laino, T. Gaudin, P. Bolgar, C.A. Hunter, C. Bekas, A.A. Lee, Molecular transformer: a model for uncertainty-calibrated chemical reaction prediction, ACS Cent. Sci. 5 (2019) 1572–1583, https://doi.org/10.1021/acscentsci.9b00576.
[41] S. Aleksić, D. Seeliger, J.B. Brown, ADMET predictability at Boehringer Ingelheim: state-of-the-art, and do bigger datasets or algorithms make a difference? Mol. Inform. 41 (2022) 2100113, https://doi.org/10.1002/minf.202100113.
[42] F. Montanari, L. Kuhnke, A. Ter Laak, D.A. Clevert, Modeling physico-chemical ADMET endpoints with multitask graph convolutional networks, Molecules 25 (2020) 44, https://doi.org/10.3390/molecules25010044.
[43] M. Awale, J.L. Reymond, The polypharmacology browser: a web-based multi-fingerprint target prediction tool using ChEMBL bioactivity data, J. Chemother. 9 (2017) 11, https://doi.org/10.1186/s13321-017-0199-x.
[44] I. Aleksić, J. Crossed, D. Signuriš, S. Ibrić, J. Parojčić, An investigation into the usefulness of different empirical modeling techniques for better control of spray-on fluidized bed melt granulation, Int. J. Pharm. 496 (2015) 627–635, https://doi.org/10.1016/j.ijpharm.2015.11.001.
[45] G.A. Landrum, S. Riniker, Combining IC50 or Ki values from different sources is a source of significant noise, J. Chem. Inf. Model. 64 (2024) 1560–1567, https://doi.org/10.1021/acs.jcim.4c00049.
[46] C. Kramer, T. Kalliokoski, P. Gedeck, A. Vulpetti, The experimental uncertainty of heterogeneous public K i data, J. Med. Chem. 55 (2012) 5165–5173, https://doi.org/10.1021/jm300131x.
[47] S. Obeid, M. Madžarević, M. Krkobabić, S. Ibrić, Predicting drug release from diazepam FDM printed tablets using deep learning approach: influence of process parameters and tablet surface/volume ratio, Int. J. Pharm. 601 (2021) 120507, https://doi.org/10.1016/j.ijpharm.2021.120507.
[48] L. Garcia-del Rio, P. Diaz-Rodriguez, G.K. Pedersen, D. Christensen, M. Landin, Sublingual boosting with a novel Mucoadhesive Thermogelling hydrogel following parenteral CAF01 priming as a strategy against chlamydia trachomatis, Adv. Healthc. Mater. 11 (2022) 2102508, https://doi.org/10.1002/adhm.202102508.
[49] R. Martinez-Borrajo, P. Diaz-Rodriguez, M. Landin, Rationalized design to explore the full potential of PLGA microspheres as drug delivery systems, Drug Deliv. 30 (2023) 2219864, https://doi.org/10.1080/10717544.2023.2219864.
[50] P. García-García, E. Briffault, M. Landin, C. Evora, P. Diaz-Rodriguez, A. Delgado, Tailor-made oligonucleotide-loaded lipid-polymer nanosystems designed for bone gene therapy, drug Deliv, Transl. Res. 11 (2021) 598–607, https://doi.org/10.1007/s13346-021-00926-5.
[51] P. Diaz-Rodriguez, C. Mariño, J.A. Vázquez, J.R. Caeiro-Rey, M. Landin, Targeting joint inflammation for osteoarthritis management through stimulus-sensitive hyaluronic acid based intra-articular hydrogels, Mater. Sci. Eng. C 128 (2021) 112254, https://doi.org/10.1016/j.msec.2021.112254.
[52] M. Orsi, S. Loh, W. Cheng, H. Ang, A. Frei, Using machine learning to predict the antibacterial activity of ruthenium complexes**, Angew. Chem. Int. Ed. 63 (2023) e202317901, https://doi.org/10.26434/chemrxiv-2023-72srv-v2.
[53] B. Hoseini, M.R. Jaafari, A. Golabpour, A.A. Momtazi-Borojeni, S. Eslami, Optimizing nanoliposomal formulations: assessing factors affecting entrapment efficiency of curcumin-loaded liposomes using machine learning, Int. J. Pharm. 646 (2023) 123414, https://doi.org/10.1016/j.ijpharm.2023.123414.
