Original Article

Chemometrics-Assisted Spectrophotometric Method Development and Validation for Simultaneous Estimation of Emtricitabine, Tenofovir Alafenamide Fumarate, and Dolutegravir Sodium in Dosage Form

Sapna M. Rathod, Paresh U. Patel[1]

Research Scholar, Ganpat University, Kherva, Gujarat, India, [1]Department of Pharmaceutical Chemistry and Quality Assurance, S.K. Patel College of Pharmaceutical Education and Research, Ganpat University, Kherva, Gujarat, India

Abstract

Aim: This study aims on the development of a chemometric-assisted spectroscopic method for the analysis of combined dosage form of emtricitabine (EMT), tenofovir alafenamide fumarate (TEN), and dolutegravir sodium (DOL). The use of a multivariate algorithm to analyse spectrophotometric data is a novel approach to estimating drug concentrations in formulations. **Materials and Methods:** The quantitative estimation of EMT, TEN, and DOL in tablets was carried out using four chemometric approaches: Classical least square (CLS), inverse least square, partial least square, and principal component regression. Thirty-two ternary mixtures of calibration sets and 16 mixtures of validation sets were prepared. The absorbance data matrix was attained by calculating absorbance at 25 different wavelengths in a range of 240–336 nm ($\Delta\lambda$ = 4 nm). The chemometric calculations were performed using Matlab2018a and Minitab software. The developed methods were validated. **Results:** The great accuracy of the current study was justified by the near-perfect recovery values (100%) and low standard deviation. For chemometrics approaches, the root mean square error of calibration (RMSEC), root mean square error of prediction (RMSEP), and root mean square error of cross-validation (RMSECV) outcomes display decent accuracy and precision. **Conclusion:** The CLS approach yielded the lowest predicted residual error sum of squares, RMSEC, RMSEP, and RMSECV scores. As a result, CLS might be regarded as the best chemometric approach among all techniques utilized. The label claim determined is in excellent accordance with the mean recoveries for EMT, TEN, and DOL. So, it can be used in quality control laboratories.

Keywords: *Chemometrics, emtricitabine, spectrophotometric, tenofovir alafenamide fumarate and dolutegravir sodium*

List of Abbreviations

Emtricitabine; EMT, Tenofovir Alafenamide Fumarate; TEN, Dolutegravir Sodium; DOL, Classical Least Square; CLS, Inverse Least Square; ILS, Partial Least Square; PLS, and Principal Component Regression; PCR, Standard deviation; SD, Root mean square error of Calibration; RMSEC, Root mean square error of Prediction; RMSEP, Root mean square error of Cross Validation; RMSECV, Predicted residual error sum of squares; PRESS, Analytical Reagent; A.R., Leave one out; LOO, correlation coefficient; R2, Figure of merits; FOM, Sensitivity; SEN, Limit of Detection; LOD, limit of quantitation; LOQ.

Introduction

In chemometrics, there are two types of data set namely, calibration and validation data set. The findings of calibration data set were utilized to determine the component concentrations in unknown sample.[1] For both classical statistics and chemometric approaches, there is currently a considerable amount of computer software readily available.[2] Chemometrics methods are particularly useful approaches for analyzing many compounds at the same time, in which the overlap of the active compounds' spectra generates an interference that makes determining the amounts of each component impossible.[3] In addition, chemometric calibration methods are simple since they can evaluate a large number of samples in a short amount of time more accurately and precisely compared with other methods. It can be described as the use of mathematical and statistical approaches to create and/or optimize measurement procedures, as well as the analysis of pertinent data to offer chemical information.[4] Multivariate calibrations such as classical least square (CLS), inverse least square (ILS), principle component regression (PCR), and partial least

How to cite this article: Rathod SM, Patel PU. Chemometrics-assisted spectrophotometric method development and validation for simultaneous estimation of emtricitabine, tenofovir alafenamide fumarate, and dolutegravir sodium in dosage form. J Rep Pharma Sci 2022;11:41-50.

Received: 23 Jul 2021
Accepted: 01 Jan 2022
Published: 29 Jun 2022

Address for correspondence:
Ms. Sapna M. Rathod, Research Scholar, Ganpat University, Kherva 384001, Gujarat, India.
E-mail: srathod456@gmail.com

Access this article online

Website: www.jrpsjournal.com

DOI:10.4103/jrptps.JRPTPS_105_21

Quick Response Code:

square (PLS) have been widely used in quantitative spectrum analysis in recent years to extract selective information from unselective data.[5,6] These approaches are commonly used since they produce the greatest outcomes when it comes to resolving complex mixtures.[7] These approaches can be used to estimate medications in pharmaceutical formulations containing two or more drug components using simultaneous spectrophotometric methods.[8] CLS and ILS are two of the most basic approaches, both based on Beer's principle and using a multivariate least square procedure. Factor analysis methods such as PCR and PLS are used to establish a link between chemical data matrices.[9,10]

Emtricitabine (EMT), also known as 2′,3′-dideoxy-5-fluoro-3′-thiacytidine (FTC), is a synthetic nucleoside reverse transcriptase inhibitor (NRTI) that is taken once a day orally.[11] Emtricitabine 5′ triphosphate, an active metabolite formed by intracellular kinases phosphorylating emtricitabine, inhibits HIV reverse transcriptase by competing for entrance into the HIV DNA chain with the endogenous substrate 2′-deoxycitidine 5′-triphosphate.[12] Because emtricitabine 5′-triphosphate lacks a hydroxyl group in the 3′ position of the sugar moiety, it causes chain termination when it is incorporated into the HIV DNA chain.[13,14]

Tenofovir alafenamide fumarate (TEN), chemically propan-2-yl(2S)-2-{[(S)-({[(2R)-1-(6-amino-9H-purin-9-yl)propan-2-yl]oxy}methyl)(phenoxy)phosphoryl]amino}propanoate.[15] TEN has a molecular weight of 476.47 g/mol and a chemical formula of $C_{21}H_{29}N_6O_5P$. TEN is a NRTI and tenofovir prodrug is used to treat HIV-1 infection. The drug has a solubility in water, methanol, and dimethyl sulfoxide.[16,17]

Dolutegravir sodium (DOL) is integrase strand transfer inhibitor. The drug prevents the viral genome from being integrated into the host cell by blocking the strand transfer stage. It is chemically sodium; (3S,7R)-13-[(2,4-difluorophenyl) methylcarbamoyl] -7-methyl-9,12-dioxo-4-oxa-1,8-diazatricyclo[8.4.0.03,8] tetradeca-10,13-dien-11-olate.[18,19]

No reported analytical methods were found for estimating EMT, TEN, and DOL in bulk and in their combination dosage form. As a result, the current study was attempted to design and validate multivariate approaches for resolving complex drug spectra.

