Int. J. Pharm. Investigation, 2027; 17(1):381-400.
https://www.jpionline.org

Original Article

Chemometric-Assisted ATR-FTIR Spectroscopy for Rapid Authentication and Quality Evaluation of Marketed Antidiabetic Tablet Formulations

Pranay Kumar Ravula, Lakshmi Manogna Parupudi, Pravalika Pottigari, Jahnavi Bandla*

Department of Pharmacy, Vishnu Institute of Pharmaceutical Education and Research, Narsapur, Tuljaraopet, Telangana, INDIA.

ABSTRACT

Background: The widespread availability of branded and generic antidiabetic tablet formulations has increased the need for rapid, reliable, and cost-effective analytical methods to ensure their authenticity and quality. Conventional quality evaluation techniques are often time-consuming, destructive, and require extensive sample preparation. Chemometric-Assisted Attenuated total Reflectance Fourier Transform Infrared (ATR-FTIR) spectroscopy offers a rapid, non-destructive, and environmentally friendly approach for the authentication and quality assessment of pharmaceutical formulations. **Objectives:** The present study aimed to develop a rapid, non-destructive, and environmentally sustainable analytical approach for the authentication and quality evaluation of marketed antidiabetic tablet formulations using ATR-FTIR spectroscopy coupled with chemometric techniques. **Materials and Methods:** Commercially available branded and generic formulations containing vildagliptin and metformin, teneligliptin and metformin, and dapagliflozin and metformin combinations were analyzed. ATR-FTIR spectra were acquired over the range of 4000-400 cm^{-1} and subjected to spectral preprocessing including baseline correction, Savitzky-Golay smoothing, standard normal variate transformation, normalization, derivative processing, and mean centering. Principal Component Analysis (PCA), Hierarchical Cluster Analysis (HCA), and Partial Least Squares-Discriminant Analysis (PLS-DA) were employed for classification and authentication of formulations. **Results:** PCA demonstrated effective discrimination among formulations, with cumulative variances of 93.5%, 79.5%, and 94.3% for vildagliptin and metformin, teneligliptin and metformin, and dapagliflozin and metformin products, respectively. HCA generated distinct clustering patterns corresponding to branded and generic products, particularly for vildagliptin and metformin formulations. PLS-DA models exhibited excellent classification performance with high accuracy, sensitivity, and specificity. **Conclusion:** The findings indicate that ATR-FTIR spectroscopy combined with chemometric modeling provides a rapid, cost-effective, and green analytical platform for pharmaceutical authentication and quality assessment. The developed approach may serve as a reliable alternative to conventional chromatographic techniques for routine quality control and counterfeit detection.

Keywords: ATR-FTIR, Chemometrics, HCA, PCA, Pharmaceutical authentication, PLS-DA.

Correspondence:
Dr. Jahnavi Bandla
Department of Pharmacy, Vishnu Institute of Pharmaceutical Education and Research, Narsapur, Tuljaraopet-502313, Telangana, INDIA.
Email: jahnavi.bandla@gmail.com
ORCID: 0000-0002-0334-5501

Received: 21-07-2026;
Revised: 11-08-2026;
Accepted: 03-09-2026.

INTRODUCTION

Diabetes Mellitus (DM) is a chronic metabolic disorder marked by persistent hyperglycaemia, arising from impaired insulin secretion, impaired insulin action, or both. It is now among the fastest-growing health concerns worldwide, with a heavy toll on morbidity, mortality, and healthcare spending. Chronic hyperglycaemia drives microvascular complications such as nephropathy, retinopathy, and neuropathy, along with macrovascular disease including coronary artery disease and stroke. Type 2 Diabetes Mellitus (T2DM) accounts for roughly 90-95% of all diabetes cases and results from progressive insulin resistance and pancreatic β-cell dysfunction, which is why most patients require long-term pharmacological treatment to control blood glucose and limit complications (Chaudhury *et al.*, 2017; Galicia-Garcia *et al.*, 2020; WHO, 2024a).

T2DM treatment has shifted toward oral agents that work through complementary mechanisms. Fixed-dose combinations of metformin with Dipeptidyl Peptidase-4 (DPP-4) inhibitors such as vildagliptin and teneligliptin, or with the Sodium-Glucose Cotransporter-2 (SGLT2) inhibitor dapagliflozin, are now widely prescribed because they improve efficacy and make it easier for patients to stay on treatment. Generic versions have made these combinations far more accessible, but differences

DOI: 10.5530/ijpi.20270240

Copyright Information :
Copyright Author (s) 2027 Distributed under Creative Commons CC-BY 4.0

Publishing Partner : Manuscript Technomedia. [www.mstechnomedia.com]

in manufacturing, excipients, storage, and production methods can still alter a product's physicochemical properties. That is why rapid, reliable, and environmentally sustainable methods for authenticating and evaluating marketed formulations matter (USFDA, 2021; WHO, 2024b).