[54] B. Hoseini, M.R. Jaafari, A. Golabpour, A.A. Momtazi-Borojeni, M. Karimi, S. Eslami, Application of ensemble machine learning approach to assess the factors affecting size and polydispersity index of liposomal nanoparticles, Sci. Rep. 13 (2023) 18012, https://doi.org/10.1038/s41598-023-43689-4.
[55] J. Jiménez-Luna, F. Grisoni, G. Schneider, Drug discovery with explainable artificial intelligence, Nat. Mach. Intell. 2 (2020) 573–584, https://doi.org/10.1038/s42256-020-00236-4.
[56] S.M. Lundberg, G. Erion, H. Chen, A. DeGrave, J.M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, S.I. Lee, From local explanations to global understanding with explainable AI for trees, Nat. Mach. Intell. 2 (2020) 56–67, https://doi.org/10.1038/s42256-019-0138-9.
[57] E.C. Giakoumatos, L. Gascoigne, B. Gumí-Audenis, Á.G. García, R. Tuinier, I. K. Voets, Impact of poly(ethylene glycol) functionalized lipids on ordering and fluidity of colloid supported lipid bilayers, Soft Matter 18 (2022) 7569–7578, https://doi.org/10.1039/d2sm00806h.
[58] D. Probst, J.L. Reymond, Visualization of very large high-dimensional data sets as minimum spanning trees, J. Chemother. 12 (2020) 13321, https://doi.org/10.1186/s13321-020-0416-x.
[59] J. Chen, D. Cheng, J. Li, Y. Wang, J.X. Guo, Z.P. Chen, B.C. Cai, T. Yang, Influence of lipid composition on the phase transition temperature of liposomes composed of both DPPC and HSPC, Drug Dev. Ind. Pharm. 39 (2013) 197–204, https://doi.org/10.3109/03639045.2012.668912.
[60] M.L. Briuglia, C. Rotella, A. McFarlane, D.A. Lamprou, Influence of cholesterol on liposome stability and on in vitro drug release, drug Deliv, Transl. Res. 5 (2015) 231–242, https://doi.org/10.1007/s13346-015-0220-8.
[61] M. Biscaia-Caleiras, N.A. Fonseca, A.S. Lourenço, J.N. Moreira, S. Simões, Rational formulation and industrial manufacturing of lipid-based complex injectables:

landmarks and trends, J. Control. Release 373 (2024) 617–639, https://doi.org/10.1016/j.jconrel.2024.07.021.
[62] G. Zhang, J. Sun, Lipid in chips: a brief review of liposomes formation by microfluidics, Int. J. Nanomedicine 16 (2021) 7391–7416, https://doi.org/10.2147/IJN.S331639.
[63] M. Maeki, N. Kimura, Y. Okada, K. Shimizu, K. Shibata, Y. Miyazaki, A. Ishida, K. Yonezawa, N. Shimizu, W. Shinoda, M. Tokeshi, Understanding the effects of ethanol on the liposome bilayer structure using microfluidic-based time-resolved small-angle X-ray scattering and molecular dynamics simulations, Nanoscale Adv. 6 (2024) 2166–2176, https://doi.org/10.1039/d3na01073b.
[64] X. Qi, J. Wang, C. Chen, L. Wang, Optimal design of micromixer for preparation of nanoliposomes, Chem. Eng. Process. (2024) 109677, https://doi.org/10.1016/j.cep.2024.109677.
[65] J. Kotouček, F. Hubatka, J. Mašek, P. Kulich, K. Velínská, J. Bezděková, M. Fojtíková, E. Bartheldyová, A. Tomečková, J. Stráská, D. Hrebík, S. Macaulay, I. Kratochvílová, M. Raška, J. Turánek, Preparation of nanoliposomes by microfluidic mixing in herring-bone channel and the role of membrane fluidity in liposomes formation, Sci. Rep. 10 (2020) 5595, https://doi.org/10.1038/s41598-020-62500-2.
[66] K. Sule, E.J. Prenner, Lipid headgroup and side chain architecture determine manganese-induced dose dependent membrane rigidification and liposome size increase, Eur. Biophys. J. 51 (2022) 205–223, https://doi.org/10.1007/s00249-022-01589-x.
[67] C. Kern, T. Klausch, F. Kreuter, Tree-based machine learning methods for survey research, Surv res, Methods 13 (2019) 73–93, https://doi.org/10.18148/srm/2019.v1i1.7395.
[68] M. Gill, R. Anderson, H. Hu, M. Bennamoun, J. Petereit, B. Valliyodan, H. T. Nguyen, J. Batley, P.E. Bayer, D. Edwards, Machine learning models outperform deep learning models, provide interpretation and facilitate feature selection for soybean trait prediction, BMC Plant Biol. 22 (2022) 180, https://doi.org/10.1186/s12870-022-03559-z.