Instrumentation and Software

To test the absorbance of all the solutions, a shimadzu model 1700 (Japan) double beam UV/Visible spectrophotometer with a spectral width of 2 nm, wavelength accuracy of 0.5 nm, and a pair of 10 mm matched quartz cells was employed. The spectra of various calibration and validation sets were recorded using UV probe software. Chemometric calculations were performed using MATLAB-R2018a Software, Minitab 16.1.1, and Microsoft Excel 2010. MVC1toolbox (with MATLAB) was used to estimate figures of merit for multivariate calibration models.

Materials and Methods

TEN was kindly gifted by Bulat Pharmaceuticals, Hyderabad, Andhra Pradesh. EMT was provided as gift samples from Amneal Pharmaceuticals, Ahmedabad, Gujarat, and DOL were provided as gift samples from Cipla Pharmaceuticals, Mumbai, Maharashtra. The local market provided the commercial combination tablet (SPEGRA). Methanol (A.R.) grade (SD fine grade chemicals Ltd.), distilled water, and other chemicals used were of analytical grade. No. 41 Whatman filter paper was utilized in the study.

Preparation of solutions

Each drug was accurately weighed and transported to a separate volumetric flask and dissolved using methanol, and the volume was brought up to the mark using methanol (1000 µg/mL). Aliquot each drug's standard stock solution to a separate volumetric flask to produce a working standard solution of 100 µg/mL, using distilled water as diluent.

Preparation of calibration set and validation set

A data set of calibration samples was created using fractional factorial design. A total 32 ternary mixture solutions were prepared by mixing known amount of drugs under study in varied proportions. Validation set consisting of 16 samples was prepared from the working solutions in the same manner as that of calibration set. The composition of calibration as well as validation set were represented in Table 1.

Optimization and selection of method parameters

Sample solutions were analyzed across 200–400 nm for calibration and validation datasets, and zero-order spectra were obtained [Figure 1]. The absorbance data from the spectrum regions of 200–220 nm with noise and 350–400 nm with zero reading were excluded because they were not essential for the chemometric approach. Wavelength in the range of 240–336 nm was chosen to produce minimal root mean square error of calibration (RMSEC) and root mean square error of cross-validation (RMSECV) values.

Classical least square

CLS is also known as K matrix. Basically, it involves the usage of multiple linear regression to represent the Beer–Lambert law of spectroscopy in a classical way.

$$A = KC$$

Calibration set comprised of concentration matrix, C, and an absorbance matrix, A for known sets of samples is constructed to generate calibration using CLS. In MATLAB2018a software, the CLS model was developed by adding absorbance (A) and concentration matrix (C) data.

The calculated K can be used to forecast the concentration of an unknown sample, C_{unk}, based on its measured spectrum, and it can be stored as an absorbance matrix, A_{unk}.

There are mainly two subclasses of CLS namely, direct CLS and indirect CLS. The K matrix is calculated in direct CLS by measuring the spectra of the pure component, either neat or in a nonabsorbing solvent. In the indirect CLS technique, pure spectra are calculated from mixture spectra rather than being measured directly.

Table 1: Composition of calibration set and validation set for mixtures

Calibration Set							
MIX	Concentrations (µg/mL)			MIX	Concentrations (µg/mL)		
	EMT	TEN	DOL		EMT	TEN	DOL
1	30	5	10	17	10	5	30
2	5	30	5	18	30	30	20
3	5	30	10	19	5	2	30
4	30	2	5	20	40	5	5
5	5	5	10	21	40	2	5
6	40	2	5	22	10	2	20
7	10	2	30	23	30	20	30
8	30	5	5	24	40	30	30
9	30	2	10	25	10	20	5
10	10	2	10	26	5	20	5
11	10	30	5	27	40	30	20
12	5	5	20	28	40	20	30
13	10	20	10	29	30	30	30
14	10	5	20	30	5	20	5
15	5	2	20	31	40	20	20
16	40	5	10	32	30	20	20
Validation Set							
MIX	Concentrations (µg/mL)			MIX	Concentrations (µg/mL)		
	EMT	TEN	DOL		EMT	TEN	DOL
1	20	2	30	9	10	20	30
2	5	30	30	10	20	30	5
3	40	10	5	11	30	10	5
4	30	20	10	12	5	2	10
5	10	5	10	13	30	5	20
6	40	2	15	14	10	20	20
7	20	10	15	15	30	30	10
8	5	5	5	16	10	10	15

EMT = emtricitabine, TEN = tenofovir alafenamide fumarate, DOL = dolutegravir sodium

Absorbance matrix *A* is comprised of zero-order spectra at 4 nm intervals between 240 and 336 nm, that is, absorbances at 25 wavelength points. The developed model comprised absorbance values of samples at 25 various wavelength points, and quantities of EMT, TEN, and DOL in the validation data set as well in tablet formulations were predicted.

Inverse least square

It is also called as *P* – matrix calibration as it originally requires the use of multiple linear regression to calculate the inverse expression of the Beer–Lambert equation of spectroscopy.

$$C = PA$$

where, *C* = concentration matrix,

P = calibration coefficient, and

A = absorbance matrix.

To determine *P*, a training set containing a concentration matrix, *C*, and an absorbance matrix, *A*, is used to create a calibration using ILS. ILS differs from the classical technique, which involves fitting a linear mixture of pure spectra to an unknown spectrum. This distinction provides ILS with several advantages. When all of the system's components aren't explicitly evaluated, CLS fails to provide accurate predictions.

The software MATLAB2018a was used to construct the approach. The samples' absorbance values were inputted into the calibrations at 25 various wavelength points in the spectral area in a range of 240–336 nm. The concentrations of EMT, TEN, and DOL in validation set as well in tablets were predicted.

