Saudi Pharmaceutical Journal 29 (2021) 516–526

Contents lists available at ScienceDirect

Saudi Pharmaceutical Journal

journal homepage: www.sciencedirect.com

Original article

Industrial application of QbD and NIR chemometric models in quality improvement of immediate release tablets

Ahmed Ramadan [a,b,*], Emad B. Basalious [a], Mohamed Abdallah [a,c]

[a] Department of Pharmaceutics and Industrial Pharmacy, Faculty of Pharmacy, Cairo University, Egypt
[b] Department of Applied Statistics, Faculty of Postgraduate Studies for Statistical Research, Cairo University, Egypt
[c] Department of Pharmaceutics and Industrial Pharmacy, School of Pharmacy, New Giza University, Egypt

ARTICLE INFO

Article history:
Received 16 January 2021
Accepted 13 April 2021
Available online 22 April 2021

Keywords:
Quality by Design (QbD)
Chemometrics
Near Infrared Spectroscopy (NIRS)
Formulation
Dissolution
Content Uniformity
Artificial Intelligence

ABSTRACT

Quality by Design (QbD) and chemometric models are different sides of the same coin. While QbD models utilize experimentally designed settings for optimization of some quality attributes, these settings can also be utilized for chemometric prediction of the same attributes. We aimed to synchronize optimization of comparative dissolution results of carvedilol immediate release tablets with chemometric prediction of dissolution profile and content uniformity of the product. As an industrial application, selection of variables for optimization was done by performing risk assessment utilizing the archived product records at the pharmaceutical site. Experimental tablets were produced with 20 different settings with the variables being contents of sucrose, sodium starch glycolate, lactose monohydrate, and avicel Ph 101. Contents of the excipients were modelled with F1 dissimilarity factor and F2 similarity factor in HCL, acetate, and USP dissolution media to determine the design space. We initiatively utilized Partial Least Square based Structural Equation Modelling (PLS-SEM) to explore how the excipients and their NIR records explained dissolution of the product. Finally, the optimized formula was utilized with varied content of carvedilol for chemometric prediction of the content uniformity.

1. Introduction

Data-driven solutions have been extensively utilized during the last two decades. Many fields widely applied these solutions, such as pharmaceutical industry, healthcare, food, and agriculture (Granato et al., 2018). Each field keeps looking for a cheap source of data that can be processed with advanced analytics to yield valuable and expensive information. In the field of pharmaceutical industry, the use of non-destructive spectral techniques along with multi-variate calibration models has delivered potential benefits to the industry (Palou et al. 2012).

Pharmaceutical industry utilizes data-driven solutions in both Research& Development (R&D) and routine manufacturing phases (Rantanen and Khinast, 2015). FDA currently supports application of improved development approaches e.g., quality by design (QbD) in R&D phase it also supports use of process analytical technology (PAT) and multi-variate calibration models as tools to support continuous manufacturing (FDA, 2019). Regulatory bodies are developing standards and controls for this conceptual shift in the pharmaceutical development and manufacturing, however, deficits in process monitoring and data collection are still limiting the use of these solutions. Hence, the guidelines require to include risk assessments for verifying how deficits in process monitoring and data collection would affect quality of the products (FDA, 2004).

ICH Quality Implementation Working Group (QIWG) points to consider for ICH Q8/Q9/Q10 guidelines categorized the mathematical models used in pharmaceutical development and manufacturing into three categories:

(a) Low impact models; supporting processes of development and manufacturing (e.g., formulation optimization),

Abbreviations: QbD, Quality by Design; NIR, Near Infra-red; PAT, Process Analytical Technology; PLS, Partial Least Square; PCA, Principal Component Analysis; PC, Principal Component; SEM, Structural Equation Modelling; RMSEP, Root Mean Square Error of Prediction.

* Corresponding author.at: Faculty of Pharmacy, Cairo University, Kasr El-Aini, Cairo 11562, Egypt.
E-mail address: aamadan@pg.cu.edu.eg (A. Ramadan).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.jsps.2021.04.012
1319-0164/

(b) Medium-impact models; contributing to assuring quality of the products but they are not the sole indicators for their quality,
(c) Quality and high impact models when prediction from the model is a significant indicator of the quality of the product.

The document also provides standards for validation of models and verification of prediction accuracy throughout the life cycle of the product. Validation of the models can be performed by several means; however, the impact and intended use of the model determines the most suitable validation method (ICH, 2012).

ICH Q8 R2 defines QbD models as multi-dimensional combination of variables which are related to quality of the product (e.g., formulation attributes and process controls). Design space incorporates a pre-defined confidence interval for every parameter associated with product quality in the design. Within these combinations of confidence intervals, the quality of the product is assured (Mishra et al., 2018; Swain et al., 2019). These models can fully rely on historical data or prior experience. However, prospective studies are often needed especially when high variability in the results of historical data exists (Singh et al., 2017). Harmonizing with QbD models, NIR spectral techniques have been utilized for optimization of various attributes of the pharmaceutical industry (Haneef and Beg, 2021; Taleuzzaman et al., 2021). Some of these attributes are process related e.g., fluidized bed granulation and tablet coating (Liu et al., 2017), drying (Pauli et al., 2018), monitoring of blending (Harting and Kleinebudde, 2019; Nagy et al., 2018; Riolo et al., 2018), other attributes are related to quality control testing of the product e.g., content uniformity testing (Arruabarrena et al., 2014; Nagy et al., 2017), dissolution testing (Galata et al., 2021, 2019; Ojala et al., 2020; Zhao et al., 2019), particle size determination (Bittner et al., 2011; Pauli et al., 2019), and detecting polymorphs and counterfeit drugs (Dégardin et al., 2016; Terra and Poppi, 2014).

One of the most important applications of these techniques is predicting in-vitro dissolution, which became a vital test since confirmed to be correlated with drug in-vivo bioavailability (Dokoumetzidis and Macheras, 2006). FDA recently recommended spectral techniques and mathematical models as a potential alternative to the convenient methods of dissolution testing (FDA, 2019). Researchers have been extensively deploying multi-variate models for prediction of dissolution using various mathematical algorithms such as Principal Component Regression (Otsuka et al., 2007), Partial Least Square (Galata et al., 2019; Zhao et al., 2019), and Artificial Neural Networks (Galata et al., 2019). In parallel, various methods were also utilized for optimization of calibration and modeling, such as mathematical preprocessing of spectral data (Martens and Stark, 1991), and wavelength selection methods (Deng et al., 2015). These methods were useful in increasing predictability of models by removing uninformative or noisy regions of the spectral data. Of the limitations of these models is that NIR spectra can detect both purposeful variables (e.g., process related and formulation related variables) which are intentionally varied for calibration, and it also can detect non-purposeful variables (e.g., water content, particle size distribution and polymorphism) (Zhong et al., 2020). This phenomenon has pros and cons, definition of purposeful and non-purposeful variables depends on identity of the preselected calibrating variables. Hence, researchers are prompted to carefully select calibration variables to get a robust, precise, and reliable calibration models.

Researchers approached different methodologies for prediction of dissolution profile of tablets, for instance, Freitas et al designed formulations with varied excipient contents (Freitas et al., 2005), Galata et al designed formulations with varied compression forces (Galata et al., 2019), while Zhao et al designed formulations with varied combination of formulation and process related variables (Zhao et al., 2019). Generally, there are substantial variables to be considered for predicting reliable results of content uniformity and dissolution testing of solid dosage forms. Some of these variables are process related (compression force, blending time, etc.), and others are formulation related (concentration of excipients, particle size distribution of constituents, etc.). Although NIR spectroscopy is reported to be capable of capturing both process and formulation related variables to a great extent, adapted modeling strategy for each product should be considered (Zhao et al., 2019).

