RSC Pharmaceutics

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: Y. Xiao, C. Chen, Y. Zhang, D. Tsaoulidis and T. Chen, *RSC Pharm.*, 2026, DOI: 10.1039/D6PM00035E.




ROYAL SOCIETY OF CHEMISTRY

DOI: 10.1039/D6PM00035E

# A Modelling-based comparison of IVRT and synthetic-membrane permeation for early formulation screening

*Yongrui Xiao[a], Chunlin Chen[a], Yu Zhang[a], Dimitrios Tsaoulidis[a], Tao Chen[a]**

[a]*School of Chemistry and Chemical Engineering, University of Surrey, Guildford GU2 7XH, UK*



**A B S T R A C T**

The development of topical drug formulations typically requires reliable assessment of both drug release and permeation characteristics. In vitro release testing (IVRT) is routinely used for quality control and early screening, while in vitro permeation testing (IVPT) has been commonly used as the standardised approach for evaluating skin permeation performance. However, the extent to which IVRT corresponds to IVPT outcomes remains insufficiently quantified when both experiments are conducted under the same conditions. Here, using fifteen ibuprofen formulations previously characterised by IVPT on a Strat-M membrane, we performed IVRT with a nylon membrane and examined the IVRT–IVPT relationship from correlation, predictability and structural-consistency perspectives. IVRT release rate exhibited a strong positive association with IVPT steady-state flux (Pearson $r = 0.95$, $R^2 = 0.92$, $p < 0.001$), and largely preserved formulation ranking (five of the top six IVRT formulations were also top-ranked in IVPT). Gaussian process regression further revealed a highly aligned similarity structure between IVRT- and IVPT-based models (kernel alignment = 0.97). This suggests that the two experimental models capture closely related representations of the formulation space, even though their response variables differ. Combined with repeatability and discriminatory capability of IVRT, these results support IVRT as an initial screening tool prior to conducting more resource-intensive permeation evaluations.

***Keywords:*** *IVRT, IVPT, ibuprofen, topical formulation, drug release, formulation optimisation, permeation testing*

** Corresponding author. E-mail address: t.chen@surrey.ac.uk (T. Chen)*

View Article Online
DOI: 10.1039/D6PM00035E



## 1. Introduction

Topical formulations are designed to deliver active ingredients to the surface or deeper layers of the skin, depending on the intended site of action [1], and have gained increasing attention due to their ability to provide localised therapeutic effects and reduce systemic side effects associated with oral administration [2]. Topical delivery can also enhance patient compliance and improve therapeutic outcomes by bypassing the gastrointestinal tract [3].

The development of topical formulations usually requires considering a range of factors, including drug release characteristics, skin permeation behaviour, and the interaction between excipients and active ingredients [4]. In vitro release test (IVRT) and in vitro permeation test (IVPT) have been commonly applied to evaluate formulation performance, enabling the selection and optimisation of candidate formulations through iterative testing. IVRT is primarily used to characterise drug release kinetics and ensure batch-to-batch consistency. In addition, it can efficiently differentiate between formulations based on variations in excipients and drug concentration [5]. IVPT, by contrast, focuses on simulating skin permeation and is widely applied in bioequivalence studies and regulatory submissions [6, 7]. It also serves as a complementary method to IVRT, and offers deeper insight into the in vivo relevance of formulation performance [8]. Key parameters obtained from IVPT, such as steady-state flux and lag time, are widely used as reference indicators during formulation development, as they reflect the efficiency and rate at which a drug penetrates the skin barrier [9].

However, IVPT is labour-intensive, time-consuming, and constrained by the limited availability of human or animal skin (which also raises ethical concerns) [10]. Even when biomimetic synthetic membranes such as Strat-M membranes are used in place of ex vivo tissue, IVPT remains substantially more time- and resource-intensive than IVRT [11, 12]. Strat-M membrane reproduces the diffusion-limiting properties of human stratum corneum through a multi-layer polyethersulfone/polyolefin matrix impregnated with synthetic lipids [13], with permeability much closer to skin than to standard hydrophilic IVRT filters [14]. Permeation studies through Strat-M membrane therefore often run over 24-48 h to reach steady-state flux [11, 12], and the membrane itself costs roughly an order of magnitude more per disc than standard IVRT filters. These differences in time and cost are the principal challenges of IVPT.

View Article Online
DOI: 10.1039/D6PM00035E



By comparison, IVRT employs inert synthetic membranes with lower barrier resistance, enabling faster permeation detection and shorter test durations [15]. The European Medicines Agency (EMA) advocates the use of inert, non-rate-limiting synthetic membranes for IVRT when evaluating product quality and equivalence [16]. Among the membranes meeting these criteria, nylon membrane has been widely adopted in both industrial and academic settings due to its chemical stability and mechanical strength [17, 18]. Comparative diffusion studies further report release profiles indistinguishable from those obtained with alternative hydrophilic membranes [14] and recovery values within the regulatory acceptance window [19], and nylon membranes have continued to be used in recent IVRT studies [20].

Establishing a reliable correlation between IVRT and IVPT results would allow researchers to use IVRT data to inform early formulation decisions, thereby reducing reliance on IVPT and enhancing development efficiency. Among the key distinctions between IVRT and IVPT is the nature of the barrier employed; regulatory guidance (e.g. FDA) emphasises that IVRT uses an inert, non-rate-limiting synthetic membrane, whereas IVPT employs biological skin where feasible [21]. Previous studies have explored correlations between synthetic membranes and biological skin to assess their suitability as surrogates in permeation tests, including PAMPA, Strat-M and cellulose acetate membranes [13, 22, 23]. While these studies demonstrate the potential of synthetic membranes as skin surrogates, they primarily focus on barrier properties and empirical comparisons.

This study establishes IVRT as a robust and model-consistent surrogate for synthetic-membrane IVPT performance in ibuprofen poloxamer-based formulations. Using fifteen formulations previously characterised by Strat-M IVPT and matched IVRT measurements on nylon membranes, we demonstrate both a strong correlation between IVRT release rate and IVPT steady-state flux, and a high degree of structural similarity between their formulation response surfaces through Gaussian process kernel alignment and trend diagnostics. These results provide a practical basis for prioritising formulations using IVRT before undertaking more resource-intensive permeation studies.