Partial least square and principal component regression

These are the most widely used methods in the multivariate calibration approach. The inverse calibration methodology is used in both procedures. The PCR is a method that operates on the principle of lowering the original data's dimensionality. The original variables are replaced with linear combinations of the variables in both PLS and PCR to solve the inversion problem (factors). The PCR employs the well-known singular value decomposition method. Using the converted data as input, this function was used to fit a PCR model. When fitting the PCR model, the leave-one-out (LOO) cross-validation approach was utilized. The optimal principal components (or eigenvectors) corresponding to the large eigenvalues are identified using cross-validation in the calibration step. The PLS employs a nonlinear

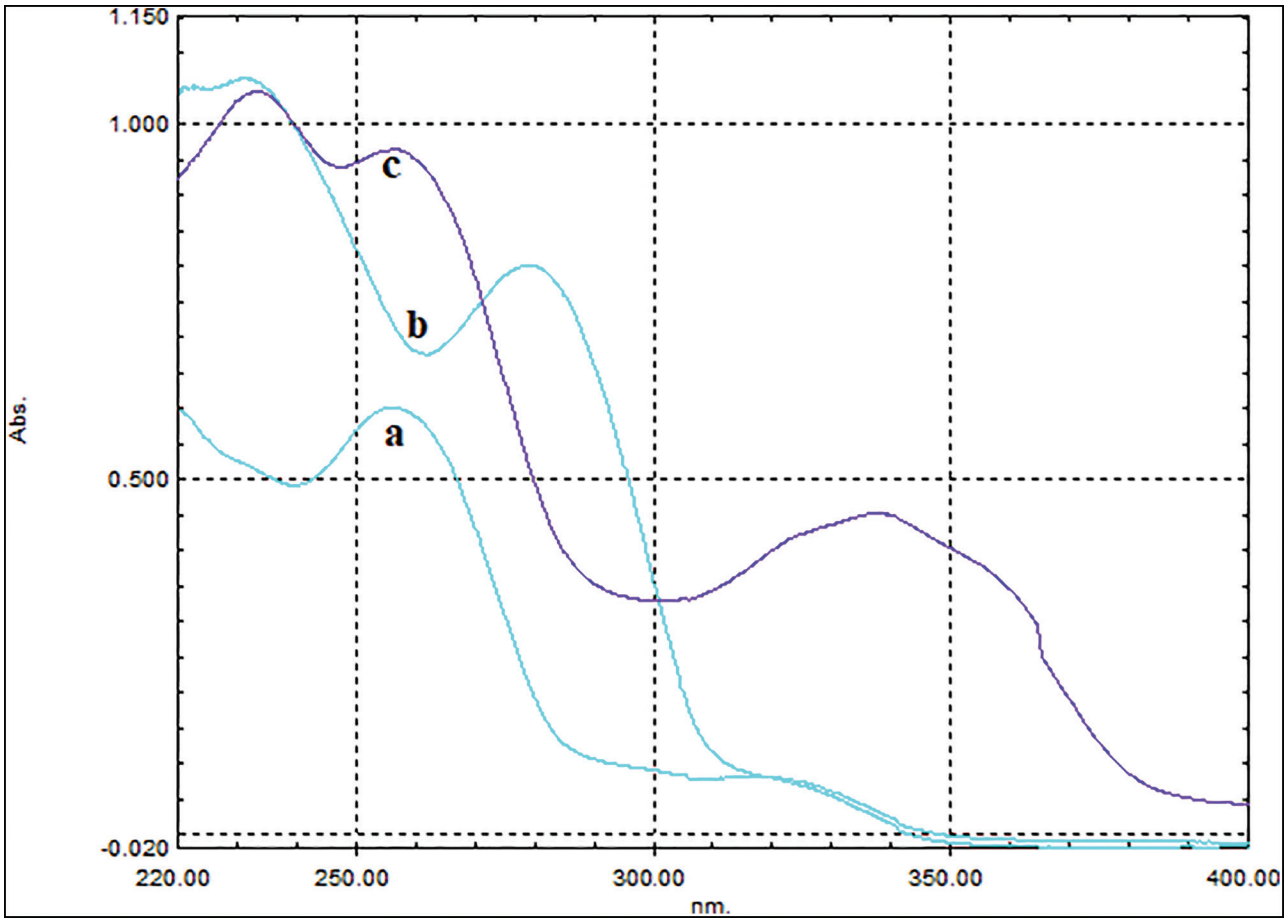

Figure 1: Overlay Spectra of 20.0 µg/mL solution of drugs; (a) tenofovir alafenamide fumarate, (b) emtricitabine, (c) dolutegravir sodium

iterative partial least square algorithm to generate the model. The independent and dependent variables are simultaneously compressed and decomposed, resulting in latent variables, in the PLS calibration using the orthogonalized PLS technique.

In Minitab 16.1.1, the A and C data matrix were incorporated in PCR and PLS models. The concentration of EMT, TEN, and DOL in the validation set and formulation were predicted. PLS and PCR calibrations were constructed by using the nonlinear iterative partial least squares (NIPALS) algorithm and standard singular value decomposition (SVD) algorithm, respectively. For PCR and PLS calibrations, an adequate number of principal components or factors must be chosen.

Validation of developed methods

Precision

Intraday and interday precision study was established by triplicate analysis of ternary mixture containing different proportions of EMT, TEN, and DOL (5/5/10 µg/mL, 10/20/5 µg/mL, and 10/20/10 µg/mL) on one day and on three successive days, respectively. The absorbance data of the ternary mixture obtained after scanning in UV spectrophotometer were incorporated in respective equations and concentrations were calculated. The results were expressed as a percent recovery ± standard deviation (SD).

Accuracy

The method's accuracy was determined by applying the analytical approach to fabricated blends of drug product components (placebo) to which known proportions of the drug ingredient to be analyzed were incorporated. Accuracy of the method was studied in triplicate at three various levels (80%, 100%, and 120%). The known amounts of standard solutions containing EMT (25.6, 32.0, and 38.4 µg/mL), TEN (3.2, 4.0, and 4.8 µg/mL), and DOL (6.4, 8.0, and 9.6 µg/mL), to achieve the various levels, were added to placebo sample solutions. The absorbance data of the ternary mixture obtained after scanning in UV spectrophotometer were incorporated in respective equations and concentrations were calculated. The results were expressed as a percent recovery ± SD.