Our current study aimed to improve quality of carvedilol immediate release tablet on industrial scale. Carvedilol belongs to biopharmaceutical classification system (BCS) class II. The drug has low solubility and relatively poor bioavailability of about 25%, hence, variation in formulation and process related parameters would potentially affect the dissolution of the drug (Fernandes et al., 2018). Another challenge was performing chemometric prediction of dissolution profile on immediate release tablets, for which dissolution results may not vary among the experimental formulations at all the tested time points. On the other hand, the product was suitable for prediction of content uniformity with NIR multivariate calibrations, as carvdilol was representing 25% of the total weight of tablet. While previous studies have only utilized experimentally designed formulations for chemometric prediction of some quality attributes, this current study initiatively aimed to utilize a set of experimentally designed formulations to synchronize optimization of comparative dissolution of Carvedilol immediate release tablets with chemometric prediction of dissolution profile and content uniformity. We primarily selected these tests for being time consuming and require high cost to be processed compared to other QC tests. As a real-life application, we conducted a risk assessment to determine predictor variables to be utilized for optimization of comparative dissolution profile. The risk assessment has underlined four excipients as potential variables for optimization. We experimentally produced formulations with varied contents of these excipients and built an QbD model for optimization. We further utilized the NIR records of these formulations for chemometric prediction of the dissolution profile. After determination of the optimized formula, it was utilized with varied content of carvedilol for chemometric prediction of the content uniformity. We initiatively built a Partial Least Square- Structural equation model (PLS-SEM) to underline the significant variables that affect dissolution results of the product.

2. Materials and methods

2.1. Materials

Carvedilol was supplied from Cadila Pharmaceuticals, Ahmedabad, India. Sucrose was supplied from Cristalco, Paris, France. Avicel Ph 101 was supplied from Mingtai chemical Co. LTD, Taiwan. Lactose monohydrate was supplied from Megelle pharmaceuticals, Waserburg, Germany. Sodium Starch glycolate was supplied from Roquette freres, Lestrem, France. Aerosil was supplied from Evonik industries, Essen, Germany. Polyvinylpyrrolidone (PVP) was supplied from Jiaozuo Zhongwei Special Products Pharmaceutical Co., Ltd, China. Croscarmellose was supplied from Blanver FARMOQUIMICA LTDA, Brazil. Magnesium Stearate was supplied from accent microcell Pvt. Ltd, India.

2.2. Methods

2.2.1. Experimental design

Based on a conducted risk assessment included different formulation related variables (risk assessment table is included in supplementary material), contents of four excipients (sucrose,

sodium starch glycolate, lactose monohydrate and avicel Ph 101) were determined as potential predictor variables for optimization of comparative dissolution profile. A mixture design model (Gervasi et al., 2019) was performed using Design Expert software (Version 10.0.0. Stat-Ease), modifying contents of the selected excipients. Sucrose content ranged from 10 to 25 mg, sodium starch glycolate content ranged from 5 to 10 mg, lactose monohydrate content ranged from 20 to 35 mg, and avicel Ph 101 content ranged from 5 to 15 mg. All experimental runs contained fixed contents of carvedilol, PVP, aerosil and Mg stearate (25, 3,4, and 2 mg respectively), maintaining the tablet weight to 100 mg. 20 runs were suggested by the software as presented in Table 1.

The experimental tablets were subjected to NIR measurement and further to comparative dissolution testing with the innovator using reference HPLC method. The calculated F1 and F2 factors for each formula in HCL, acetate and USP medium were added to the QbD model as responses. The recorded NIR spectra along with HPLC dissolution results in USP medium were utilized for building a chemometric model for prediction of dissolution results. Four settings with varied comparative dissolution results were formulated on larger scale (500 tablets) and were utilized for external testing of the chemometric model.

Based on design space of QbD models, a formulation setting with optimum comparative dissolution results was determined. This formulation was further utilized with varied content of Carvedilol to construct NIR chemometric model for prediction of content uniformity. Five different settings were manufactured modifying content of carvedilol to 90%, 95%, 100%, 105% and 110% of the label claim. Three settings (90%, 100%, 110% of the label claim) were formulated on a larger scale (500 tablets) for external validation of the chemometric model.

2.2.1.1. Tablet manufacturing. The experimental tablets consisted of eight components: carvedilol (25 mg, as API), sucrose (10–25 mg, as binding agent for wet granulation), lactose monohydrate (20–35 mg, as diluent), avicel pH 101 (5–15 mg, as diluent), sodium starch glycolate (5–10 mg, as disintegrating agent), PVP K30 (3 mg, as binder), aerosil 200 (4 mg, as glidant), and Mg stearate (2 mg, as glidant). API was blended manually with grinded sucrose, avicel pH 101, PVP k30 & lactose monohydrate. This powder mixture was sieved through 0.5 mm sized sieve (no. 35) and then granulated with water. The wet granules were dried in Memmert oven (Memmert GmbH, Germany) at 45 °C for approximately 3 h. Aerosil 200 was added to the dried granules and the mixture was sieved through 0.5 mm sized sieve. Afterwards, sodium starch glycolate was added, and this blend was then lubricated with Magnesium stearate. Finally, compression was performed using 6 mm round concave punch on Vanguard single tablet press (Vanguard, USA) with 100 mg average tablet weight and hardness of 6–8 kp. For content uniformity testing, the tablets consisted of same components modifying content of the API to 90%, 95%, 100%, 105% and 110% of the label claim.

2.2.1.2. NIR spectroscopy. NIR spectroscopy measurements were performed using Bruker Optics MPA (Multi-Purpose Analyzer) FT-NIR spectrometer (Bruker Optik GmbH, Ettlingen, Germany) with high intensity tungsten NIR source and InGaAs detector equipped with a fiber optic probe for measurement of solids. Spectra were recorded in range of 12000–4000 cm^{-1} wavenumber range setting for 16 cm^{-1} spectral resolution. Each tablet was scanned 16 times per each single measurement. Reflectance measurements were carried out for dissolution testing while absorbance measurements were carried out for content uniformity testing (Shi and Anderson, 2010).

2.2.1.3. NIR for dissolution testing. Three repeated NIR measurements (reflectance mode, Bruker MPA with solid probe) were performed for three tablets of each of the twenty experimental formulations. 180 NIR spectra were recorded and further utilized along with HPLC dissolution results of these experimental tablets to build a chemometric model for prediction of dddissolution results in USP medium. NIR spectra were also recorded for four settings of experimental tablets manufactured on larger scale for external validation of the chemometric model.

2.2.1.3.1. NIR for content uniformity. Five repeated NIR measurements (absorbance mode, Bruker MPA with solid probe) were performed for five tablets of each experimental formulation settings. The recorded 125 NIR spectra were further utilized along with HPLC assay results of these tablets to build a chemometric model for prediction of content uniformity. NIR spectra were also recorded for three settings of experimental tablets manufactured on larger scale for external testing of this chemometric model.

2.2.1.4. In vitro dissolution testing. Dissolution profiles of the experimental tablets were recorded utilizing a Hanson SR8-Plus dissolution tester (Chatsworth, CA, USA) under the following conditions: 900 ml dissolution medium, paddle type, 50 rpm. Samples were taken at 7 time points (5, 10, 15, 20, 30, 45 and 60 min). Compar-

Table 1
Applied experimental settings for optimization. The amounts are presented in mg.

Run	Avicel Ph 101	Lactose Monohydrate	Sucrose	Sodium Starch Glycolate
1	15	31	10	10
2	10.8	25.44	24.75	5
3	15	25.25	17.64	8.12
4	12.16	29.22	14.62	10
5	5	28	25	8
6	11	20	25	10
7	5.52	30.89	20.11	9.48
8	9.11	30.43	21.46	5
9	15	25.25	17.64	8.12
10	10	35	13.26	7.74
11	15	25.25	17.64	8.12
12	15	20	25	6
13	10	35	13.26	7.74
14	10	35	13.26	7.74
15	5	35	21	5
16	5	35	16	10
17	9.61	26.27	20.13	10
18	5	28	25	8
19	15	35	11	5
20	13.54	30.35	17.12	5

ative dissolution testing with the innovator was performed in three dissolution media as follows: USP medium; 0.7% (7 ml/L) HCL adjusted by sodium hydroxide 50% to pH 1.4 (USP, 2018), acetate pH 4.5, and 0.1 N HCL. HPLC analysis for dissolution samples was performed using a Waters Alliance system with UV detector following in-house validated method; the mobile phase was prepared by dissolving potassium phosphate (2.7 g/L) adjusted to pH 2.0, adding 450 ml buffer solution to 550 ml of acetonitrile. An L1 packing 4.6 mm × 25 mm column was utilized for the analysis, at a flow rate of 1 ml/min and injection volume of 20 μl. Absorbance of samples was measured at wavelength of 240 nm.