## 2. Materials and methods

View Article Online
DOI: 10.1039/D6PM00035E





### *2.1 Materials*

Ibuprofen and medium chain triglycerides (MCT) were purchased from Fagron (the Netherlands); ethanol, propylene glycol (PG) and poloxamer 407 (P407) were purchased from Sigma Aldrich (US). Phosphate-buffered saline (PBS) was purchased from Sigma-Aldrich (Germany). Nylon membranes were purchased from Millipore (USA). All aqueous solutions were prepared with ultrapure water (ELGA Maxima, High Wycombe, UK). All chemicals and reagents used were of analytical or high-performance liquid chromatography (HPLC) grade and were used as received, without further purification.

### *2.2 Preparation of ibuprofen-loaded poloxamer 407-based formulations*

The ibuprofen-loaded poloxamer 407-based formulations were prepared following the method described by previous work [24]. Briefly, the formulations were prepared by thoroughly blending poloxamer 407, PG, ethanol, and ultrapure water in a mortar. The mixture was subsequently allowed to equilibrate for at least one hour with intermittent stirring, after which MCT was gradually incorporated under continuous stirring. To maintain consistent formulation conditions, any ethanol evaporation occurring during the preparation process was promptly replenished. Once prepared, formulations were immediately sealed to minimise ethanol loss, thus ensuring comparable preparation conditions and enabling direct comparative analysis.

The formulations used in this study were identical to those reported in our previous work, where drug content uniformity, pH value, rheological behaviour, and microstructure were thoroughly characterised [24]. These properties were found to be consistent and within acceptable ranges across all formulations.

### *2.3 In vitro release tests (IVRT)*

IVRT was carried out using a Phoenix DB-6 Diffusion apparatus (Teledyne Hanson, Chatsworth, CA, USA) fitted with a 25 mm Nylon membrane. A 15 mm diffusion diameter was used, providing an effective permeation area of 1.76 $cm^2$. After placing a magnetic stir bar stirring at 500 rpm in the receptor chamber, the final receptor volume was approximately

View Article Online
DOI: 10.1039/D6PM00035E

21.56 ml. Degassed phosphate-buffered saline (PBS, pH 7.4), preheated to 32 °C, served as the receptor medium, in which ibuprofen has a saturation solubility of 1.34 ± 0.01 mg/mL [24]. The apparatus was maintained at 32 ± 1 °C, with a 30-minute equilibration period prior to sample application.

Following equilibration, approximately 0.30 g of the test formulation was uniformly applied to the membrane surface in the donor chamber, which was then sealed with a glass cap to minimise evaporation. The experiment was conducted over 6 hours, with 1.0 ml samples withdrawn from the receptor medium at 1, 2, 3, 4, 5 and 6 hours, each immediately replaced with an equal volume of fresh, preheated PBS. Samples were then analysed by ultraviolet spectrophotometry to determine drug concentrations.

### *2.4 UV-spectral analysis*

The spectrophotometric analysis was carried out on a UV–Vis spectrophotometer (Evolution 201, Thermo Fisher Scientific, Waltham, MA, USA) using 10 mm path-length quartz cells to quantify the concentration of ibuprofen. PBS was used as the blank for baseline correction.

A multi-wavelength UV spectrophotometric method was adopted for the determination of ibuprofen concentration. Given that ibuprofen exhibits absorbance at multiple wavelengths, a full UV scan (200–400 nm) was performed using a 500 µg/ml standard solution to identify the principal absorption peaks, which were found at 203, 209, 213, 219, 226, 229, 264 and 272 nm [25-27]. Standard solutions from 0.5 to 500 µg/ml were prepared by dissolving ibuprofen in PBS, and wavelength-specific calibration curves were constructed to assess linearity, sensitivity and potential interference.

Since sample concentration in this study ranged from below 10 µg/ml to above 200 µg/ml, absorbance readings could fall outside the recommended optimal linear range (approximately 0.2-0.8 absorbance units). To ensure accurate quantification across all concentration levels, measurements were conducted over a broad wavelength range (203–272 nm). For each sample, concentration was obtained from the corresponding wavelength-specific calibration [28]. When multiple wavelengths met the criterion, concentrations from each were required to agree within ±5%; otherwise, the sample was re-measured (or diluted) and the mean value

View Article Online
DOI: 10.1039/D6PM00035E

reported.

Although HPLC offers high sensitivity and selectivity for multi-component and trace-level analyses [24, 29, 30], its use is often accompanied by greater operational complexity, longer analysis times, and higher cost. In this single-analyte diffusion cell context, matrix effects were assessed using receptor blanks and placebo formulations, and no significant interference was observed within the analytical window. UV spectrophotometry was therefore selected as a practical and sufficiently sensitive method for rapid, routine quantification in this study.

### *2.5 Calculation of release rates*

To calculate the drug release rates, the cumulative amount released into the receptor per unit area at each sampling time point $Q_n$(μg/cm$^2$) was computed from the measured concentration $C_n$(μg/ml) while explicitly accounting for the reduction in drug content resulting from successive sampling [31]. The calculation incorporated both the amount present in the receptor at the current sampling point and the drug amount removed by previous samples. Specifically, the cumulative amount $Q_n$ released per unit area at sampling time $t_n$ was calculated as:

$$Q_n = C_n \frac{V_c}{A_c} + \frac{V_s}{A_c} \sum_{i=1}^{n-1} C_i$$

where $Q_n$ is the cumulative released amount per unit area at sampling time $t_n$, $C_n$ is the measured drug concentration in the receptor at time $t_n$, $V_c$ is the total receptor volume, $V_s$ is the volume withdrawn at each sampling, and $A_c$ is the diffusion area of the cell. The summation term $\sum_{i=1}^{n-1} C_i$ explicitly accounts for the cumulative reduction in drug content caused by sampling at previous time points. Unlike the Higuchi model which plots cumulative release versus the square root of time, the release rate in this study was determined by plotting $Q_n$ directly against time and estimating the slope via ordinary least squares (OLS) linear regression to maintain consistency with the IVPT analysis [32].

View Article Online
DOI: 10.1039/D6PM00035E



### *2.6 Correlation analysis*

To investigate whether IVRT can serve as a preliminary screening tool for formulation optimisation, we quantified the association between the drug release rates obtained from the current IVRT experiments and the steady-state flux previously obtained via IVPT testing on Strat-M membranes [24]. Briefly, the previous IVPT used the same vertical diffusion-cell approach as the present IVRT but with Strat-M synthetic membranes over 30 h, with HPLC quantification at 11 sampling points (1, 2, 3, 4, 6, 8, 22, 24, 26, 28, 30 h). Correlation and regression analyses were employed to evaluate the degree of association between the two datasets [33].