Assay of formulation

Weigh 10 spegra tablets and determine the average content of blend. The tablet powder equivalent to 100 mg EMT, 26.3 mg DOL (≈ 25 mg dolutegravir), and 15.6 mg TEN (≈ 12.5 mg tenofovir alafenamide) was transferred to volumetric flask and dissolved in methanol by sonication for 20 min and the volume was made up to the mark with methanol. Filter paper No. 41 (Whatman) was used to filter the solution. Aliquot required amount from the above solution to achieve 32 µg/mL EMT,

8.4 µg/mL DOL, and 5 µg/mL TEN. At the specified wavelengths, the absorbance of the sample solutions was recorded, and the amount of individual component was measured.

Results and Discussion

Classical least square

The value of calibration coefficient can be calculated by using the equation:

$$K = \text{pinv}(c) \times A$$

where, pinv(c) is the pseudo inverse of concentration matrix and A is matrix of absorbance of mixture.

$$K_{cal} = \text{pinv}(K)$$

where, pinv(K) is pseudo inverse of K matrix concentration of unknown:

$$C = K_{cal} \times A$$

Spectra of solutions containing unknown concentrations of drugs were recorded in the optimized range of wavelength and absorbance matrix A were generated. Using the calibration coefficient matrix K, the concentration was computed.

$$\underset{\text{Conc matrix}}{\begin{bmatrix} C_{EMT} \\ C_{TEN} \\ C_{DOL} \end{bmatrix}} = \underset{\substack{\text{Kcal Matrix}^{\#} \\ EMT \quad TEN \quad DOL}}{\begin{bmatrix} 2.5817 & -3.0432 & 1.3716 \\ 1.912 & 1.583 & -0.7103 \\ 0.2503 & 7.4894 & -2.5124 \\ -1.7189 & 12.0367 & -3.1954 \\ -3.4584 & 13.9309 & -2.6006 \\ -4.065 & 13.5792 & -1.8539 \\ -3.2107 & 11.2118 & -1.4085 \\ -1.0937 & 6.2445 & -0.7398 \\ 1.8526 & 0.5876 & -0.5498 \\ 4.6662 & -5.1288 & -0.2128 \\ 6.836 & -9.6667 & 0.0777 \\ 7.7676 & -11.5561 & 0.0938 \\ 7.4444 & -11.235 & 0.1188 \\ 6.323 & -10.0885 & 0.4681 \\ 4.6298 & -8.8943 & 1.3534 \\ 2.5101 & -7.5833 & 2.5921 \\ 0.2221 & -6.6484 & 4.2212 \\ -1.4785 & -6.4007 & 5.6972 \\ -2.4641 & -6.5054 & 6.7231 \\ -3.0195 & -7.1512 & 7.6593 \\ -3.3878 & -8.1195 & 8.6026 \\ -3.669 & -9.2132 & 9.5248 \\ -3.8339 & -10.5526 & 10.4599 \\ -4.008 & -12.2572 & 11.6066 \\ -4.2149 & -13.9382 & 12.7694 \end{bmatrix}} \times \underset{\text{Abs matrix } (Mixture)}{\begin{bmatrix} A_{240} \\ A_{244} \\ A_{248} \\ A_{252} \\ A_{256} \\ A_{260} \\ A_{264} \\ A_{268} \\ A_{272} \\ A_{276} \\ A_{280} \\ A_{284} \\ A_{288} \\ A_{292} \\ A_{296} \\ A_{300} \\ A_{304} \\ A_{308} \\ A_{312} \\ A_{316} \\ A_{320} \\ A_{324} \\ A_{328} \\ A_{332} \\ A_{336} \end{bmatrix}}$$

[#]For representation of matrix conveniently, Kcal values are shown in transposed form.

Where *A* is the absorbance values at 25 points corresponding to the 240–336 nm spectral range at an interval of 4 nm. C_{EMT}, $C_{TEN,}$ and C_{DOL} represent the concentrations of EMT, TEN, and DOL, respectively.

Inverse least square

The value of calibration coefficient can be calculated by using the following equation:

$$P = \text{pinv}(A) \times C$$

Where, P is the matrix of the unknown calibration coefficients relating the concentrations to the spectral intensities. Spectra of solutions containing unknown concentrations of drugs mixture were recorded in the optimized range of wavelength and absorbance matrix A was generated. Using the calibration coefficient matrix P, the concentration was computed using the equation:

$$C = P \times A$$

$$\underset{\text{Conc matrix}}{\begin{bmatrix} C_{EMT} \\ C_{TEN} \\ C_{DOL} \end{bmatrix}} = \underset{\substack{P \text{ matrix} \\ EMT \quad TEN \quad DOL}}{\begin{bmatrix} 284.4683 & 40.2716 & 110.4284 \\ -244.7356 & -150.4681 & -98.9464 \\ 243.2853 & -88.8418 & -140.6718 \\ -39.4257 & 144.8259 & 72.1366 \\ -139.9503 & 66.0755 & 51.2964 \\ -505.1327 & 110.8386 & 112.3062 \\ 322.7066 & -39.7513 & 36.8206 \\ 79.2171 & -73.261 & -180.7986 \\ 198.853 & -81.6778 & -20.4717 \\ 323.0868 & 40.1419 & 34.6374 \\ -771.0162 & 43.4324 & -70.437 \\ -42.2148 & 24.783 & -29.6429 \\ 433.2921 & 20.8334 & 80.0165 \\ -447.1916 & 13.222 & 108.1144 \\ 247.1869 & -93.1218 & 46.8237 \\ 211.0631 & 3.6621 & -106.2136 \\ -162.7434 & 81.8123 & 90.8483 \\ -128.5652 & -21.5662 & -212.2727 \\ 359.2401 & 23.6529 & 21.4378 \\ -41.094 & 4.9311 & 74.593 \\ -327.1792 & 16.5191 & -39.0939 \\ 555.3401 & 2.1758 & -25.5125 \\ -433.031 & -25.9353 & -36.763 \\ -138.3896 & 27.1467 & 126.4679 \\ 109.6559 & -90.0582 & 34.8071 \end{bmatrix}} \times \underset{\text{Abs matrix } (Mixture)}{\begin{bmatrix} A_{240} \\ A_{244} \\ A_{248} \\ A_{252} \\ A_{256} \\ A_{260} \\ A_{264} \\ A_{268} \\ A_{272} \\ A_{276} \\ A_{280} \\ A_{284} \\ A_{288} \\ A_{292} \\ A_{296} \\ A_{300} \\ A_{304} \\ A_{308} \\ A_{312} \\ A_{316} \\ A_{320} \\ A_{324} \\ A_{328} \\ A_{332} \\ A_{336} \end{bmatrix}}$$

Where, *A* is the absorbance values at 25 points corresponding to the 240–336 nm spectral range at interval of 4 nm. C_{EMT}, $C_{TEN,}$ and C_{DOL} represent the concentrations of EMT, TEN, and DOL, respectively.