2.2.1.5. Content uniformity testing. HPLC assay of tablets was performed following the method stated in the monograph of carvedilol tablets in the USP Pharmacopoeia. A sample of powered tablets -equivalent to 25 mg carvedilol- were utilized for the analysis. Potassium phosphate buffer adjusted to pH 3.0 ± 0.1 was utilized as buffer solution, the mobile phase was prepared by dissolving sodium dodecyl sulphate in the buffer solution (1.04 g/150 ml), adding 720 ml of acetonitrile and complete to 2L with water. An L1 packing 4.6 mm × 25 mm column was utilized for the analysis, at a flow rate of 1 ml/min and injection volume of 25 μl. The absorbance of samples was measured at wavelength of 240 nm (USP, 2018).

2.3. Data analysis

QbD models were generated by Design expert 10 software. Optimization of wavelength selection was performed with OPUS 7 Quant software (Bruker optics, Germany). The descriptive statistics and chemometrics were performed using The Unscrambler X (Version 10.5, Camo Analytics, Norway). Structural Equation Modelling (SEM) was performed using SmartPLS 3 software (Version 3, Germany).

3. Results and discussion

3.1. Formula optimization

After dissolution profiles were recorded in HCL, acetate and USP dissolution media, dissimilarity factors F1 and similarity factors F2 were calculated for the experimental tablets versus innovator using reference HPLC method. ANOVA models were utilized to assess contribution of the varied contents of excipients to variance of the calculated similarity and dissimilarity factors. Hence, 6 models were performed, the results are presented in Table 2.

Values of predicted R^2 of USP models were higher than 50%, indicating a reasonable explanation of the calculated F1 and F2 factors with varying contents of the excipients. On the other hand, values of predicted R^2 of HCL and acetate models were lower than 50%, indicating that the change in contents of the excipients did not explain much of the variance of the calculated F1 and F2 factors for these media. Accordingly, we could not estimate confidence intervals for contents of excipients for optimizing the calculated F1 and F2 factors in HCL and acetate media. Hence, we primarily utilized the calculated F1 and F2 factors in USP medium as response variables for optimization of formulation and determination of the design space. Contents of Sucrose and Sodium starch glycolate were significant predictors for F1 dissimilarity factor ($p < 0.0001$, and $p < 0.000$ respectively) and F2 similarity factor ($p < 0.000$, and $p < 0.000$ respectively) in the USP medium.

As we primarily aimed to synchronize optimization of comparative dissolution results with chemometric prediction of USP dissolution results, we proceeded to select a time point, at which the USP dissolution results of the experimental tablets were most varied, and further integrate it as response in both chemometric and QbD models. The variation in dissolution results was assuring effect of the varied contents of excipients. In contrast to similarity factors as response variables in the QbD model, this newly integrated response should be mainly calculated for each single tablet of the experimental formulations, so it was suitable to be coupled with NIR records of each tablet for chemometric prediction of dissolution. Variations in USP dissolution results at all tested time points for the twenty experimental runs are presented in Fig. 1. Variance of dissolution results across the tested formulations was increasing starting from the first time point (5 min), reaching maximum variance at the third time point (15 min), and then start to decrease till it reaches the lowest variance at the last time point (60 min). Among USP dissolution results at each time point, dissolution at 15 min was the most correlated variable (Pearson correlation) with the F1 and F2 factors in USP medium ($r = 0.7$, $p < 0.000$ and -0.73, $p < 0.000$ respectively), it had also good correlation with the results at the other time points (correlation matrix between dissolution results is supplied in supplementary material).

To verify selection of USP dissolution result at 15 min as a response in the QbD model, dissolution results at the first four time points of testing -time points with the highest variances- along with USP similarity factors were subjected to two step cluster analysis using SPSS software (V25.0, IBM, USA). This clustering algorithm optimally divided the results into three clusters based on their mean values of the input variables. Again, dissolution result at 15 min was the most important predictor, indicating it has the highest variance among the input variables. 18.3% of the experimental tablets (cluster 3) were meeting the criteria of F1 and F2 factors in the USP medium (mean values of 6.00 and 64.91 respectively). It is worth noting that Cluster 3 was having a mean percent of drug dissolved at 15 min of 43.65% (the lowest mean value among the three clusters), matching with the innovator results at this time point which ranged from 35 to 65%. Run 2 was the most predominant formula (27.3%) in cluster 3, the results are presented in Fig. 2.

Accordingly, amounts of excipients were modelled with new response variable -dissolution at 15 min- using Design Expert software. ANOVA model was significant ($p = 0.0003$) with predicted R^2 of 0.5051, adjusted R^2 of 0.6284 and the lack of fit test was insignificant ($p = 0.8030$). Amounts of sucrose and sodium starch glycolate were significant predictors ($p < 0.000$ and 0.0012, respectively).

Finally, the calculated F1 and F2 factors in USP medium along with the dissolution results at 15 min were utilized as three response variables for determination of the design space using Design Expert software. The target range for F1 dissimilarity factor was (1–15), the target range for F2 similarity factor was (50–100), while the target range for dissolution at 15 min was (35–65%) matching with that of the innovator. Overlay and desirability plots for these responses are presented in Fig. 3.

The overlay plots clearly showed that the contents of sucrose were negatively associated to F1 dissimilarity factor and positively associated with F2 similarity factor in USP medium. Content of sodium starch glycolate had the same association pattern to a lesser extent, while the contents of avicel Ph 101 and lactose monohydrate did not show association with the response variables, matching with the results of the ANOVA models. The desirability plot suggested that the target values of the response variables can be achieved by utilizing sucrose in range from 18 to 25 mg, and sodium starch glycolate in range from 5 to 7.22 mg with any convenient content of avicel Ph 101 and lactose monohydrate. Hence, the software suggested 64 solutions to achieve the target response variables, modifying content of the excipients within the design space. The optimized formulation was designed by modifying experimental formulation no. 18 which met the criteria

Table 2
The calculated dissimilarity (F1) and similarity (F2) factors, ANOVA models summary and regression coefficients.

Run	F2 in USP	F1 in USP	F2 in HCL	F1 in HCL	F2 in Acetate	F1 in Acetate
1	29	36	36	33	75	2
2	63	6	40	24	41	12
3	37	24	30	45	83	1
4	39	21	30	40	58	5
5	55	10	69	7	48	8
6	46	18	59	10	48	7
7	46	15	53	13	44	9
8	66	6	64	7	49	7
9	41	21	40	27	49	10
10	37	24	25	56	85	1
11	45	19	40	29	62	3
12	73	4	67	6	43	10
13	43	18	37	29	51	9
14	46	18	34	36	55	7
15	58	8	41	23	44	8
16	37	27	26	54	66	4
17	37	27	48	19	65	5
18	44	16	56	11	62	5
19	36	28	28	49	79	4
20	41	22	30	45	68	5
Models' summary*						
Predicted R^2	0.5253	0.5983	0.4004	0.3896	0.1641	0.0706
Adjusted R^2	0.6611	0.6976	0.5627	0.5392	0.3095	0.2162
Lack of fit (p value)	0.2427	0.1716	0.1836	0.5195	0.9758	0.9869
Standardized coefficients (p value)						
Sucrose	78.97 (<0.0001)	−7.16 (<0.0001)	81.21 (0.0002)	−15.44 (0.0003)	26.47 (0.0042)	12.64 (0.0117)
Sod. Starch glycolate	−32.07 (0.0006)	71.66 (0.0008)	37.33 (0.8203)	32.42 (0.8715)	84.23 (0.4332)	−5.27 (0.1561)
Lactose monohydrate	41.6 (0.4761)	20.12 (0.7046)	20.98 (0.0128)	53.28 (0.0144)	67.66 (0.3709)	5.01 (0.6542)
Avicel pH 101	40.99 (0.5657)	28.41 (0.1022)	22.89 (0.0920)	50.28 (0.8715)	80.44 (0.1396)	3.52 (0.4578)