Pharmaceutical quality control still relies mainly on chromatographic techniques such as High-Performance Liquid Chromatography (HPLC), Gas Chromatography (GC), and Liquid Chromatography-Mass Spectrometry (LC-MS). These methods are sensitive and accurate, but they demand extensive sample preparation, expensive instrumentation, trained personnel, and large volumes of hazardous solvents. That makes them poorly suited to rapid screening or high-throughput analysis, and has pushed the field toward simpler, faster, non-destructive, and greener alternatives (Stuart, 2004; Bunaciu *et al.*, 2010; Magalhães *et al.*, 2021).

Among spectroscopic techniques, Fourier Transform Infrared (FTIR) spectroscopy stands out for pharmaceutical characterization because it captures a molecular fingerprint based on how chemical bonds vibrate. It can identify functional groups, characterize active pharmaceutical ingredients, check drug-excipient compatibility, and assess formulation integrity with little or no sample preparation. Attenuated total reflectance Fourier transform infrared (ATR-FTIR) in particular has become popular in pharmaceutical analysis for its simplicity, reproducibility, speed, and fit with green analytical chemistry. Because ATR-FTIR spectra are so reproducible, the technique works well for routine quality control and counterfeit drug detection (Bunaciu *et al.*, 2010; Magalhães *et al.*, 2021; Bro & Smilde, 2014; Brereton & Lloyd, 2014; Rohman & Che Man, 2012; Sherif *et al.*, 2020).

FTIR spectra, though, contain hundreds of highly correlated variables that conventional univariate statistics cannot interpret well. Chemometrics gets around this by combining multivariate statistics, mathematics, and computation to pull meaningful information out of complex analytical datasets. Unsupervised methods such as Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA) reveal similarities and natural groupings among samples, while supervised techniques such as Partial Least Squares-Discriminant Analysis (PLS-DA) sharpen classification accuracy and point to which spectral variables actually drive sample discrimination. Pairing FTIR with chemometrics has become a standard approach for pharmaceutical authentication, formulation classification, adulteration detection, and quality evaluation (Fatmarahmi *et al.*, 2022; Ahmad, 2025; Smith, 2011; Griffiths & de Haseth, 2007; Skoog *et al.*, 2018; Brereton, 2018; Beebe *et al.*, 1998).

FTIR and ATR-FTIR coupled with chemometrics have already been used to authenticate pharmaceutical formulations, herbal medicines, edible oils, and other complex matrices. PCA, HCA, PLS regression, and PLS-DA classify samples quickly and accurately while cutting down on sample preparation and solvent use. These spectroscopic methods generally agree well with conventional chromatographic results and fit naturally within green analytical chemistry. Portable instrumentation has also pushed the technique toward real-time quality monitoring and counterfeit detection. Even so, few studies have compared and classified commercially available fixed-dose antidiabetic formulations that combine metformin with vildagliptin, teneligliptin, or dapagliflozin using an integrated ATR-FTIR-chemometric approach. That gap matters because rapid discrimination between marketed formulations can meaningfully strengthen pharmaceutical quality assurance and regulatory oversight (Otto, 2017; Miller & Miller, 2018; Eriksson *et al.*, 2013; Wold *et al.*, 2001; Yuliani *et al.*, 2018; Rohman *et al.*, 2017; Darbandi & Sohrabi, 2024; Attia *et al.*, 2022; Palur *et al.*, 2020; Madbouly *et al.*, 2024; Pruksapha *et al.*, 2023; Tiernan *et al.*, 2020).

This study set out to develop and evaluate a rapid, non-destructive, and environmentally sustainable strategy, based on ATR-FTIR spectroscopy coupled with chemometrics, for authenticating and classifying commercially available antidiabetic tablet formulations. PCA, HCA, and PLS-DA were used to examine spectral variability, classify the formulations, and pinpoint which spectral regions drove the discrimination among the different antidiabetic combinations. The goal was a method that offers a fast, affordable, and reliable alternative for routine pharmaceutical quality assessment and counterfeit drug detection (Figure 1).