Correlation analysis aims to identify potential relationships between two variables and quantify their strength using correlation coefficients. In this study, we focused on the Pearson correlation coefficient, which is commonly used to assess the linear relationship between two continuous variables that are approximately normally distributed. If $X$ and $Y$ represent the two variables under investigation, the Pearson correlation coefficient $r$ is calculated as follows:

$$r = \frac{\sum(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum(X_i - \overline{X})^2 \sum(Y_i - \overline{Y})^2}} (2)$$

The resulting value of $r$ ranges between –1 and +1, with the magnitude reflecting the strength of association. Specifically, values of $r \leq 0.4$ are generally considered to indicate weak correlation, values between $0.4 < r < 0.8$ indicate moderate correlation, and values $r \geq 0.8$ suggest a strong correlation.

### *2.7 Gaussian process regression analysis*

Gaussian process regression (GPR) is a non-parametric Bayesian regression method that places a Gaussian-process prior over the unknown function, with the covariance between any two function evaluations specified by a kernel $k(x, x')$ [34]. Given training data $D = \{(x_i, y_i)\}_{i=1}^{n}$, the posterior predictive distribution at a test input $x_*$ is Gaussian, $f(x_*)|D \sim \mathcal{N}(\mu(x_*), \sigma^2(x_*))$ with

$$\mu(x_*) = k_*^\top (K + \sigma_n^2 I)^{-1} y (3)$$

$$\sigma^2(x_*) = k(x_*, x_*) - k_*^\top (K + \sigma_n^2 I)^{-1} k_* (4)$$

View Article Online
DOI: 10.1039/D6PM00035E



where $K \in \mathbb{R}^{n\times n}$ is the Gram matrix with $K_{ij} = k(x_i,x_j)$, $k_* = [k(x_*,x_1),...,k(x_*,x_n)]^\top$ is the vector of kernel evaluations between $x_*$ and the training inputs; $\sigma_n^2$ is the observation-noise variance; and $I$ is the $n \times n$ identity matrix. We adopted GPR here to characterise the formulation–response relationships for the IVRT release rate and the IVPT steady-state flux beyond the parametric form imposed by polynomial regression, and to enable a kernel-level comparison between the two response surfaces.

Two separate GPR models were trained on the 15-formulation dataset using the GaussianProcessRegressor class of the scikit-learn library in Python: one model used the IVPT steady-state flux as the output, the other used the IVRT release rate. Both models shared the same three-dimensional input $x$ = (P407, ethanol, PG, in %w/w). Inputs and outputs were standardised to zero mean and unit variance prior to training. The covariance function was a constant-scaled squared exponential (radial basis function, RBF) kernel:

$$k(x,x') = \sigma_f^2 \exp\left(-\frac{\|x-x'\|^2}{2l^2}\right) (5)$$

where $l$ is the isotropic length scale and $\sigma_f^2$ is the signal variance.

Partial dependence plots (PDPs) display how a model's predicted response varies with a single input while the remaining inputs are held fixed at their mean values [35]. For each of the three formulation variables, a 50-point grid was constructed across its experimental range, the other two variables were set to their (standardised) means, and the GPR posterior mean $\mu(x_*)$ was evaluated at each grid point. The resulting curves were plotted with an uncertainty band of ±0.5 standard-deviation combining the GPR predictive variance $\sigma^2(x_*)$ and the mean of the experimental replicate-level variances.)

A GPR model encodes pairwise relationships among its training inputs through the fitted kernel (Gram) matrix $K$, so the structural similarity between two GPR models trained on the same input set can be quantified by comparing their kernel matrices. We used the kernel alignment score for this purpose [36]. Let $K_1$ and $K_2$ denote the fitted kernel matrices obtained from the IVRT and IVPT models, respectively. The alignment is defined as:

$$A(K_1,K_2) = \frac{\langle K_1,K_2\rangle_F}{\|K_1\|_F \|K_2\|_F} (6)$$

View Article Online
DOI: 10.1039/D6PM00035E

where $\langle \cdot, \cdot \rangle_F$ denotes the Frobenius inner product, and $\| \cdot \|_F$ is the Frobenius norm. This metric ranges from 0 to 1, with higher values indicating greater alignment of the underlying structures.



### *2.8 Data analysis*

All experimental data are presented as mean ± standard deviation (SD). Data analysis and graphing were performed using OriginPro 2025 software (version 10.2.0.188, Northampton, MA, USA). Differences in release rate between formulations were assessed by one-way ANOVA followed by Tukey's HSD post-hoc test ($\alpha = 0.05$), implemented in Python using the scipy.stats library.

## 3. Results and discussion

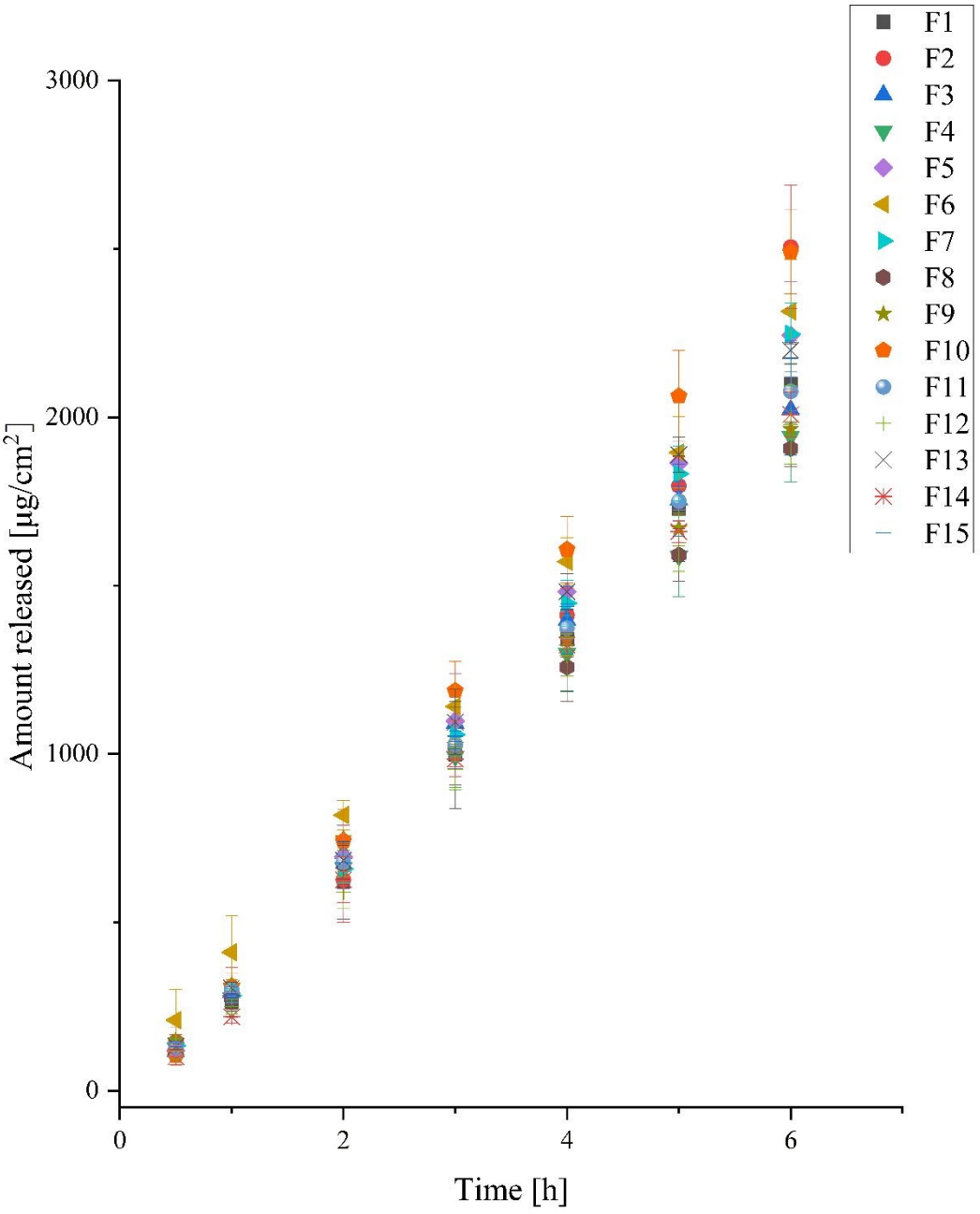
### *3.1 Linearity, precision, reproducibility, and discriminatory analysis*