* All models were linear, and significant P < 0.001

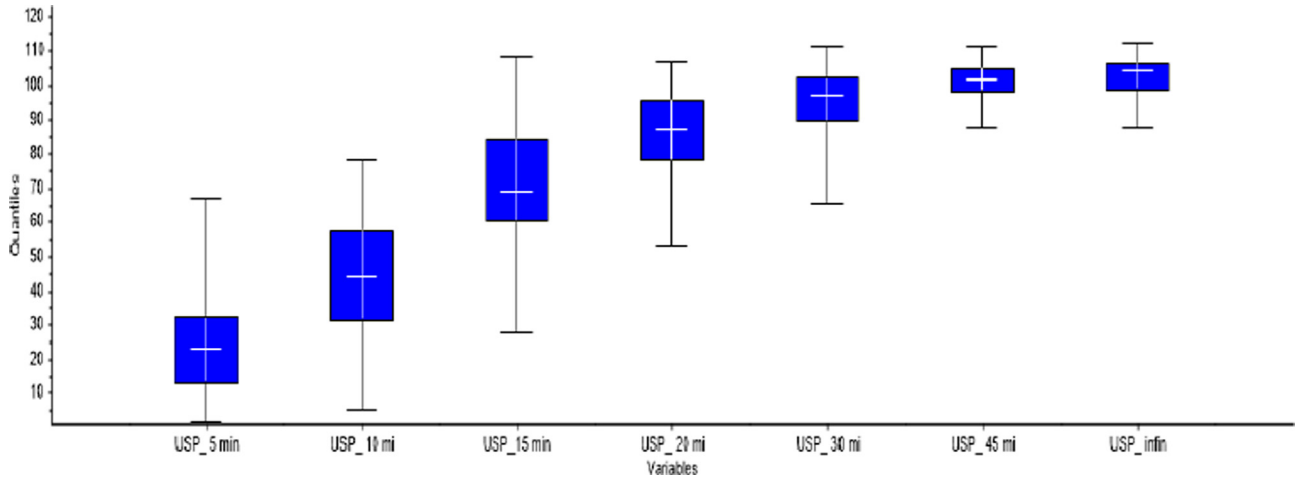

Fig. 1. Variances of dissolution results in USP medium at each time point.

for F1 and F2 factors in HCL and acetate media. Content of Sodium starch glycolate was modified from 8 to 7 mg while content of lactose monohydrate was modified from 28 to 29 mg, to pull contents of sucrose and sodium starch glycolate into the design space. The optimized formulation was subjected to comparative dissolution testing with the innovator, and it met the criteria for F1 and F2 factors in HCL medium (F1 = 9 and F2 = 58), acetate medium (F1 = 3 and F2 = 63) and USP medium (F1 = 12, F2 = 52).

3.2. Chemometric model for prediction of USP dissolution profile

As discussed in the previous section, results of USP dissolution were not varying at all tested time points in a way that allows for chemometric prediction. After being verified to be the most correlated with the values of F1 and F2 factors in USP medium, we used the dissolution result at 15 min as the response variable of the chemometric prediction.

Upon NIR measurement of experimental tablets, each spectrum consisted of 2203 spectral points. After calculation of the corresponding dissolution results, we created a calibration set with NIR spectral points as predictors' matrix and percentage of drug dissolved in the USP medium at 15 min as response variable. OPUS® Quant software was used for wavelength selection and mathematical preprocessing. The software optimally selected 1427 spectral points (9500–4000 cm^{-1}) for calibration. Further, orthogonal signal correction was applied as mathematical preprocessing method.

The processed datapoints was subjected to Principal Component Analysis (PCA) and Partial Least Square (PLS) regression analysis. PC1 and PC2 retained 65% and 8% respectively of the total

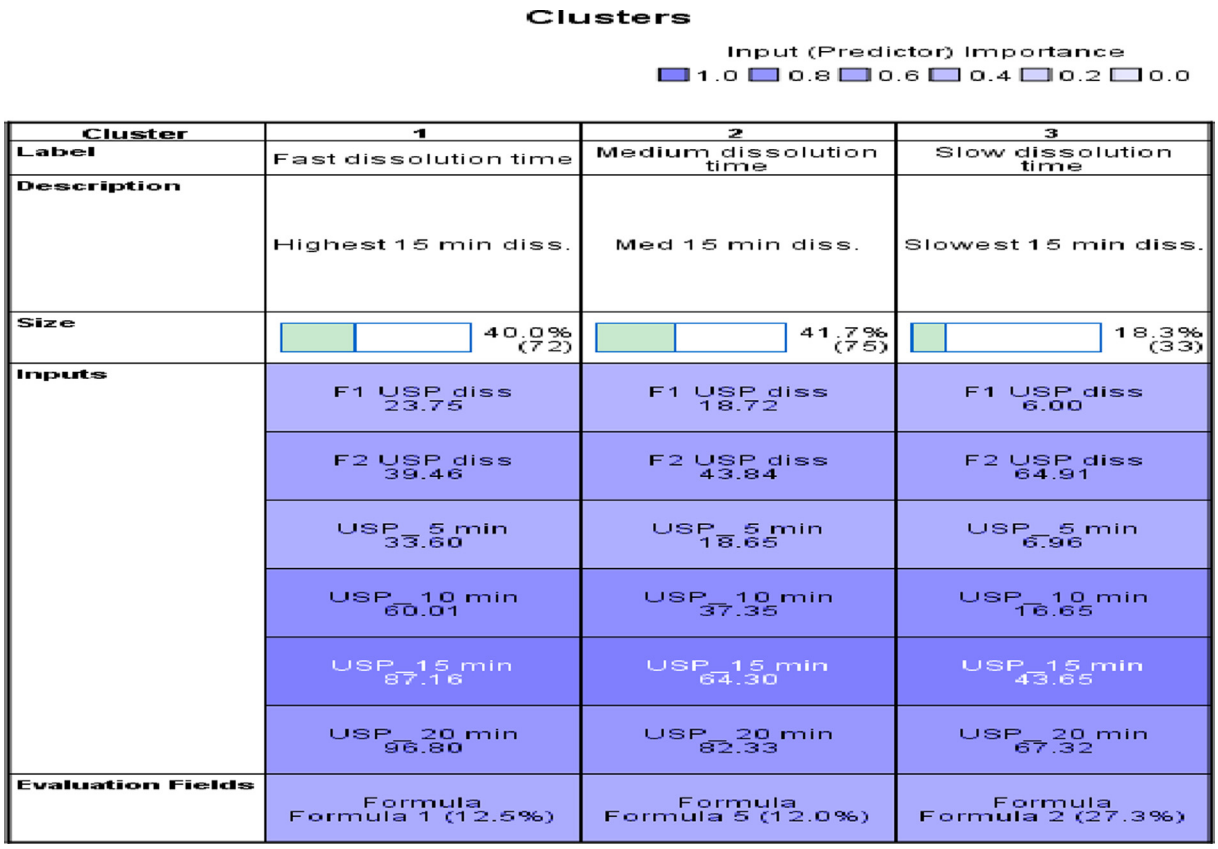

Fig. 2. Two step cluster analysis for four dissolution time points with the calculated similarity factors in USP medium.

variance of the processed data points, indicating that the principal components kept most of the information within processed spectral data. The samples were classified into 3 levels based on the value of the response variable (27.45–54.26%, 54.27–81.07% and 81.08–107.88%, respectively). The first two principal components were able to discriminate the 3 levels of the response variable Fig. 4.

PLS regression model was preliminary validated with leave one out cross validation method. The model had R^2 of cross validation of 0.721 and RMSECV of 9.403. The results were adequate for application as a medium impact model, which can efficiently predict the level of dissolution at 15 min, indicate the similarity with the innovator in USP medium, and contributes to assuring quality of the products but they are not the sole indicators.

Four formulations (Run no. 1, 2, 12 and 15) with varied dissolution results at 15 min were prepared at larger scale (500 tablets) for external testing of the calibration model. NIR measurements of these experimental tablets were processed with the same mathematical preprocessing method of the calibration set, and then the PLS model was used to predict the response variable. The experimental tablets were further subjected to dissolution analysis with reference HPLC method, the prediction vs HPLC results are presented in Table 3.

We further aimed to investigate how NIR explained the presented extent of variance of the response variable, and how we could improve the percentage of explained variance and relatively decrease prediction errors. Utilizing PCA and PLS methods, we modeled the NIR records of the twenty experimental formulations against the content of excipients within the experimental tablets as response variables. Interestingly, the processed spectral data was able to explain the variance of the content of excipients to a great extent. Models were validated with Leave one out cross validation method. PCA classification showed excellent discrimination of different contents of the excipients with the processed NIR data. The R^2 of cross validation were 0.966, 0.8361, 0.963, and 0.962 for Ssucrose, sodium starch glycolate, lactose monohydrate and avicel Ph 101, respectively. indicating for a good prediction power. PCA classification for each of the four excipients are presented in Fig. 5.