Partial least square and principle component regression

The component number for the experimental data should be chosen in such a way that overfitting is avoided. The number of principal components determined using the following approaches:

For PCR, two PCs were selected based on retaining components with eigenvalues greater than 1 (and retain components that cumulatively explain 90% of the variance.) and it was confirmed using scree plot [Figure 2]. Scree plot shows steep curve up to three PCs, followed by a bend and then a flat line. A number of PCs in PLS selected using a model selection plot; scatterplot of the cross-validated R^2 and fitted R^2 values as a function of the number of components [Figure 2]. Three numbers of components were selected based on retaining components with identical R^2 values of validated R^2 and fitted R^2. The components selected are also assessed using score plot. Here, the first two components make up the majority of the variance in the data, and there are no outliers in this data set; the points are spread randomly around zero.

The equations for the PLS method were obtained as:

C_{EMT} = 0.107+9.485 × **A1** + 28.863 × **A2** + 18.46 × **A3** − 11.686 × **A4** −11.264 × **A5** − 18.456 × **A6** −10.271 × **A7** + 3.293 × **A8** − 11.557 × **A9** − 9.722 × **A10** − 3.309 × **A11** − 4.031 × **A12** − 8.275 × **A13** − 5.134 × **A14** + 52.76 × **A15** − 6.275 × **A16** − 64.879 × **A17** + 126.418 × **A18** − 55.385× **A19** − 26.903 × **A20** + 33.011 × **A21** − 33.954× **A22** − 0.849× **A23** – 109.399 × **A24** + 89.03× **A25**

C_{TEN} = 0.035 – 36.553 × **A1** − 47.925× **A2** − 10.068× **A3** + 58.874× **A4** + 44.355× **A5** + 30.864× **A6** + 2.532 × **A7** − 29.872× **A8** − 32.414× **A9** + 11.905× **A10** +9.001× **A11** + 10.04× **A12** + 29.664× **A13** + 44.557× **A14** − 95.969× **A15** − 15.952× **A16** + 102.497× **A17** − 211.513 × **A18** + 80.741× **A19** − 61.178 × **A20** + 6.754× **A21** + 155.712× **A22** − 20.279 × **A23** – 111.113 × **A24** − 260.844 × **A25**

C_{DOL} = − 0.172+19.94 × **A1** − 10.77× **A2** − 27.725× **A3** − 9.676× **A4** − 3.541× **A5** + 24.839 × **A6** + 11.394 × **A7** − 9.1× **A8** + 1.65 × **A9** + 7.97 × **A10** + 0.778 × **A11** − 4.315 × **A12** + 4.747 × **A13** + 24.486 × **A14** + 0.708 × **A15** − 71.69 × **A16** + 150.532 × **A17** − 267.055× **A18** + 53.179 × **A19** + 155.129 × **A20** − 23.697 × **A21** − 168.71 × **A22** − 30.047× **A23** + 201.696× **A24** − 1.725× **A25**

The equations for the PCR method were obtained as:

C_{EMT} = − 1.128+1.519 × **A1** + 0.937 × **A2** − 0.215 × **A3** − 1.501 × **A4** −2.545 × **A5** − 2.937 × **A6** −2.406 × **A7** – 0.914 × **A8** + 1.271 × **A9** -+ 3.656 × **A10** + 5.860 × **A11** + 7.234 × **A12** + 7.563 × **A13** + 7.241 × **A14** + 6.397 × **A15** + 4.780 ×