The in vitro release profiles of ibuprofen formulations were evaluated using nylon membranes over a 6-hour period under standard IVRT conditions. The cumulative amount of drug released per unit area (μg/cm²) was plotted against the time (h) in Figure 1, and a linear relationship was observed for all formulations. The linearity of each profile was confirmed by regression analysis, with an average $R^2$ value of 0.99 across all formulations.

Table 1 presents a summary of the release rate and the cumulative amount released at 6 hours for all 15 formulations. Among the 15 formulations, F10 exhibited the highest mean release rate of 435.25±23.99 μg/cm²/h. In contrast, F8 showed the lowest release rate at 322.96±15.21 μg/cm²/h. The cumulative release at 6 hours was also reported in Table 1. F2 showed the highest total release (2505.70±183.71 μg/cm²), while F8 showed the lowest value (1907.55±54.47 μg/cm²). Consistent release patterns across replicates per formulation (n=2-5) support the reproducibility of the analytical method. These results validate the robustness of the release testing method and its applicability for comparative evaluation of formulation performance.

The discriminatory capability of the IVRT method was further evaluated by one-way ANOVA, which confirmed that release rates differed significantly between the 15

formulations ($F(14, 60) = 12.66$, $p < 0.001$). Tukey HSD post-hoc analysis ranked the formulations into five overlapping statistical groups (compact letter display in Table 1), demonstrating that the IVRT method has sufficient sensitivity to differentiate between excipient compositions. F10 exhibited the highest release rate and was significantly different from all formulations except F2, while F9, F4 and F8 formed the lowest group and were significantly different from the top six formulations (F10, F2, F5, F7, F13, F6); the remaining formulations occupied intermediate positions with overlapping letter codes reflecting the continuum of release rates.



View Article Online
DOI: 10.1039/D6PM00035E

Figure 1. The cumulative release profiles of 15 Ibuprofen formulations

Table 1. IVRT release rate and cumulative released. Compact letter display (Sig. column) indicates statistical groups by Tukey HSD post-hoc test (α = 0.05).

| Formulation | Components (Eth: PG: P407) (%w/w) | Cumulative released at 6h (μg/cm²) | Release rate (μg/cm²/h) | Sig. |
|---|---|---|---|---|
| F1 | 20:15:30 | 2099.99±117.82 | 362.90±20.69 | cde |
| F2 | 15:10:30 | 2505.70±183.71 | 412.19±14.58 | ab |
| F3 | 20:20:25 | 2022.76±134.89 | 351.82±27.33 | cde |
| F4 | 15:15:25 | 1944.10±136.62 | 328.49±22.38 | e |
| F5 | 15:15:25 | 2243.84±160.05 | 389.91±30.22 | bc |
| F6 | 15:15:25 | 2314.58±153.68 | 378.65±36.26 | bcd |
| F7 | 10:20:25 | 2247.38±93.39 | 386.77±17.86 | bcd |
| F8 | 20:10:25 | 1907.55±54.47 | 322.96±15.21 | e |
| F9 | 20:15:20 | 1965.40±22.00 | 329.85±2.91 | e |
| F10 | 10:15:30 | 2491.52±124.45 | 435.25±23.99 | a |
| F11 | 15:10:20 | 2077.19±59.18 | 356.61±10.14 | cde |
| F12 | 10:10:25 | 1977.51±116.96 | 342.55±20.31 | de |
| F13 | 15:20:20 | 2199.60±23.34 | 383.90±6.01 | bcd |
| F14 | 15:20:30 | 2008.20±67.43 | 351.32±11.86 | cde |
| F15 | 10:15:20 | 2175.59±53.42 | 374.52±6.10 | bcd |

### *3.2 Correlation between IVRT and IVPT performance*

As outlined in the methodology section, a correlation analysis was performed to determine whether IVRT could serve as an effective early screening method for identifying formulations with promising transdermal delivery potential. A linear regression was conducted between the IVRT release rates and the IVPT steady-state flux values across all fifteen formulations to assess the feasibility of using IVRT data to inform formulation selection ahead of more resource-intensive IVPT studies.

A strong positive correlation was observed, with a Pearson correlation coefficient of r = 0.95 and a coefficient of determination $R^2$ = 0.92 (Figure 2). This indicates that approximately 92% of the variability in IVPT flux could be explained by the IVRT release rate. The regression model was statistically significant ($p < 0.001$). These findings suggest that formulations exhibiting higher drug release in IVRT tend to correspond to greater transdermal permeation in IVPT, supporting the relevance of IVRT as a screening tool in formulation

View Article Online
DOI: 10.1039/D6PM00035E

development.