At this point, we further proceeded to utilize Structural Equation PLS Modelling (PLS-SEM) to underline the significant paths that explain the variance of dissolution test results. PLS-SEM model was able to combine the conducted PCA and PLS models for prediction of dissolution and contents of excipients into one single model and further assess the significant predictors for USP dissolution results at 15 min. Using smartPLS 3 software, we constructed a formative PLS-SEM model in which content of excipients along with the principal components resulted from NIR chemometric models (content of excipients and dissolution models) were utilized as observed variables. The observed variables in turn form latent variables for which we were assessing the interactions Fig. 6. Values inside the circles are the R^2, the values on arrows directed from observed (yellow rectangles) to latent variables (blue circles) are factor loadings, while the values on arrows directed latent variable to another are regression coefficients.

The interesting finding was that the content of excipients along their corresponding NIR principal components explained 80.4% of the variance of NIR principal components of dissolution model. Since it is well known that NIR is sensitive to both physical and chemical attributes of the pharmaceutical formulations (Siddiqui et al., 2017), we may refer the unexplained variance (19.6%) to the unobserved calibrated variables with varying amounts of excipients. It was clear that capability of NIR to determine concentration of excipients contributes greatly and significantly to how it

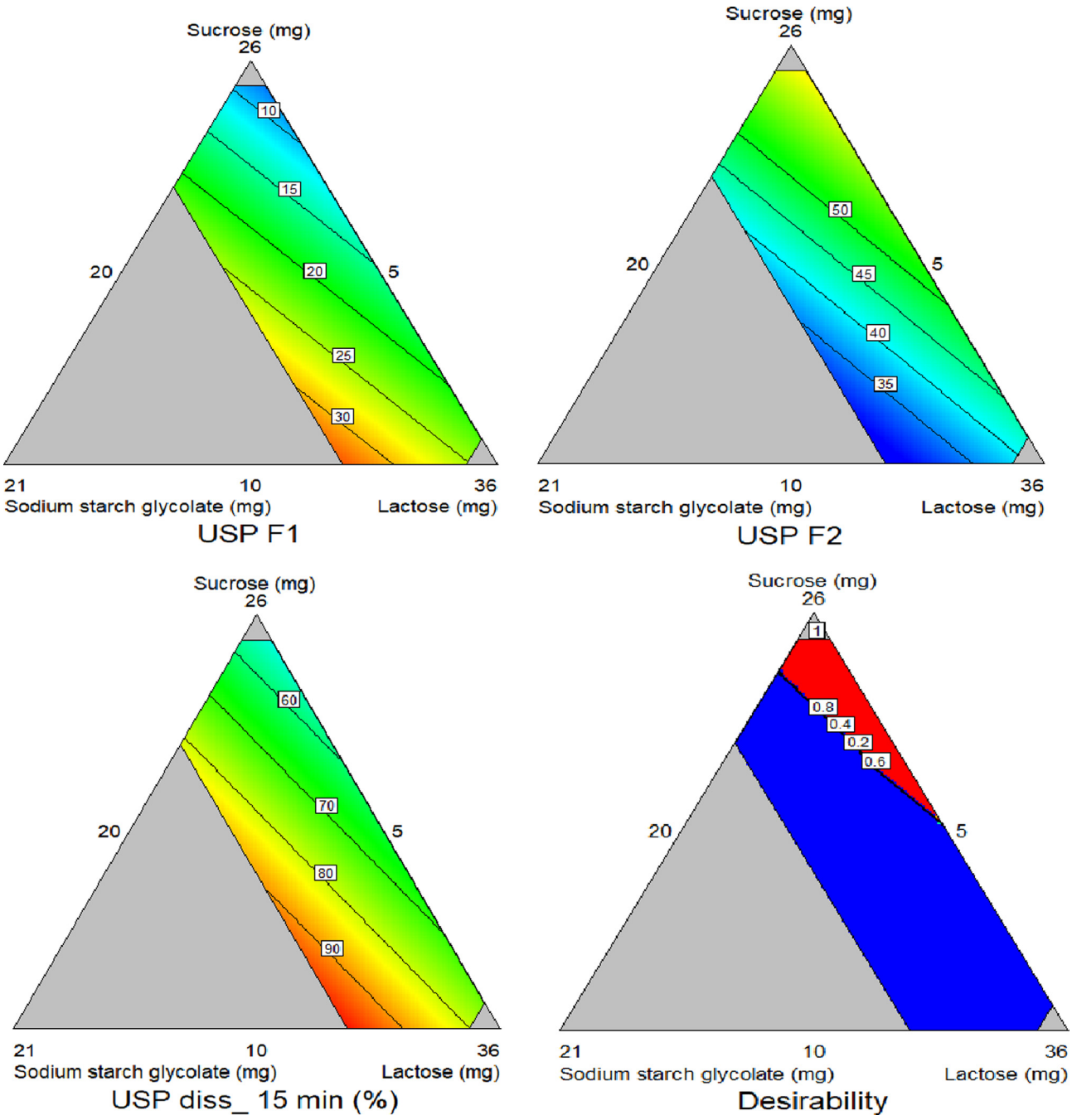

Fig. 3. Overlay and desirability plots (at 15 mg of Avicel pH 101).

explains the dissolution phenomenon, confirming the theory that NIR is very sensitive to pharmaceutical excipients (Hédoux, 2016).

Towards NIR principal component of dissolution model, NIR Principal components of sucrose model and sodium starch glycolate model had the highest standardized regression coefficients of −0.488 and 0.370, respectively. At the final part of model, SEM-PLS model explained 81.8% of the variance of USP dissolution results at 15 min. We performed non-parametric bootstrapping (5000 samples) to determine the statistically significant paths in the model. The results are presented in Table 4.

Confirming on the results of QbD models, paths showing effect of NIR PC scores of contents of sodium starch glycolate and sucrose towards NIR PC scores of dissolution testing were significant ($p < 0.000$ and $p = 0.049$, respectively). In contrast to Artificial Neural Network (ANN) techniques, PLS-SEM was able to quantify contributions of predictors and determine the significant pathways. Hence, it could be used in later studies to quantify the contribution of variables in mixed models which contain formulation and process variables as predictors.

3.3. Chemometric model for prediction of content uniformity

Carvedilol was representing a considerable proportion of the tablet weight (25%), making it suitable for accurate and precise prediction of API content. After designing optimized formula as discussed in section 3.1, it was utilized with different contents of carvedilol (90%, 95%, 100%, 105%, 110% of the label claim) for NIR prediction of content uniformity. Upon NIR measurement of exper-

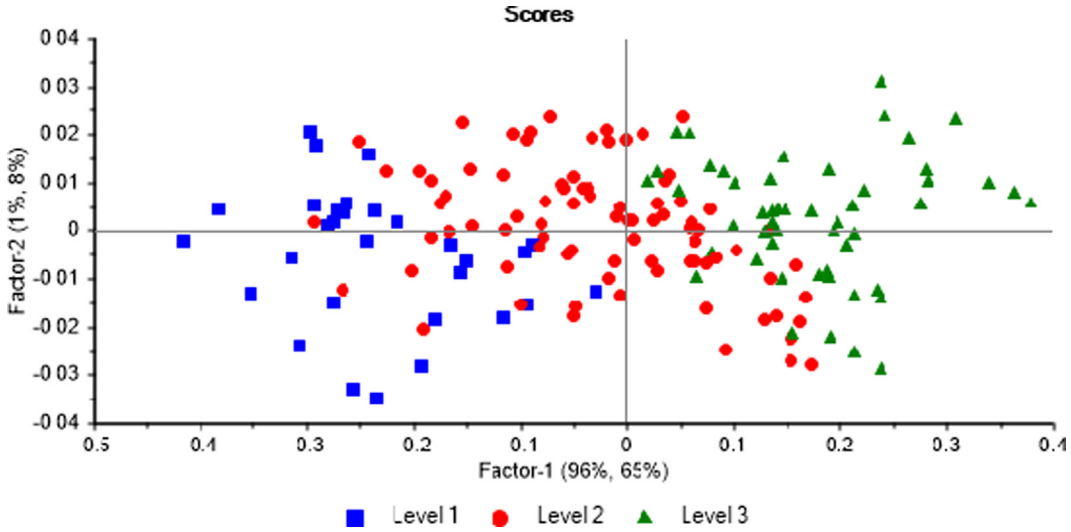

Fig. 4. PC scores of dissolution experimental samples on PC1 and PC2 of the process spectral points.

Table 3
Results of HPLC vs. NIR predictions for dissolution testing at 15 min.