Pairing vibrational spectroscopy with chemometrics has become a practical way to assess pharmaceutical quality without the cost and solvent use of chromatography, and ATR-FTIR combined with PCA, HCA, PLSR, and PLS-DA has already been applied to pharmaceuticals, herbal products, and food matrices with good predictive accuracy. Few studies, however, have looked specifically at fixed-dose antidiabetic formulations combining metformin with vildagliptin, teneligliptin, or dapagliflozin using this combined approach - the gap this study addresses (Haaren *et al.*, 2023; Beć *et al.*, 2021; Pasquini, 2018; Deidda *et al.*, 2019).

Therefore, the present study aimed to investigate the applicability of ATR-FTIR spectroscopy combined with multivariate chemometric techniques for rapid authentication, classification, and quality evaluation of marketed antidiabetic tablet formulations containing vildagliptin and metformin, teneligliptin and metformin, and dapagliflozin and metformin.

MATERIALS AND METHODS

Materials

Marketed antidiabetic tablet formulations containing vildagliptin (50 mg) + metformin (500 mg), teneligliptin (20 mg) + metformin (500 mg), and dapagliflozin (10 mg) + metformin (1000 mg) were

procured from licensed pharmacies. Both branded and generic formulations were included for comparative evaluation.

Instrumentation

ATR-FTIR spectra were acquired using a Nicolet iS50 FTIR spectrophotometer (Thermo Fisher Scientific, USA) equipped with a diamond ATR accessory and OMNIC software. Spectra were recorded over the range of 4000-400 cm^{-1} at a resolution of 4 cm^{-1} using 10 scans per sample.

Ethical Statement

Ethical approval was not required for this study because it involved only the analysis of commercially available pharmaceutical tablet formulations without the involvement of human participants, animals, or patient-derived materials.

Sample Preparation

Twenty tablets from each formulation were randomly selected and finely powdered. The powders were homogenized and passed through a 60-mesh sieve. Approximately 10-20 mg of sample was placed directly on the ATR crystal and analyzed without additional sample preparation.

Spectral Acquisition

A background spectrum was collected before each analysis. Triplicate spectra were recorded for every formulation and averaged to improve reproducibility. Spectral data were exported for subsequent chemometric analysis.

Spectral Preprocessing

To minimize instrumental and sample-related variations, spectra were subjected to baseline correction, Savitzky-Golay smoothing, standard normal variate transformation, min-max normalization, derivative processing, and mean centering.

Chemometric Analysis

PCA was performed to reduce dimensionality and visualize clustering patterns among formulations. HCA was employed to evaluate similarities among formulations using Euclidean distance and Ward's linkage method. PLS-DA was used as a supervised classification technique for discrimination of branded and generic formulations.

Model Validation

FTIR spectral method was validated using specificity, precision, correlation coefficient (r), robustness and reproducibility. PLS-DA models were validated using leave-one-out cross-validation. Performance was assessed using accuracy, sensitivity, specificity, precision, F1-score, confusion matrix.

Statistical Analysis

All FTIR spectra were preprocessed using baseline correction and vector normalization before chemometric analysis. PCA, HCA, and PLS-DA were performed using Unscrambler X software. Leave-One-Out Cross-Validation (LOOCV) was employed for model validation. Model performance was assessed using sensitivity, specificity, precision, F1-score, and classification accuracy.

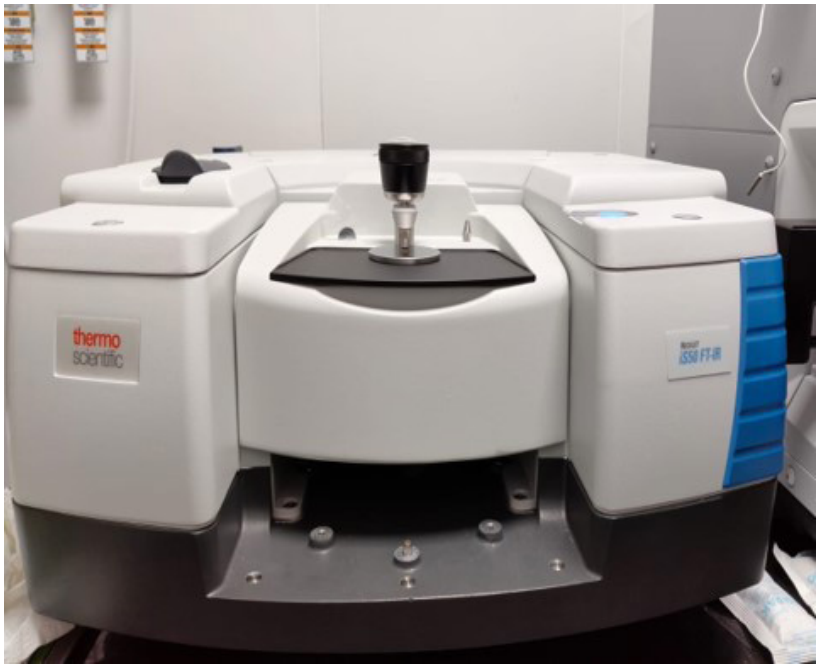


Figure 1: FTIR Instrumentation. shows the FTIR instrument used in this work.