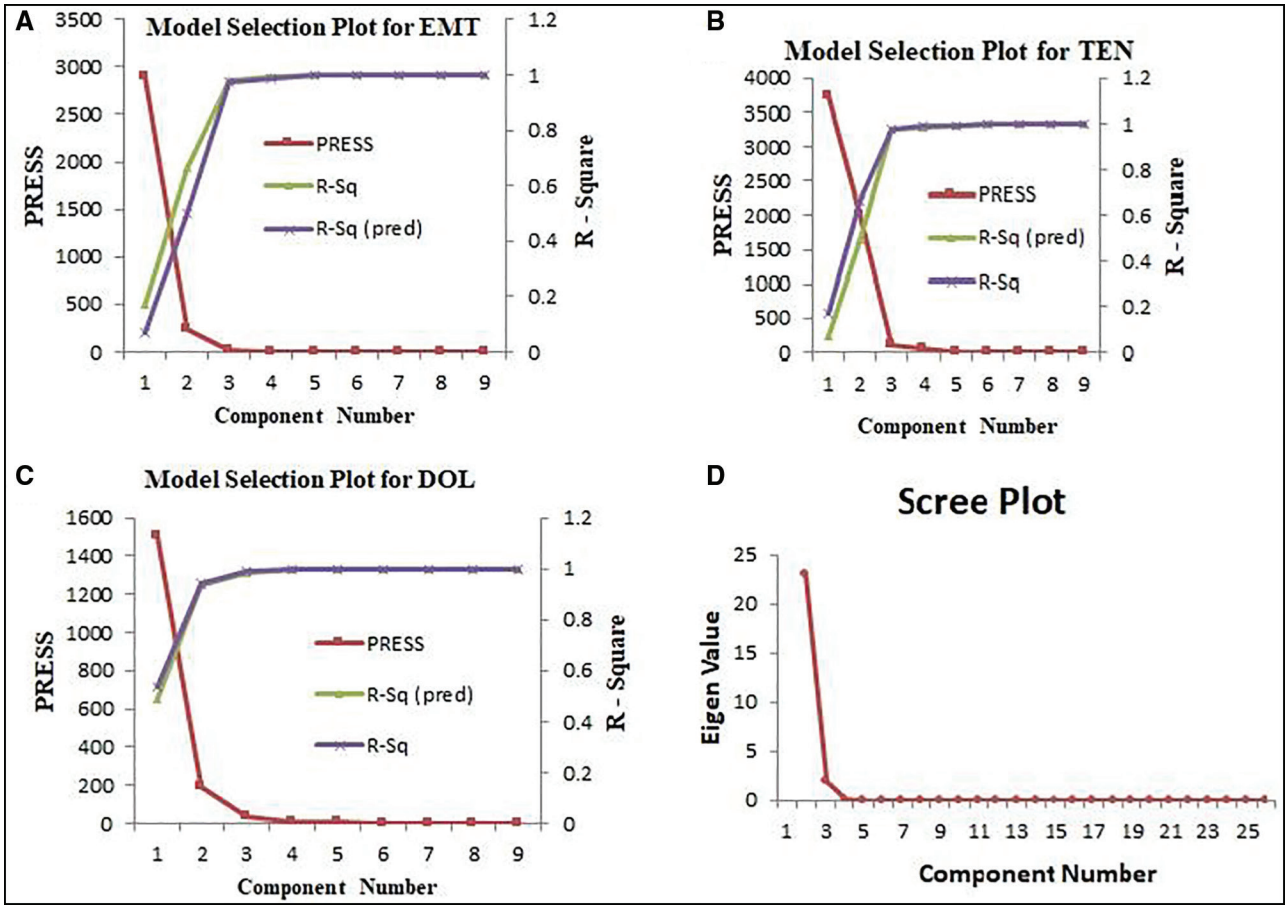

Figure 2: Plots for selection of number of component in PLS method for ternary mixture tenofovir alafenamide fumarate (TEN) (a), emtricitabine (EMT) (b), dolutegravir sodium (DOL) (c) and scree plot; for selection of component in PCR (d)

A16 + 2.080 × **A17** − 0.865 × **A18** − 2.814 × **A19** − 3.660 × **A20** - 3.981 × **A21** − 4.131× **A22** − 4.072 × **A23** – 3.995 × **A24** − 3.966 × **A25**

C_{TEN} = − 0.021+0.660 × **A1** + 2.985× **A2** + 6.505 × **A3** + 9.676× **A4** + 11.730 × **A5** + 12.157 × **A6** + 10.802 × **A7** + 7.162 × **A8** + 2.310 × **A9** − 3.218 × **A10** − 8.297 × **A11** − 11.407 × **A12** − 12.227 × **A13** − 11.941 × **A14** − 11.261 × **A15** − 9.945 × **A16** − 8.157 × **A17** − 6.689 × **A18** − 5.560× **A19** − 5.667 × **A20** − 6.437× **A21** − 7.693 × **A22** − 9.922 × **A23** – 12.885 × **A24** − 15.807 × **A25**

C_{DOL} = 2.687 − 0.298 × **A1** − 0.962 × **A2** − 1.717 × **A3** − 2.172× **A4** − 2.274× **A5** − 2.136 × **A6** − 1.924 × **A7** − 1.417 × **A8** − 0.927 × **A9** − 0.285 × **A10** + 0.283 × **A11** + 0.599 × **A12** + 0.713 × **A13** + 0.874 × **A14** + 1.317 × **A15** − 2.154 × **A16** + 3.740 × **A17** + 5.659 × **A18** + 6.827 × **A19** + 7.627 × **A20** + 8.297 × **A21** + 9.060 × **A22** + 10.125× **A23** + 11.545× **A24** + 12.980 × **A25**

Where, *A* is the absorbance values at 25 points corresponding to the 240–336 nm spectral range at intervals of 4 nm and $C_{EMT,}$ $C_{TEN,}$ and C_{DOL} are the concentrations of EMT, TEN, and DOL, respectively.

Method validation

Precision

Method reproducibility for each title ingredient was demonstrated by repeatability and intermediate precision measurements. The obtained results within and between days trials are represented in Tables 2 and 3. The recovery values were close to 100% with low SD justified the good precision of the proposed methods.

Accuracy

The mean percent recoveries for EMT, TEN, and DOL are reported in Table 4. The remarkable accuracy of the

Table 2: Precision data (intraday) by chemometric methods

Chemometric model	Ternary mixture (µg/mL) (EMT/TEN/DOL)	% Recovery, mean ± SD (*n* = 3),		
		EMT	TEN	DOL
CLS	1 (5/5/10)	99.6±0.801	100.2±1.709	99.74±0.709
	2 (10/20/5)	99.87±0.803	99.92±0.326	99.47±1.223
	3 (10/20/10)	99.94±0.809	100.07±0.417	100.27±0.851
ILS	1 (5/5/10)	99.87±1.617	99.4±1.201	99.84±0.751
	2 (10/20/5)	100.3±0.755	99.69±0.576	99.74±1.528
	3 (10/20/10)	99.84±0.612	100.09±0.476	99.5±0.755
PLS	1 (5/5/10)	100.54±1.007	99.94±1.528	99.8±0.529
	2 (10/20/5)	99.6±0.954	99.8±0.399	100.34±1.748
	3 (10/20/10)	100.37±1.041	100.12±0.276	100.67±0.908
PCR	1 (5/5/10)	99.54±1.102	99.6±1.778	100.14±1.202
	2 (10/20/5)	99.57±1.107	100.22±0.653	100.27±1.102
	3 (10/20/10)	100.24±0.951	100.09±0.389	100.57±1.151

EMT = emtricitabine, TEN = tenofovir alafenamide fumarate, DOL = dolutegravir sodium, CLS = classical least square, ILS = inverse least square, PLS = partial least square, PCR = principle component regression, SD = standard deviation

Table 3: Precision data (interday) by chemometric methods

Chemometric model	Ternary mixture (µg/mL) (EMT/TEN/DOL)	% Recovery, mean ± SD (*n* = 3),		
		ET	TEN	DOL
CLS	1 (5/5/10)	100.32±1.659	100.34±1.446	99.69±1.069
	2 (10/20/5)	99.94±0.809	99.91±0.563	100.38±1.509
	3 (10/20/10)	100.05±0.775	99.89±0.615	99.81±1.089
ILS	1 (5/5/10)	99.98±1.069	100.19±1.327	99.73±0.954
	2 (10/20/5)	99.62±0.968	99.66±0.412	99.89±1.201
	3 (10/20/10)	99.39±1.059	99.82±0.476	99.53±0.759
PLS	1 (5/5/10)	100.03±1.348	99.96±1.208	99.85±1.004
	2 (10/20/5)	99.77±1.036	100.04±0.539	100.14±1.229
	3 (10/20/10)	100.17±1.012	99.88±0.645	99.91±1.245
PCR	1 (5/5/10)	99.96±1.253	100.05±1.334	99.66±0.987
	2 (10/20/5)	100.14±0.945	99.91±0.539	100.47±1.158
	3 (10/20/10)	100.04±1.023	99.97±0.409	99.58±0.902

EMT = emtricitabine, TEN = tenofovir alafenamide fumarate, DOL = dolutegravir sodium, CLS = classical least square, ILS = inverse least square, PLS = partial least square, PCR = principle component regression, SD = standard deviation

recommended approaches was justified by the near-100% recovery values with low SD.