Run	Average result of NIR prediction (%)	Average result of HPLC analysis (%)	Error (%)	RMSEP
1	115.4	99.86	15.54	7.084457
2	43.95	40.16	3.79	
12	56.61	46.91	9.7	
15	34.27	35.02	−0.75	

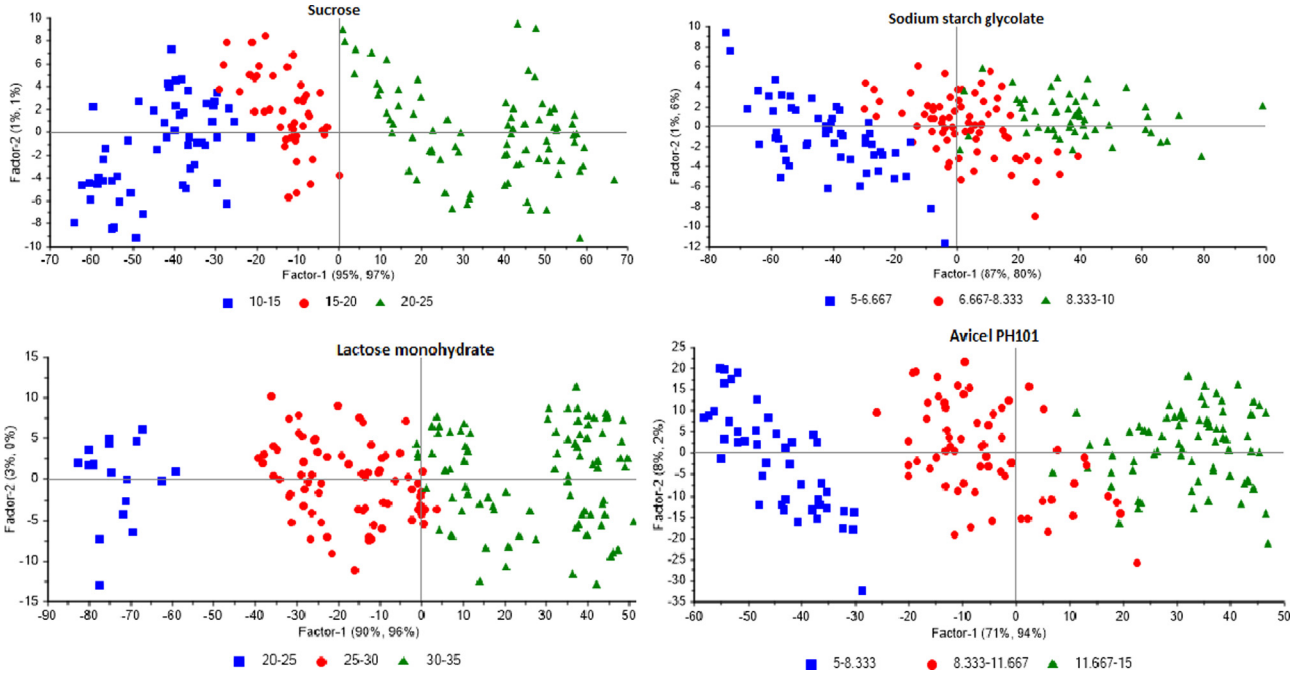

Fig. 5. PCA classification of excipients' content with processed NIR spectral points.

imental tablets, each spectrum consisted of 4407 spectral points. OPUS® software optimally selected 988 spectral points ranged from 9401 cm^{-1} to 7498 cm^{-1}. Mean centering and vector normalization was performed for mathematical preprocessing for selected spectral data points. The processed data was modelled using PCA classification and PLS regression against the HPLC analysis results which ranged from 87.9 to 113.54%.

Upon performing PCA on processed NIR data, PC1 retained 77% of the total variance within the spectral data, while PC2 retained 7% (Fig. 7). The Fig. shows clustering the experimental sample into five levels of API contents in (mg). The data were further subjected to PLS regression, the model was preliminary validated by leave one out cross validation method. The model had R^2 of cross validation of 0.909 and RMSECV of 2.26.

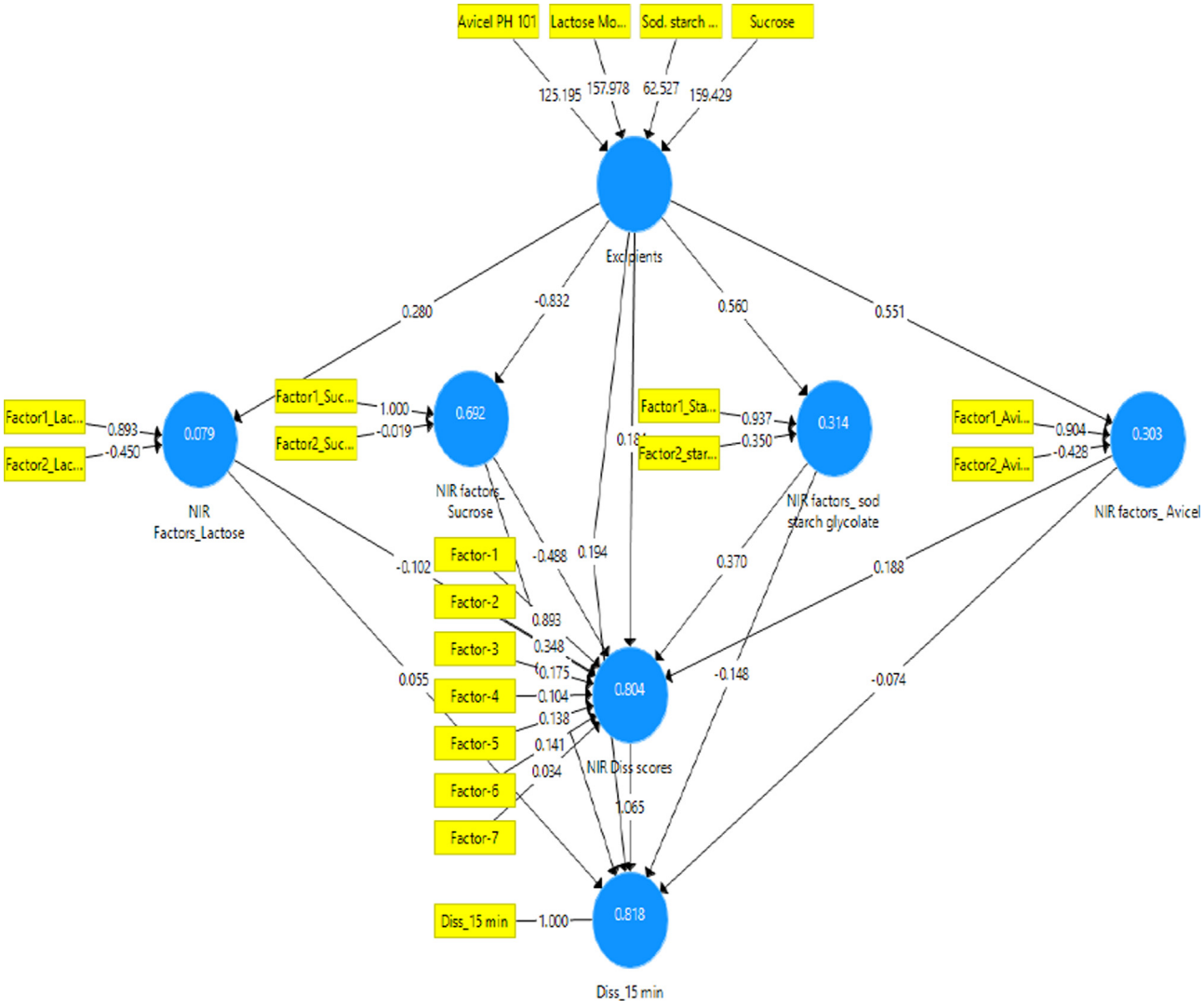

Fig. 6. PLS-SEM formative model.

Table 4
Bootstrapping results for SEM-PLS Model.