RESULTS

ATR-FTIR Spectral Characterization

Figures 2-4 illustrate the stacked (a) and overlay (b) ATR-FTIR spectra of all tablet formulations recorded over the spectral range of 4000-400 cm^{-1}.

The spectra exhibited characteristic absorption bands corresponding to both Active Pharmaceutical Ingredients (APIs) and excipients.

Principal Component Analysis

PCA was performed on FTIR spectral data obtained from marketed tablet formulations to investigate similarities and differences among branded and generic products. Figures 5-7 present the PCA results, including the scree plot (a), loading plots (b), and score plots (c) of three marketed antidiabetic tablet formulations.

Hierarchical Cluster Analysis

HCA was performed on the FTIR spectral data of marketed tablet formulations using Ward's linkage method and it consists of HCA dendrogram and HCA heatmap analysis. Figures 8-10 present the HCA dendrograms (a) and corresponding heatmaps (b) for the three marketed antidiabetic tablet formulations evaluated in this study.

Partial Least Squares-Discriminant Analysis

Partial Least Squares-Discriminant Analysis (PLS-DA) was applied to the FTIR spectral dataset of branded and generic tablet formulations to evaluate their classification and identify the spectral regions responsible for discrimination. Figures 11-13 illustrate the PLS-DA results for the three marketed antidiabetic tablet formulations. Each Figure comprises the PLS-DA score plot (a), LOOCV confusion matrix (b), Variable Importance in Projection (VIP) score spectrum (c), and the top discriminating wavenumber bands (d), respectively.

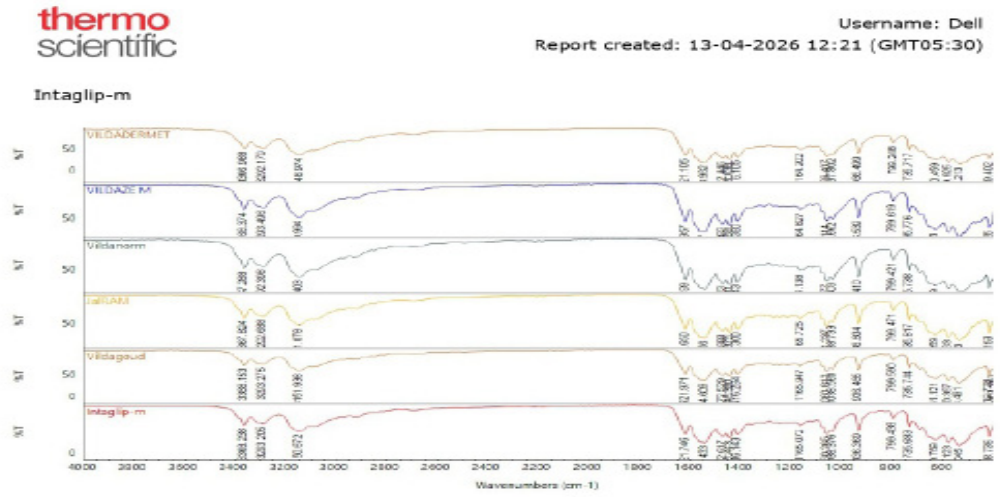


Figure 2a: Stacked FTIR spectrum of vildagliptin + metformin (50 mg/500 mg) tablets.