Chemometric methods

LOO approach was employed for cross validation technique using calibration set of 32 mixtures. Each calibration sample's predicted concentrations were tested to the known concentrations of compounds. RMSECV and RMSEP were calculated to validate the model. For a given model, these values must be as low as possible. For assessing the inaccuracies in the predicted concentrations, the RMSECV value was utilized as a screening test. It denotes the precision as well as the accuracy of predictions.

The prediction capabilities of developed methods (CLS, ILS, PCR, and PLS) are evaluated by using two different methods. Plotting the known concentration against the predicted concentration was the first approach used. The aforesaid chemometric procedures [Tables 5-7] yielded a reasonable correlation coefficient (R^2) value for each drug, and the second way was the calculation of RMSECV and RMSEP.

The analytical figure of merits (FOM) is critical for quantifying the quality of an approach or comparing methods. Several FOM has been observed in multivariate calibration, including sensitivity (SEN), analytical sensitivity, the limit of detection (LOD), and limit of quantitation (LOQ) [Tables 5-7].

Assay of formulation

EMT, TEN, and DOL in tablet formulations were assessed using the proposed chemometric approach. The results were satisfactory and in line with the label claim. The assay results [Table 8] show that the approach is acceptable for simultaneous quantification of EMT, TEN, and DOL without intervention from common excipients.

Table 4: Accuracy data of chemometric methods (CLS, ILS, PLS, and PCR)

Drug	Level (%)	Std. spiked (µg/mL)	% Recovery mean ± SD (n = 3)			
			CLS	ILS	PLS	PCR
EMT	80	25.6	99.99±0.314	99.99±0.216	100.11±0.239	100.07±0.294
	100	32	100.12±0.408	99.97±0.355	100.05±0.338	100.11±0.144
	120	38.4	100.14±0.209	100.09±0.289	100.19±0.291	100.08±0.105
TEN	80	3.2	99.48±1.302	100.53±0.955	99.48±1.302	99.48±0.955
	100	4	99.92±1.259	99.42±1.259	99.59±1.259	100.17±1.377
	120	4.8	100.56±1.148	99.94±1.273	100.56±1.148	99.94±1.273
DOL	80	6.4	100.21±0.942	100.42±0.651	99.64±0.549	99.59±0.861
	100	8	100.29±0.711	100.17±0.878	100.38±0.573	100.09±0.764
	120	9.6	100.08±0.684	100.18±0.637	100.39±0.47	99.69±0.478

EMT = emtricitabine, TEN = tenofovir alafenamide fumarate, DOL = dolutegravir sodium, CLS = classical least square, ILS = inverse least square, PLS = partial least square, PCR = principle component regression, SD = standard deviation

Table 5: Statistical parameters and figure of merits for emtricitabine

Component	Emtricitabine			
Parameters/model	CLS	ILS	PLS	PCR
RMSEC	0.1928	0.8965	0.2348	0.2597
RMSECV	0.1866	0.9965	0.1906	0.2235
RMSEP	0.9797	1.02923	1.1929	1.3193
R^2 Calibration	0.9998	0.9994	0.9998	0.9997
R^2 Prediction	0.995	0.993	0.999	0.996
Intercept	0.427	0.397	−0.019	0.225
Slope	0.992	0.976	0.998	0.999
PRESS	1.114	6.268	1.168	1.599
Sensitivity (SEN)	0.00858	0.0108	0.01641	0.01969
Selectivity (SEL)	0.0328	0.0418	0.0627	0.00752
LOD, µg/mL	0.222	0.196	0.153	0.121
LOQ, µg/mL	0.675	0.537	0.465	0.368
Analytical sensitivity (γ), mL/µg	13.5	14.7	19.5	24.7

CLS = classical least square, ILS = inverse least square, PLS = partial least square, PCR = principle component regression, RMSEC = root mean square error of calibration, RMSECV = root mean square error of cross validation, RMSEP = root mean square error of prediction, PRESS = predicted residual error sum of squares, LOD = limit of detection, LOQ = limit of quantitation

Table 6: statistical parameters and figure of merits for tenofovir alafenamide fumarate

Component	Tenofovir alafenamide fumarate			
Parameters/model	**CLS**	**ILS**	**PLS**	**PCR**
RMSEC	0.368	0.9136	0.412	0.462
RMSECV	0.433	1.13837	0.361	0.389
RMSEP	2.7919	1.175703	3.1251	3.5054
R^2 Calibration	0.9985	0.9982	0.9989	0.9987
R^2 Prediction	0.992	0.991	0.999	0.991
Intercept	0.657	-0.185	−0.025	0.597
Slope	0.964	1.044	1.001	0.965
PRESS	6.018	8.639	4.181	4.844
Sensitivity (SEN)	0.00345	0.0053	0.0067	0.0071
Selectivity (SEL)	0.0103	0.039	0.0203	0.0213
LOD, µg/mL	0.616	0.569	0.3637	0.3369
LOQ, µg/mL	1.87	1.29	1.1	1.02
Analytical sensitivity (γ), mL/µg	4.87	5.38	8.25	8.91

CLS = classical least square, ILS = inverse least square, PLS = partial least square, PCR = principle component regression, RMSEC = root mean square error of calibration, RMSECV = root mean square error of cross validation, RMSEP = root mean square error of prediction, PRESS = predicted residual error sum of squares, LOD = limit of detection, LOQ = limit of quantitation

Table 7: Statistical parameters and figure of merits for dolutegravir sodium

Component	Dolutegravir sodium			
Parameters/model	**CLS**	**ILS**	**PLS**	**PCR**
RMSEC	0.582	0.976	0.441	0.492
RMSECV	0.452	1.284	0.581	0.644
RMSEP	3.963	1.326	3.003	3.353
R^2 Calibration	0.9997	0.9988	0.9983	0.9994
R^2 Prediction	0.996	0.982	0.999	0.987
Intercept	0.627	0.663	0.076	−0.186
Slope	1.007	0.991	1.001	0.983
PRESS	6.555	15.364	10.819	13.31
Sensitivity (SEN)	0.00246	0.0068	0.00744	0.009179
Selectivity (SEL)	0.00692	0.0156	0.02094	0.0258
LOD, µg/mL	0.847	0.734	0.327	0.26
LOQ, µg/mL	2.57	1.69	0.991	0.79
Analytical sensitivity (γ), ml/µg	3.59	7.86	9.17	11.15