[69] A. Jahn, S.M. Stavis, J.S. Hong, W.N. Vreeland, D.L. Devoe, M. Gaitan, Microfluidic mixing and the formation of nanoscale lipid vesicles, ACS Nano 4 (2010) 2077–2087, https://doi.org/10.1021/nn901676x.
[70] E. Ledell, M. Petersen, M. Van Der Laan, Computationally efficient confidence intervals for cross-validated area under the ROC curve estimates, Electron. J. Stat. 9 (2015) 1583–1607, https://doi.org/10.1214/15-EJS1035.
[71] M.O. Jara, M. Landin, J.O. Morales, Screening of critical variables in fabricating polycaprolactone nanoparticles using neuro fuzzy logic, Int. J. Pharm. 601 (2021) 120558, https://doi.org/10.1016/j.ijpharm.2021.120558.
[72] M.O. Jara, J. Catalan-Figueroa, M. Landin, J.O. Morales, Finding key nanoprecipitation variables for achieving uniform polymeric nanoparticles using neurofuzzy logic technology, drug Deliv, Transl. Res. 8 (2018) 1797–1806, https://doi.org/10.1007/s13346-017-0446-8.
[73] I. Roshanski, M. Kalech, L. Rokach, Automatic feature engineering for learning compact decision trees, Expert Syst. Appl. 229 (2023) 120470, https://doi.org/10.1016/j.eswa.2023.120470.
[74] S. Lundberg, S.-I. Lee, A Unified Approach to Interpreting Model Predictions, Long Beach, https://github.com/slundberg/shap, 2017.
[75] Food and Drug Administration, Liposome Drug Products: Chemistry, Manufacturing, and Controls, Human Pharmacokinetics and Bioavailability, and Labeling Documentation; Guidance for Industry, Silver Spring. http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm, 2018.
[76] S. Choi, B. Kang, E. Yang, K. Kim, M.K. Kwak, P.S. Chang, H.S. Jung, Precise control of liposome size using characteristic time depends on solvent type and membrane properties, Sci. Rep. 13 (2023), https://doi.org/10.1038/s41598-023-31895-z.
[77] O. Garbuzenko, Y. Barenholz, A. Priev, Effect of grafted PEG on liposome size and on compressibility and packing of lipid bilayer, Chem. Phys. Lipids 135 (2005) 117–129, https://doi.org/10.1016/j.chemphyslip.2005.02.003.
[78] R. Xu, M.A. Tomeh, S. Ye, P. Zhang, S. Lv, R. You, N. Wang, X. Zhao, Novel microfluidic swirl mixers for scalable formulation of curcumin loaded liposomes for cancer therapy, Int. J. Pharm. 622 (2022) 121857, https://doi.org/10.1016/j.ijpharm.2022.121857.
[79] S. Khadke, C.B. Roces, R. Donaghey, V. Giacobbo, Y. Su, Y. Perrie, Scalable solvent-free production of liposomes, J. Pharm. Pharmacol. 72 (2020) 1328–1340, https://doi.org/10.1111/jphp.13329.
[80] M. Zhao, H. Wei, Y. Mao, C. Zhang, T. Liu, W. Liao, Predictions of additive manufacturing process parameters and molten Pool dimensions with a physics-informed deep learning model, Engineering 23 (2023) 181–195, https://doi.org/10.1016/j.eng.2022.09.015.
[81] H. Rouco, P. Diaz-Rodriguez, S. Rama-Molinos, C. Remuñán-López, M. Landin, Delimiting the knowledge space and the design space of nanostructured lipid carriers through artificial intelligence tools, Int. J. Pharm. 553 (2018) 522–530, https://doi.org/10.1016/j.ijpharm.2018.10.058.
[82] F. Weber, L. Rahnfeld, P. Luciani, Analytical profiling and stability evaluation of liposomal drug delivery systems: a rapid UHPLC-CAD-based approach for phospholipids in research and quality control, Talanta 220 (2020) 121320, https://doi.org/10.1016/j.talanta.2020.121320.
[83] S. Peretz Damari, D. Shamrakov, M. Varenik, E. Koren, E. Nativ-Roth, Y. Barenholz, O. Regev, Practical aspects in size and morphology characterization of drug-loaded nano-liposomes, Int. J. Pharm. 547 (2018) 648–655, https://doi.org/10.1016/j.ijpharm.2018.06.037.