The additional ranking results of observed release rates and flux values across all formulations further support the observed IVRT-IVPT correlation. As shown in Table 2, five of the top six formulations by IVRT release rate were also among the top six by IVPT flux. This strong overlap in ranking reinforces that formulations demonstrating superior drug release ability in IVRT tend to exhibit enhanced flux in IVPT. Despite minor discrepancies in individual rankings, the general alignment across both testing modalities substantiates the potential of IVRT as an early-stage screening tool in topical formulation development.

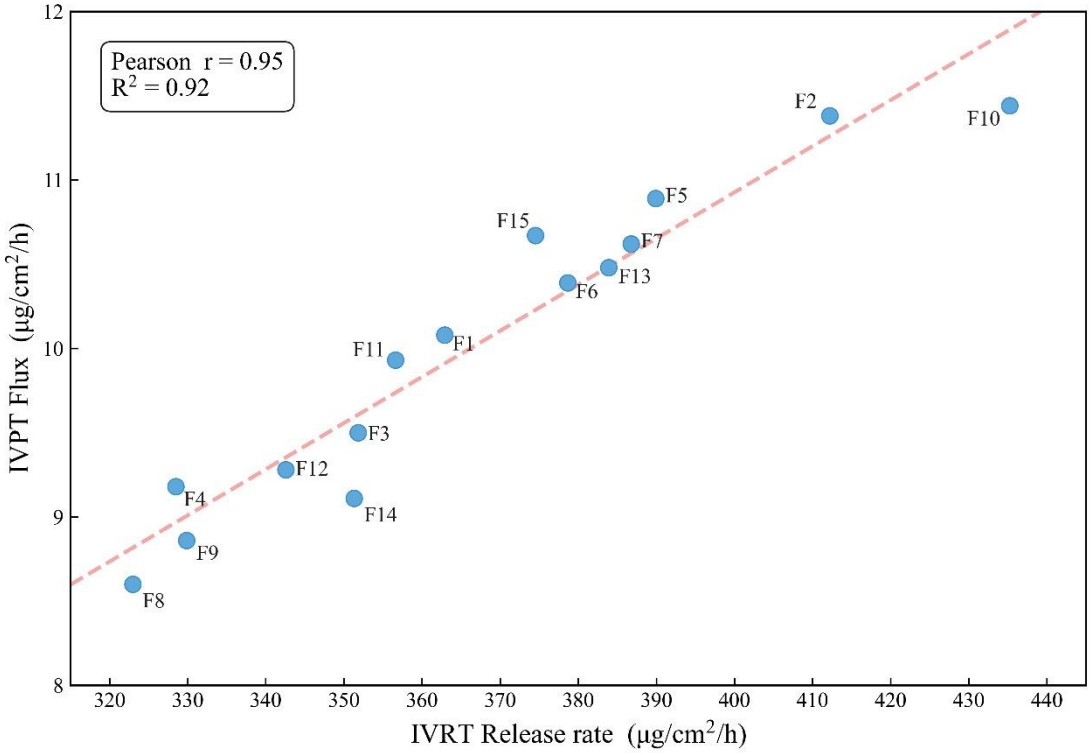



Figure 2. Correlation between IVRT release rates and IVPT flux values across fifteen formulations.

Table 2. Top six formulations by IVRT release rate and their corresponding IVPT flux ranks

| Rank (IVRT) | Formulation | Release rate(μg/cm²/h) | Rank (IVPT) | Flux (μg/cm²/h) |
|---|---|---|---|---|
| 1 | F10 | 435.25±23.99 | 1 | 11.44±0.69 |
| 2 | F2 | 412.19±14.58 | 2 | 11.38±0.71 |
| 3 | F5 | 389.91±30.22 | 3 | 10.89±0.41 |
| 4 | F7 | 386.77±17.86 | 5 | 10.62±1.69 |
| 5 | F13 | 383.90±6.01 | 6 | 10.48±0.53 |
| 6 | F6 | 378.65±36.26 | 7 | 10.39±0.35 |

View Article Online
DOI: 10.1039/D6PM00035E





### *3.3 Comparison of polynomial regression-based optimisation*

To assess the potential predictive capability of IVRT data in identifying the optimal formulation compared with IVPT outcomes, a polynomial regression model was developed based on the release rate and the concentrations of three excipients: P407 ($x_1$), ethanol ($x_2$) and PG ($x_3$).

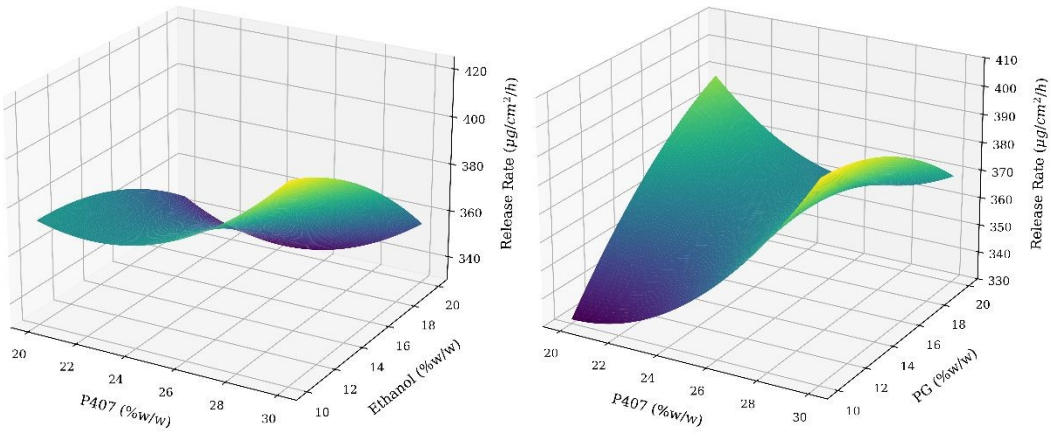
$$y = 177.903 - 14.63x_1 + 13.95\,x_2 + 33.90\,x_3 + 0.69x_1^2 - 0.30x_2^2 - 0.28x_3^2 - 0.28x_1x_2 - 0.88x_1x_3 - 0.15x_2x_3 \quad (7)$$

where $y$ is the release rate. The coefficients reflect the direction and magnitude of each variable's influence on the release rate. P407 ($x_1$) has a relatively high negative linear impact, whereas ethanol ($x_2$) and PG ($x_3$) show positive contributions.