Path	T Statistics	p values
Excipients -> Diss_15 min	0.967	0.334
Excipients -> NIR Diss scores	0.921	0.357
Excipients -> NIR Factors_ Lactose	0.845	0.398
Excipients -> NIR factors_ Avicel pH 101	1.477	0.14
Excipients -> NIR factors_ Sucrose	6.63	0.000
Excipients -> NIR factors_ sod starch glycolate	4.203	0.000
NIR Diss scores -> Diss_15 min	12.034	0.000
NIR Factors_ Lactose -> Diss_15 min	0.22	0.826
NIR Factors_ Lactose -> NIR Diss scores	0.451	0.652
NIR factors_ Avicel pH 101 -> Diss_15 min	0.443	0.658
NIR factors_ Avicel pH 101 -> NIR Diss scores	1.004	0.316
NIR factors_ Sucrose -> Diss_15 min	0.905	0.366
NIR factors_ Sucrose -> NIR Diss scores	1.965	0.049
NIR factors_ sod starch glycolate -> Diss_15 min	1.252	0.211
NIR factors_ sod starch glycolate -> NIR Diss scores	3.599	0.000

3 formulations (90%, 100% and 110% of the label claim) were prepared on larger scale (500 tablets) for external validation of the model. We aimed to test the points at the extremes of the calibration curve along with the target point at 100% of the label claim. After NIR measurement of the experimental tablets, spectral data was processed with the same methods and utilized for prediction of carvedilol content. Results are presented in Table 5, it revealed that the model is good for application as a medium impact model, which contributes to assuring quality of the products, but they are not sole indicators.

4. Conclusion

Pharmaceutical industry is convoying the great advancements in spectral techniques and mathematical processing of spectral data. There is a remarkable amount of data that can be produced upon in-line spectral monitoring of the pharmaceutical products. However, selection of calibrating variables remains to be the greatest challenge in face of these revolutionary techniques. In real-life pharmaceutical application of process analytical technology (PAT), the calibrating variables should be defined depending on the intended impact (Low-impact, Medium-impact, and High-impact) of the model along with the chemical and physical attributes of the pharmaceutical product. The current work aimed to synchronize the QbD model for optimization of the comparative dissolution results with medium-impact chemometric models that support in-line monitoring of content uniformity and dissolution

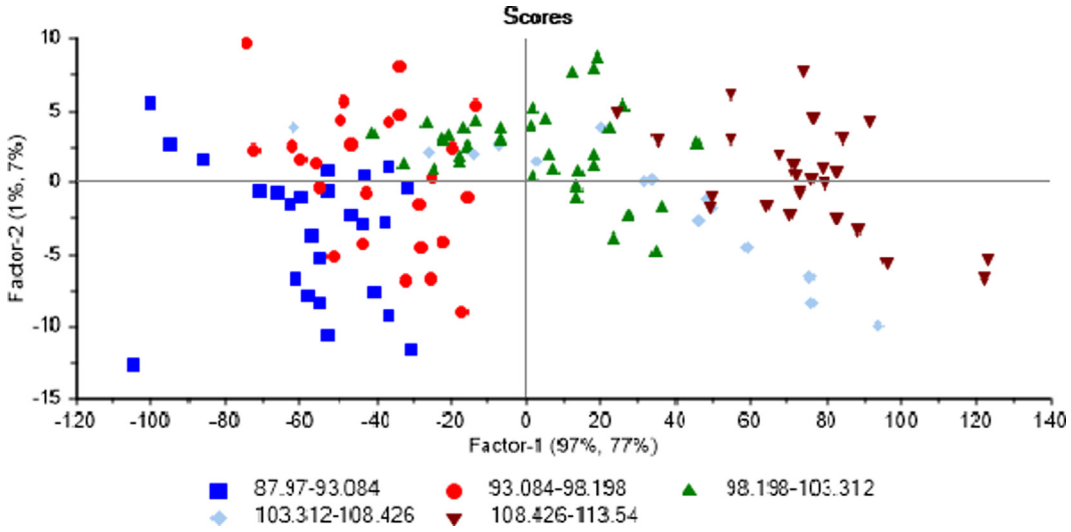

Fig. 7. PCA classification of content uniformity samples with NIR processed spectral points.

Table 5
Actual and predicted results of content uniformity test.

Average result of NIR prediction (%)	Average result of HPLC analysis (%)	Error (%)	RMSEP
91.95	90.53	1.42	2.418684
102.61	99.79	2.82	
108.87	110.76	−1.89	

testing of the product. Selection of calibrating variables of QbD model was carried out by investigating critical process parameters (CPPs) and the historical quality control results of the product. Twenty different settings of varied concentrations of excipients were utilized for optimization of the formulation and for calibrating the in-vitro USP dissolution chemometric model. The optimized formula was utilized with varied contents of the API for calibrating the content uniformity chemometric model. Four settings of dissolution experimental tablets and three settings of content uniformity tablets were produced on larger scale for external testing of the chemometric models. The produced chemometric models were suitable as medium-impact models. NIR spectra also perfectly predicted the contents of the excipients. Combining PCA and PLS models for prediction of USP dissolution and the contents of excipients into one PLS structural equation modeling (PLS-SEM) revealed the significant paths that contributes to explanation of dissolution phenomenon.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors would like to thank R&D team at Marcyrl pharmaceuticals for their valuable support. We want to express deepest gratitude to Dr. Hany Soliman and Dr. Ahmed Soliman at Hikma pharmaceuticals for their appreciated support through years of study conduction, and to Dr. Abdelhamid Elshabrawy for his support in conceptualizing the PLS-SEM model.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsps.2021.04.012.

References

Arruabarrena, J., Coello, J., Maspoch, S., 2014. Raman spectroscopy as a complementary tool to assess the content uniformity of dosage units in break-scored warfarin tablets. Int. J. Pharm. 465 (1-2), 299–305. https://doi.org/10.1016/j.ijpharm.2014.01.027.

Bittner, L.K.H., Heigl, N., Petter, C.H., Noisternig, M.F., Griesser, U.J., Bonn, G.K., Huck, C.W., 2011. Near-infrared reflection spectroscopy (NIRS) as a successful tool for simultaneous identification and particle size determination of amoxicillin trihydrate. J. Pharm. Biomed. Anal. 54 (5), 1059–1064. https://doi.org/10.1016/j.jpba.2010.12.019.

Dégardin, K., Guillemain, A., Guerreiro, N.V., Roggo, Y., 2016. Near infrared spectroscopy for counterfeit detection using a large database of pharmaceutical tablets. J. Pharm. Biomed. Anal. 128, 89–97. https://doi.org/10.1016/j.jpba.2016.05.004.

Deng, B.-C., Yun, Y.-H., Ma, P., Lin, C.-C., Ren, D.-B., Liang, Y.-Z., 2015. A new method for wavelength interval selection that intelligently optimizes the locations, widths and combinations of the intervals. Analyst 140 (6), 1876–1885. https://doi.org/10.1039/C4AN02123A.

Dokoumetzidis, A., Macheras, P., 2006. A century of dissolution research: From Noyes and Whitney to the Biopharmaceutics Classification System. Int. J. Pharm. 321 (1-2), 1–11. https://doi.org/10.1016/j.ijpharm.2006.07.011.

FDA, 2019. Quality Considerations for Continuous Manufacturing Guidance for Industry. Food Drug Adm. 1–27.

Fernandes, G.J., Kumar, L., Sharma, K., Tunge, R., Rathnanand, M., 2018. A Review on Solubility Enhancement of Carvedilol—a BCS Class II Drug. J. Pharm. Innov. 13 (3), 197–212. https://doi.org/10.1007/s12247-018-9319-z.

FDA, 2004. PAT—A Framework for Innovative Pharmaceutical Manufacturing and Quality Assurance. Food and Drug Administration. Guid. Ind, Washington, DC.

Freitas, M.P., Sabadin, A., Silva, L.M., Giannotti, F.M., do Couto, D.A., Tonhi, E., Medeiros, R.S., Coco, G.L., Russo, V.F.T., Martins, J.A., 2005. Prediction of drug dissolution profiles from tablets using NIR diffuse reflectance spectroscopy: A rapid and nondestructive method. J. Pharm. Biomed. Anal. 39 (1-2), 17–21. https://doi.org/10.1016/j.jpba.2005.03.023.

Galata, D.L., Farkas, A., Könyves, Z., Mészáros, L.A., Szabó, E., Csontos, I., Pálos, A., Marosi, G., Nagy, Z.K., Nagy, B., 2019. Fast, spectroscopy-based prediction of in vitro dissolution profile of extended release tablets using artificial neural networks. Pharmaceutics 11, 1–18. https://doi.org/10.3390/pharmaceutics11080400.