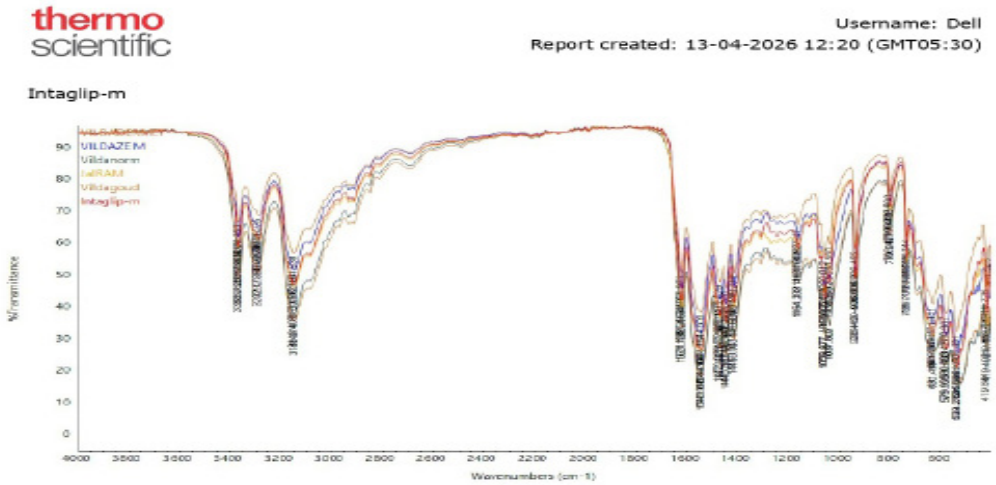


Figure 2b: Overlay FTIR spectrum of vildagliptin + metformin (50 mg/500 mg) tablets.

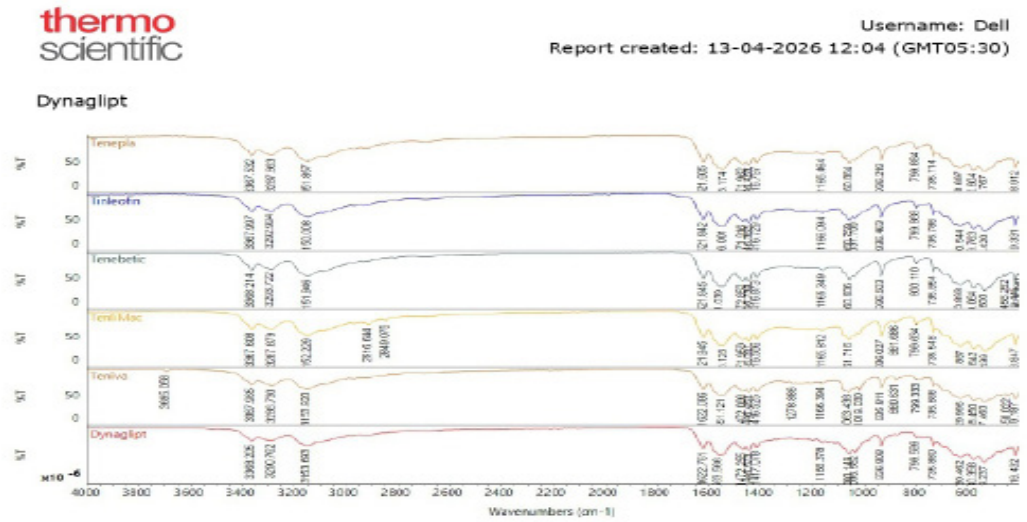


Figure 3a: Stacked FTIR spectrum of teneligliptin + metformin (20 mg/500 mg) tablets.

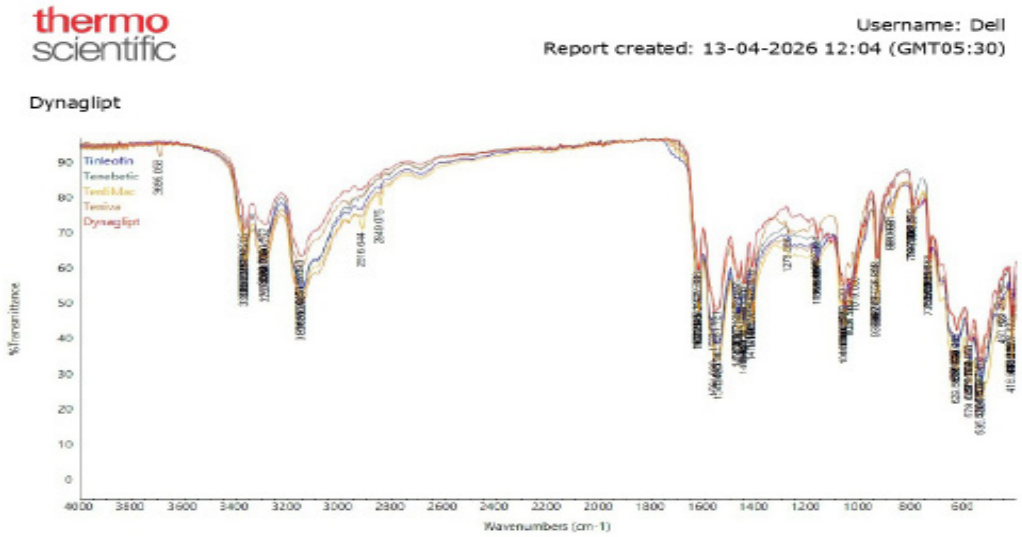


Figure 3b: Overlay FTIR spectrum of teneligliptin + metformin (20 mg/500 mg) tablets.