For comparison, the IVPT-based polynomial model was previously established as follows [24]:

$$y = -3.6 + 0.491x_1 + 0.958\,x_2 - 0.0347x_1x_2$$

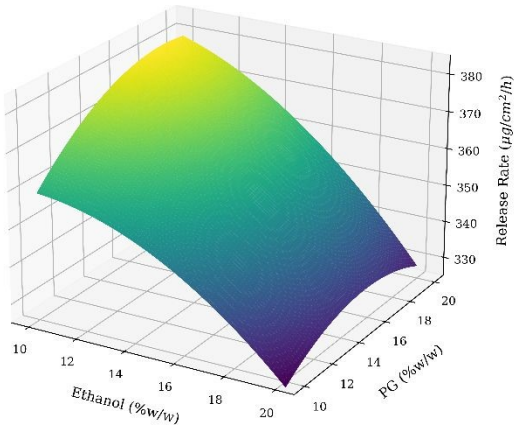
The maximum predicted release rate from the IVRT model was 424.75 μg/cm²/h, obtained with a formulation of 30% P407, 10% ethanol, and 10% PG. Three-dimensional response surface plots were generated to visualise the influence and interaction of pairwise component interaction under fixed third-variable conditions. Figure 3a shows the relationship between P407 and ethanol at a fixed PG level, Figure 3b illustrates P407 versus PG at a constant ethanol level; Figure 3c displays ethanol versus PG with P407 held constant.




a) b)

View Article Online
DOI: 10.1039/D6PM00035E






c)

Figure 3. Response surface plots. a) P407 (%w/w) vs ethanol (%w/w) at fixed PG = 15%; b) P407 (%w/w) vs PG (%w/w) at fixed ethanol = 15%; c) Ethanol (%w/w) vs PG (%w/w) at fixed P407 = 25%.

The IVPT model achieves its maximum at a composition of 20% P407, 20% ethanol, and 10% PG, corresponding to a predicted flux of 11.49 µg/cm²/h. When the optimal IVRT composition (30% P407, 10% ethanol, 10% PG) is applied to the IVPT model, the predicted flux is 10.22 µg/cm²/h, lower than the IVPT-predicted maximum and below the top six observed IVPT formulations in Table 2. The two polynomial models therefore do not locate the same optimum within the explored excipient space. Despite this divergence, the strong empirical correlation reported in section 3.2 (Pearson r = 0.95) together with the substantial overlap in top-six rankings indicate that the two assays rank candidate formulations consistently. The divergent optima also reflect the limited extrapolation reliability of low-order polynomial response surfaces near the boundaries of the design space, motivating the kernel-based analysis in section 3.4.

### *3.4 Gaussian process regression-based analysis*

In our previous study, we reported that Bayesian optimisation based on GPR achieved better results compared to traditional response surface methodology (RSM) [24]. In this section, GPR models were used to evaluate the similarity between the IVRT and IVPT datasets. Two separate GPR models were trained based on two datasets.

To visualise how the IVRT and IVPT models respond to individual formulation variables, we generated partial dependence plots (PDPs) for the three input features: Ethanol, PG, and

P407. These plots depict the predicted standardised response of each GPR model as each individual variable changes, while holding the others at their mean values. Shaded regions around each prediction curve represent the predictive uncertainty, illustrated as ± 0.5 standard deviations.

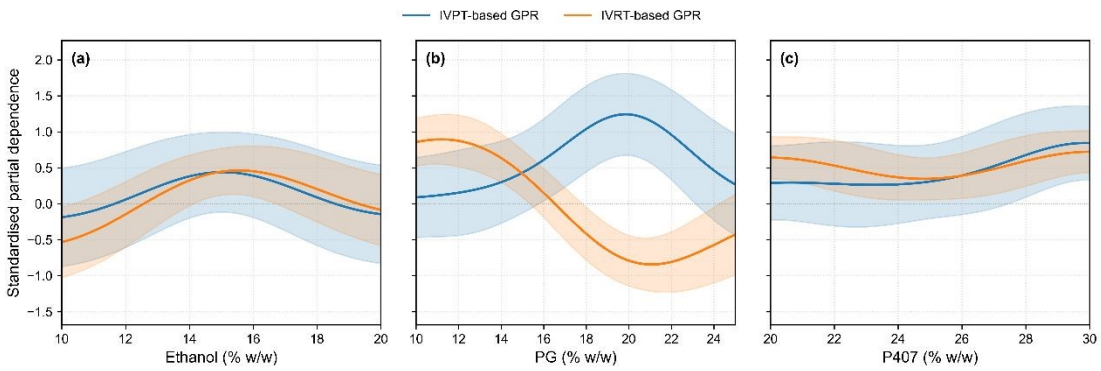



Figure 4. Partial dependence comparison of IVPT-based and IVRT-based GPR models for (a) Ethanol, (b) PG, and (c) P407, with shaded ±0.5 SD predictive bands.

As shown in Figure 4, both models exhibited consistent response patterns for ethanol and P407. Ethanol displayed relatively mild fluctuations in both models, with peaks near 15 %w/w. PG showed peak responses at different concentrations in the two models, with the IVRT-based GPR peaking near PG ≈ 11% w/w and the IVPT-based GPR peaking near PG ≈ 20 %w/w. P407 responses in both models stayed relatively flat at lower concentrations and rose toward 30 %w/w. These coherent trends suggest that, despite having different target outputs, the two systems display similar marginal responses for ethanol and P407. Although the PG marginal responses diverge, partial dependence captures only one-variable-at-a-time effects. To assess overall structural agreement between the two assays, the kernel alignment analysis below evaluates similarity at the level of pairwise sample relationships.

GPR captures the similarity between input samples through their kernel functions, which define a covariance structure over the formulation space. When trained on the same set of input formulations, the resulting kernel matrices provide insight into how each model internally encodes pairwise relationships among samples.

To quantify the degree of structural similarity between the GPR models for IVRT and IVPT,

we computed the kernel alignment score between their respective fitted kernel matrices. In our case, the alignment score between the two models was 0.97, indicating a very high degree of structural agreement. Despite being trained on different outputs, the two models learn closely aligned notions of similarity across the formulation space. This strong alignment supports the hypothesis that the IVRT and IVPT responses share a common latent structure, and that the underlying formulation–response relationship is captured similarly by both models.

### *3.5 Discussion*

This study provides quantitative evidence supporting the utility of IVRT as a surrogate for IVPT in transdermal drug formulation evaluation. Linear regression analysis revealed a strong correlation between IVRT release rates and IVPT flux values (r = 0.95), with 92% of the variation in flux explained by release rate. Rank analysis further confirmed this association: five of the top six formulations ranked by IVRT were also in the top six for IVPT, with the remainder still among the highest-performing formulations.