Galata, D.L., Könyves, Z., Nagy, B., Novák, M., Mészáros, L.A., Szabó, E., Farkas, A., Marosi, G., Nagy, Z.K., 2021. Real-time release testing of dissolution based on surrogate models developed by machine learning algorithms using NIR spectra, compression force and particle size distribution as input data. Int. J. Pharm. 597, 120338. https://doi.org/10.1016/j.ijpharm.2021.120338.

Gervasi, V., Cullen, S., McCoy, T., Crean, A., Vucen, S., 2019. Application of a mixture DOE for the prediction of formulation critical temperatures during lyophilisation process optimisation. Int. J. Pharm. 572, 118807. https://doi.org/10.1016/j.ijpharm.2019.118807.

Granato, D., Putnik, P., Kovačević, D.B., Santos, J.S., Calado, V., Rocha, R.S., Cruz, A.G. D., Jarvis, B., Rodionova, O.Y., Pomerantsev, A., 2018. Trends in Chemometrics: Food Authentication, Microbiology, and Effects of Processing. Compr. Rev. Food Sci. Food Saf. 17 (3), 663–677. https://doi.org/10.1111/crf3.2018.17.issue-310.1111/1541-4337.12341.

Haneef, J., Beg, S., 2021. Chapter 9 - Quality by design-based development of nondestructive analytical techniques, in: Beg, S., Hasnain, M.S., Rahman, M., Almalki, W.H.B.T.-H. of A.Q. by D. (Eds.), . Academic Press, pp. 153–166. DOI: DOI: 10.1016/B978-0-12-820332-3.00006-6

Harting, J., Kleinebudde, P., 2019. Optimisation of an in-line Raman spectroscopic method for continuous API quantification during twin-screw wet granulation and its application for process characterisation. Eur. J. Pharm. Biopharm. 137, 77–85. https://doi.org/10.1016/j.ejpb.2019.02.015.

Hédoux, A., 2016. Recent developments in the Raman and infrared investigations of amorphous pharmaceuticals and protein formulations: A review. Adv. Drug Deliv. Rev. 100, 133–146. https://doi.org/10.1016/j.addr.2015.11.021.

ICH, 2012. Points to consider for ICH Q8/Q9/Q10 guidelines 44.

Liu, R., Li, L., Yin, W., Xu, D., Zang, H., 2017. Near-infrared spectroscopy monitoring and control of the fluidized bed granulation and coating processes—A review. Int. J. Pharm. 530 (1-2), 308–315. https://doi.org/10.1016/j.ijpharm.2017.07.051.

Martens, H., Stark, E., 1991. Extended multiplicative signal correction and spectral interference subtraction: New preprocessing methods for near infrared spectroscopy. J. Pharm. Biomed. Anal. 9 (8), 625–635. https://doi.org/10.1016/0731-7085(91)80188-F.

Mishra, V., Thakur, S., Patil, A., Shukla, A., 2018. Quality by design (QbD) approaches in current pharmaceutical set-up. Expert Opin. Drug Deliv. 15 (8), 737–758. https://doi.org/10.1080/17425247.2018.1504768.

Nagy, B., Farkas, A., Gyürkés, M., Komaromy-Hiller, S., Démuth, B., Szabó, B., Nusser, D., Borbás, E., Marosi, G., Nagy, Z.K., 2017. In-line Raman spectroscopic monitoring and feedback control of a continuous twin-screw pharmaceutical powder blending and tableting process. Int. J. Pharm. 530 (1-2), 21–29. https://doi.org/10.1016/j.ijpharm.2017.07.041.

Nagy, B., Farkas, A., Magyar, K., Démuth, B., Nagy, Z.K., Marosi, G., 2018. Spectroscopic characterization of tablet properties in a continuous powder blending and tableting process. Eur. J. Pharm. Sci. 123, 10–19. https://doi.org/10.1016/j.ejps.2018.07.025.

Ojala, K., Myrskyranta, M., Liimatainen, A., Kortejärvi, H., Juppo, A., 2020. Prediction of drug dissolution from Toremifene 80 mg tablets by NIR spectroscopy. Int. J. Pharm. 577, 119028. https://doi.org/10.1016/j.ijpharm.2020.119028.

Otsuka, M., Tanabe, H., Osaki, K., Otsuka, K., Ozaki, Y., 2007. Chemoinformetrical evaluation of dissolution property of indomethacin tablets by near-infrared spectroscopy. J. Pharm. Sci. 96 (4), 788–801. https://doi.org/10.1002/jps.20704.

Palou, A., Cruz, J., Blanco, M., Tomàs, J., de los Ríos, J., Alcalà, M., 2012. Determination of drug, excipients and coating distribution in pharmaceutical tablets using NIR-CI. J. Pharm. Anal. 2 (2), 90–97. https://doi.org/10.1016/j.jpha.2011.11.003.

Pauli, V., Elbaz, F., Kleinebudde, P., Krumme, M., 2018. Methodology for a Variable Rate Control Strategy Development in Continuous Manufacturing Applied to Twin-screw Wet-Granulation and Continuous Fluid-bed Drying. J. Pharm. Innov. 13 (3), 247–260. https://doi.org/10.1007/s12247-018-9320-6.

Pauli, V., Roggo, Y., Kleinebudde, P., Krumme, M., 2019. Real-time monitoring of particle size distribution in a continuous granulation and drying process by near infrared spectroscopy. Eur. J. Pharm. Biopharm. 141, 90–99. https://doi.org/10.1016/j.ejpb.2019.05.007.

Rantanen, J., Khinast, J., 2015. The Future of Pharmaceutical Manufacturing Sciences. J. Pharm. Sci. 104 (11), 3612–3638. https://doi.org/10.1002/jps.24594.

Riolo, D., Piazza, A., Cottini, C., Serafini, M., Lutero, E., Cuoghi, E., Gasparini, L., Botturi, D., Marino, I.G., Aliatis, I., Bersani, D., Lottici, P.P., 2018. Raman spectroscopy as a PAT for pharmaceutical blending: Advantages and disadvantages. J. Pharm. Biomed. Anal. 149, 329–334. https://doi.org/10.1016/j.jpba.2017.11.030.

Shi, Z., Anderson, C.A., 2010. Pharmaceutical Applications of Separation of Absorption and Scattering in Near-Infrared Spectroscopy (NIRS). J. Pharm. Sci. 99 (12), 4766–4783. https://doi.org/10.1002/jps.22228.

Siddiqui, M.R., AlOthman, Z.A., Rahman, N., 2017. Analytical techniques in pharmaceutical analysis: A review. Arab. J. Chem. 10, S1409–S1421. https://doi.org/10.1016/j.arabjc.2013.04.016.

Singh, H., Khurana, L.K., Singh, R., 2017. Pharmaceutical Development. Pharm. Med. Transl. Clin. Res. 4, 33–46. https://doi.org/10.1016/B978-0-12-802103-3.00003-1.

Swain, S., Parhi, R., Jena, B.R., Babu, S.M., 2019. Quality by Design: Concept to Applications. Curr. Drug Discov. Technol. 16 (3), 240–250. https://doi.org/10.2174/1570163815666180308142016.

Taleuzzaman, M., Kala, C., Alam, J., Rahat, I., Beg, S., 2021. Chapter 8 - Quality by design–based development of vibrational spectroscopy methods, in: Beg, S., Hasnain, M.S., Rahman, M., Almalki, W.H.B.T.-H. of A.Q. by D. (Eds.), . Academic Press, pp. 133–151. DOI: DOI: 10.1016/B978-0-12-820332-3.00001-7

Terra, L.A., Poppi, R.J., 2014. Monitoring the polymorphic transformation on the surface of carbamazepine tablets generated by heating using near-infrared chemical imaging and chemometric methodologies. Chemom. Intell. Lab. Syst. 130, 91–97. https://doi.org/10.1016/j.chemolab.2013.10.009.

United States Pharmacopeial Convention, 2018. The United States Pharmacopeia: USP 41 : The National Formulary: NF 36.

Zhao, Y., Li, W., Shi, Z., Drennen, J.K., Anderson, C.A., 2019. Prediction of Dissolution Profiles From Process Parameters, Formulation, and Spectroscopic Measurements. J. Pharm. Sci. 108 (6), 2119–2127. https://doi.org/10.1016/j.xphs.2019.01.023.

Zhong, L., Gao, L., Li, L., Zang, H., 2020. Trends-process analytical technology in solid oral dosage manufacturing. Eur. J. Pharm. Biopharm. 153, 187–199. https://doi.org/10.1016/j.ejpb.2020.06.008.