CLS = classical least square, ILS = inverse least square, PLS = partial least square, PCR = principle component regression, RMSEC = root mean square error of calibration, RMSECV = root mean square error of cross validation, RMSEP = root mean square error of prediction, PRESS = predicted residual error sum of squares, LOD = limit of detection, LOQ = limit of quantitation

Table 8: Content of emtricitabine, tenofovir alafenamide fumarate and dolutegravir sodium by chemometrics method

Drug		EMT	TEN	DOL
Label claim (mg in tablet)		200	25	50
% Label claim, mean ± SD, (n = 3)	CLS	100.09±0.284	100.81±0.835	100.38±0.819
	ILS	100.02±0.239	99.67±1.011	99.71±0.711
	PLS	100.08±0.141	100.42±1.128	100.34±0.506
	PCR	99.94±0.256	99.75±1.392	99.84±1.003

EMT = emtricitabine, TEN = tenofovir alafenamide fumarate, DOL = dolutegravir sodium, CLS = classical least square, ILS = inverse least square, PLS = partial least square, PCR = principle component regression, SD = standard deviation

Conclusion

The chemometric method is more accurate and precise than conventional methods as the total absorbance of the ternary mixture was measured. The developed method holds an acceptable degree of precision and accuracy in accordance with international guidelines. With great recoveries and precision, the proposed approach was successfully used to the assay of formulation. As a result, the current method can be used to estimate EMT, TEN, and DOL in formulation simultaneously.

Acknowledgement

The authors express gratitude towards Bulat Pharmaceuticals, Amneal Pharmaceuticals, Cipla Pharmaceuticals for gifting

drug samples. The authors also acknowledge APMC College of Pharmaceutical Education and Research, Himatnagar (Managed by Himatnagar Kelvani Mandal) for providing facilities to conduct this research.

Financial support and sponsorship

Nil.

Conflicts of interest

I/we certify that no actual or potential conflict of interest in relation to this article exists.

References

1. Kramer R. Chemometric Techniques for Quantitative Analysis. New York: Marcel Dekker Inc; 1998.
2. Beebe KR, Pell RJ, Seasholtz MB, Chemometrics: A Practical Guide. 1st ed. New York:Wiley Interscience; 1998.
3. Miller JN, Miller JC. Statistics and Chemometrics for Analytical Chemistry. 6th ed. Harlow: Pearson Education Limited; 2005.
4. Goicoechea HC, Olivieri AC. Simultaneous determination of rifampicin, isoniazid and pyrazinamide in tablet preparations by multivariate spectrophotometric calibration. J Pharm Biomed Anal 1999;20:681-6.
5. Patel NS, Nandurbarkar VP, Patel AJ, Patel SG. Simultaneous spectrophotometric determination of celecoxib and diacerein in bulk and capsule by absorption correction method and chemometric methods. Spectrochim Acta A Mol Biomol Spectrosc 2014;125:46-52.
6. Nejem RA, Issa MM, Shanab AA, Shat NA. Double divisor mean centering of ratio spectra as a developed spectrophotometric method for the analysis of five-component mixture in drug analysis. J Saudi Chem Soc 2017;21:S283-92.
7. Rathod SM, Patel PU. Chemometrics assisted spectroscopic method development and validation for simultaneous determination of sofosbuvir and daclatasvir dihydrochloride in tablet formulation. Indian Drugs 2020;57:37-46.
8. Patel NC, Patel AP, Patel JK. Development and validation of chemometrics assisted UV spectrophotometric method for epigallocatechin gallate and curcumin in tablet formulation. Indian Drugs 2020;57:45-54.
9. Sankar ASK, Vetrichelvan T, Venkappaya D, Divya O. Simultaneous estimation of ramipril, aspirin and atorvastatin calcium by classical least squares regression in capsule dosage form. Res J Pharm Technol 2011;4:398-401.
10. Patel A, Gohel M, Soni T. Partial least square analysis and mixture design for the study of the influence of composition variables on nanoemulsions as drug carriers. Res J Pharm Technol 2014;7:1446-55.
11. Pathi PJ, Reddy PR, Raju NA. Visible spectrophotometric estimation of emtricitabine in pharmaceutical formulations. Res J Pharm Technol 2011;4:437-40.
12. Halde S, Mungantiwar A, Chintamaneni M. Determination of emtricitabine in human plasma by LC-MS/MS. Res J Pharm Technol 2012;5:133-7.
13. Shah AU, Kotadiya VV, Ajmera AA. Analytical method development and validation for simultaneous estimation of emtricitabine and tenofovir disoproxil fumarate in tablet dosage form. Res J Pharm Technol 2016;9:463-8.
14. Narendra A, Deepika D, Annapurna M. Determination of emtricitabine by derivative spectrophotometric method in pharmaceutical dosage forms. Chem Sci Trans 2013;2:978-82.
15. Aggarwal NN, Bhat KI, Jacob JT. Stability indicating assay method development and validation for tenofovir alafenamide fumarate by RP-HPLC. Pharm Anal Acta 2018;9:1-6.
16. Bhavyasri K, Manisha M. UV-spectrophotometric estimation and forced degradation studies of tenofovir alafenamide fumarate (TAF) in its bulk and tablet dosage form. Res J Pharm Technol 2020;13:1231-5.
17. Mastanamma SK, Saidulu P, Srilakshmi B, Ramadevi N, Prathyusha D, Vidya Rani M. Analytical quality by design approach for the development of UV-spectophotometric method in the estimation of tenofovir alafenamide in bulk and its laboratory synthetic mixture. Res J Pharm Technol 2018;11:499-503.
18. Rathod SM, Patel PU. Development and validation of RP–HPLC method for estimation of lamivudine and dolutegravir sodium in synthetic mixture. Res J Pharm Technol 2020;3:2864-8.
19. Devanna N, Dudekula B, Ramachandraiah C. Development and validation for the simultaneous estimation of dolutegravir and lamivudine in drug product by RP-HPLC. Int J Res Pharm Nano Sci 2017;6:173-80.