Polynomial regression and GPR models were then used to characterise the formulation–response relationships. The polynomial models captured local response trends within the observed dataset but located their optima in different regions of the design space, reflecting the limited extrapolation reliability of low-order response surfaces. GPR addressed this limitation by providing improved predictive accuracy together with predictive uncertainty, consistent with prior findings supporting GPR-based Bayesian optimisation strategies [24].

Partial dependence analysis showed that both IVRT and IVPT models exhibited similar trends with respect to ethanol and P407 concentrations, with peak responses at different PG concentrations. Kernel alignment analysis quantitatively confirmed that the two GPR models learned structurally consistent response surfaces, further supporting the existence of a shared similarity structure between IVRT and IVPT.

These results support IVRT as a practical screening tool. The combination of strong empirical correlation, rank consistency, aligned response trends, and model structural similarity indicates that IVRT data correlate closely with IVPT outcomes, thereby reducing the

View Article Online
DOI: 10.1039/D6PM00035E





experimental burden in formulation development.

Nevertheless, several limitations should be acknowledged. The present work focuses on a single active ingredient (ibuprofen) and a specific poloxamer 407-based gel system, and the number of formulations is relatively limited. Moreover, the IVPT data were generated using Strat-M synthetic membranes rather than ex vivo human or animal skin. While Strat-M has been proposed as a useful skin surrogate, validation on ex vivo biological skin will be required before extending these conclusions to broader topical formulation development contexts. Because the nylon membrane used for IVRT is non-rate-limiting whereas Strat-M is rate-limiting, both the release rate and the flux of the present formulations (which differ only in vehicle composition) are governed largely by drug release from the vehicle, and part of the observed correlation reflects this shared dependence. The correlation would therefore be expected to weaken for formulations that differ in components acting on the barrier rather than on release, such as penetration enhancers. The IVRT–IVPT correlation reported here is established within a single excipient family (ibuprofen in poloxamer 407-based gels with ethanol and PG) and a relatively narrow compositional range (P407 20–30% w/w, ethanol 10–20% w/w, PG 10–20% w/w); extrapolation to formulations with substantially different excipient classes or wider compositional ranges would require independent validation. Additional studies on different actives, formulation types and barrier models are also needed to confirm the generality of the observed IVRT–IVPT relationships.

## 4. Conclusion

This study aimed to quantify whether IVRT can prioritise candidate formulations ahead of more resource-intensive IVPT. IVRT and IVPT were systematically compared using correlation metrics, rank consistency, and predictive modelling. The results consistently support a strong relationship between the two experimental methods. Polynomial regression captured the local response trends, while GPR-based partial dependence analysis and kernel alignment confirmed structural consistency between the two assays. Together, these analyses reinforce the value of IVRT for early-stage formulation screening and prioritisation. IVRT thus emerges as a valuable tool for early formulation screening. Its consistency with IVPT across multiple analytical layers supports its use for prioritising formulations prior to IVPT

View Article Online
DOI: 10.1039/D6PM00035E

confirmation. This approach can reduce experimental effort by limiting the number of resource-intensive IVPT studies required, thereby decreasing time and labour demands and lowering overall development costs in transdermal drug development.



## Author contributions

Yongrui Xiao and Chunlin Chen carried out the investigation, analysis, methodology development and data visualisation. Tao Chen was responsible for conceptualisation of the research, project administration and supervision. Yongrui Xiao wrote the original draft of the manuscript and Yongrui Xiao, Chunlin Chen, Yu Zhang, Dimitrios Tsaoulidis and Tao Chen were responsible for review and editing.

## Acknowledgements

Yongrui Xiao was funded by a PhD studentship from the China Scholarship Council (No. 202206050047).

During the preparation of this work the authors used ChatGPT in order to polish text by correcting grammatical mistakes and enhancing the readability of the text. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the published article.



## References

[1] F. Casanova and L. Santos, "Encapsulation of cosmetic active ingredients for topical application–a review," *Journal of microencapsulation,* vol. 33, no. 1, pp. 1-17, 2016.

[2] B. R. Da Costa *et al.*, "Effectiveness and safety of non-steroidal anti-inflammatory drugs and opioid treatment for knee and hip osteoarthritis: network meta-analysis," *bmj,* vol. 375, 2021.

[3] B. R. Jasti, W. Abraham, and T. K. Ghosh, "Transdermal and Topical drug delivery systems," in *Theory and practice of contemporary pharmaceutics*: CRC Press, 2021, pp. 423-454.

[4] M. B. Brown, R. Turner, and S. T. Lim, "Topical product formulation development," *Transdermal and topical drug delivery: principles and practice. Hoboken: Wiley,* pp. 255-84, 2012.

[5] V. P. Shah, D. S. Miron, F. Ș. Rădulescu, J.-M. Cardot, and H. I. Maibach, "In vitro release test (IVRT): Principles and applications," *International Journal of Pharmaceutics,* vol. 626, p.

View Article Online
DOI: 10.1039/D6PM00035E



122159, 2022.

[6] M. E. Lane, "In vitro permeation testing for the evaluation of drug delivery to the skin," *European Journal of Pharmaceutical Sciences,* vol. 201, p. 106873, 2024.

[7] R. S. Nair, N. Billa, and A. P. Morris, "Optimizing In Vitro Skin Permeation Studies to Obtain Meaningful Data in Topical and Transdermal Drug Delivery," *AAPS PharmSciTech,* vol. 26, no. 5, pp. 1-12, 2025.

[8] M. I. S. C. d. M. Rodrigues, "Topical bioequivalence: experimental and regulatory considerations," 00500:: Universidade de Coimbra, 2022.

[9] L. L. Santos, N. J. Swofford, and B. G. Santiago, "In vitro permeation test (IVPT) for pharmacokinetic assessment of topical dermatological formulations," *Current Protocols in Pharmacology,* vol. 91, no. 1, p. e79, 2020.

[10] B. Finnin, K. A. Walters, and T. J. Franz, "In vitro skin permeation methodology," in *Transdermal and topical drug delivery: principles and practice*: John Wiley & Sons, 2012, pp. 85-108.

[11] J. Klebeko *et al.*, "Permeability of ibuprofen in the form of free acid and salts of l-valine alkyl esters from a hydrogel formulation through strat-m™ membrane and human skin," *Materials,* vol. 14, no. 21, p. 6678, 2021.

[12] B. Milanowski *et al.*, "Optimization and evaluation of the in vitro permeation parameters of topical products with non-steroidal anti-inflammatory drugs through Strat-M® membrane," *Pharmaceutics,* vol. 13, no. 8, p. 1305, 2021.

[13] A. Haq, M. Dorrani, B. Goodyear, V. Joshi, and B. Michniak-Kohn, "Membrane properties for permeability testing: Skin versus synthetic membranes," *International journal of pharmaceutics,* vol. 539, no. 1-2, pp. 58-64, 2018.

[14] S. Nallagundla, S. Patnala, and I. Kanfer, "Comparison of in vitro release rates of acyclovir from cream formulations using vertical diffusion cells," *AAPS PharmSciTech,* vol. 15, no. 4, pp. 994-999, 2014.

[15] S. D'Souza, "A review of in vitro drug release test methods for nano-sized dosage forms," *Advances in pharmaceutics,* vol. 2014, no. 1, p. 304757, 2014.

[16] European Medicines Agency, "Draft guideline on quality and equivalence of topical products," ed: European Medicines Agency Amsterdam, The Netherlands, 2018.

[17] M. Shakiba *et al.*, "Nylon—A material introduction and overview for biomedical applications," *Polymers for advanced technologies,* vol. 32, no. 9, pp. 3368-3383, 2021.

[18] R. R. Klein, J. Q. Tao, S. Wilder, K. Burchett, Q. Bui, and K. D. Thakker, "Development of an in vitro release test (IVRT) for a vaginal microbicide gel," *Dissolufion Technol,* vol. 17, no. 4, pp. 6-10, 2010.

[19] S. Rath and I. Kanfer, "A validated IVRT method to assess topical creams containing metronidazole using a novel approach," *Pharmaceutics,* vol. 12, no. 2, p. 119, 2020.

[20] M. d. S. Brighenti *et al.*, "In vitro drug release and ex vivo dermal drug permeation studies of selected commercial benzoyl peroxide topical formulations: correlation between human and porcine skin models," *Molecular Pharmaceutics,* vol. 22, no. 3, pp. 1365-1372, 2025.

[21] (2022). *In Vitro Release Test Studies for Topical Drug Products Submitted in ANDAs Guidance for Industry* [Online] Available: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/in-vitro-release-test-studies-topical-drug-products-submitted-andas

View Article Online
DOI: 10.1039/D6PM00035E



[22] Y. Zhang, M. E. Lane, J. Hadgraft, M. Heinrich, T. Chen, G. Lian, and B. Sinko, "A comparison of the in vitro permeation of niacinamide in mammalian skin and in the Parallel Artificial Membrane Permeation Assay (PAMPA) model," *International Journal of Pharmaceutics,* vol. 556, pp. 142-149, 2019.

[23] B. Sinkó *et al.*, "Use of an in vitro skin parallel artificial membrane assay (Skin-PAMPA) as a screening tool to compare transdermal permeability of model compound 4-phenylethyl-resorcinol dissolved in different solvents," *Pharmaceutics,* vol. 13, no. 11, p. 1758, 2021.

[24] Y. Xiao, T. Ilić, A. Tošić, B. Ivković, D. Tsaoulidis, S. Savić, and T. Chen, "Topical drug formulation for enhanced permeation: A comparison of Bayesian optimisation and response surface methodology with an ibuprofen-loaded poloxamer 407-based formulations case study," *International Journal of Pharmaceutics,* p. 125306, 2025.

[25] B. R. Kesur, V. Salunkhe, and C. Magdum, "Development and validation of UV spectrophotometric method for simultaneous estimation of ibuprofen and famotidine in bulk and formulated tablet dosage form," *Int J Pharm Pharm Sci,* vol. 4, no. 4, pp. 271-274, 2012.

[26] S. Sunaric, M. Petkovic, M. Denic, S. Mitic, and A. Pavlovic, "Determination of ibuprofen in combined dosage forms and cream by direct UV spectrophotometry after solid-phase extraction," *Acta Poloniae Pharm. Drug Res,* vol. 70, no. 3, pp. 401-411, 2013.

[27] M. Khoshayand, H. Abdollahi, M. Shariatpanahi, A. Saadatfard, and A. Mohammadi, "Simultaneous spectrophotometric determination of paracetamol, ibuprofen and caffeine in pharmaceuticals by chemometric methods," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy,* vol. 70, no. 3, pp. 491-499, 2008.

[28] N. Kumari, A. Chaudhary, K. K. Verma, K. Verma, G. Saini, M. Vyas, and B. Singh, "Development and validation of an analytical method for the simultaneous estimation of Etoricoxib and Thiocolchicoside in tablet dosage form by UV Spectrophotometric method," *Research Journal of Pharmacy and Technology,* vol. 15, no. 7, pp. 3051-3056, 2022.

[29] Z. Han, L. Lu, L. Wang, Z. Yan, and X. Wang, "Development and validation of an HPLC method for simultaneous determination of ibuprofen and 17 related compounds," *Chromatographia,* vol. 80, pp. 1353-1360, 2017.

[30] M. R. Payán, M. Á. B. López, R. Fernández-Torres, J. L. P. Bernal, and M. C. Mochón, "HPLC determination of ibuprofen, diclofenac and salicylic acid using hollow fiber-based liquid phase microextraction (HF-LPME)," *Analytica chimica acta,* vol. 653, no. 2, pp. 184-190, 2009.

[31] M. R. Marques, "Performance Tests--Update on USP Activities," *Dissolution Technologies,* vol. 28, no. 1, pp. 40-42, 2021.

[32] A. Boyko, V. Kukartsev, V. Tynchenko, L. Korpacheva, N. Dzhioeva, A. Rozhkova, and S. Aponasenko, "Using linear regression with the least squares method to determine the parameters of the Solow model," in *Journal of Physics: Conference Series*, 2020, vol. 1582, no. 1: IOP Publishing, p. 012016.

[33] R. Shi and S. A. Conrad, "Correlation and regression analysis," *Ann Allergy Asthma Immunol,* vol. 103, no. 4, pp. S34-S41, 2009.

[34] V. L. Deringer, A. P. Bartók, N. Bernstein, D. M. Wilkins, M. Ceriotti, and G. Csányi, "Gaussian process regression for materials and molecules," *Chemical reviews,* vol. 121, no. 16, pp. 10073-10141, 2021.

[35] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of*

View Article Online
DOI: 10.1039/D6PM00035E

*statistics,* pp. 1189-1232, 2001.

[36] N. Cristianini, J. Shawe-Taylor, A. Elisseeff, and J. Kandola, "On kernel-target alignment," *Advances in neural information processing systems,* vol. 14, 2001.



View Article Online
DOI: 10.1039/D6PM00035E

## Data Availability Statement

Data for this article are available at [https://doi.org/10.1016/j.ijpharm.2025.125306].