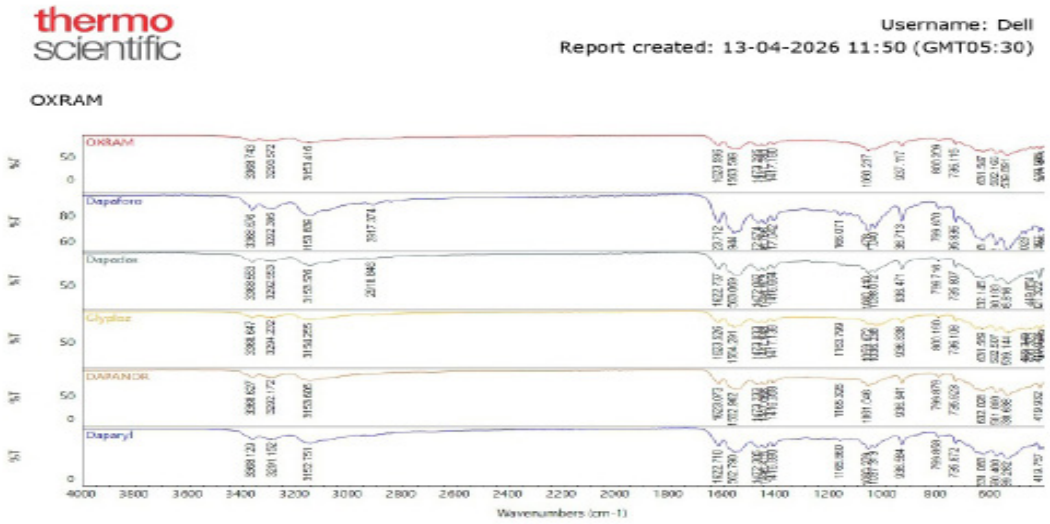


Figure 4a: Stacked FTIR spectrum of dapagliflozin + metformin (10 mg/1000 mg) tablets.

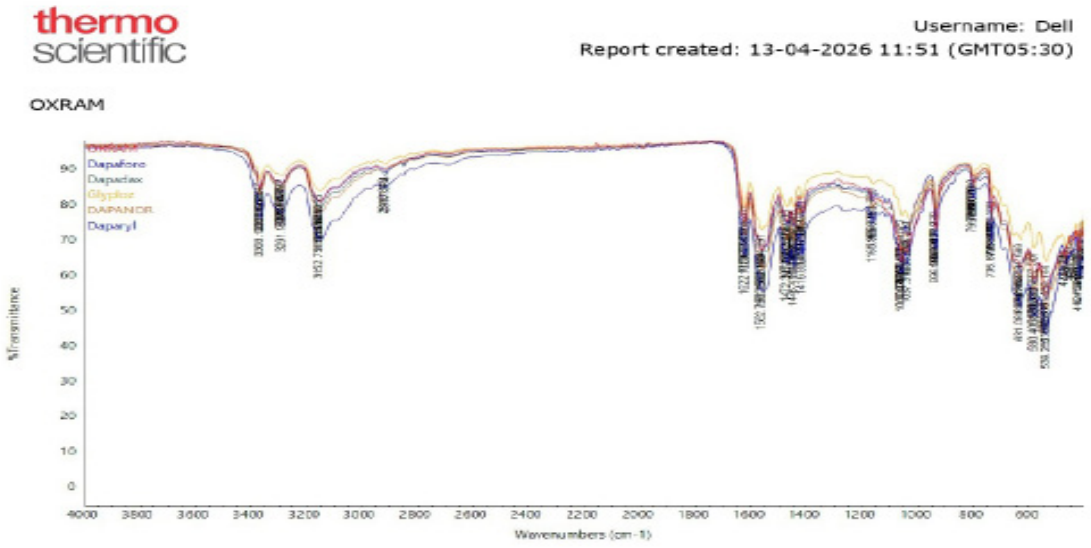


Figure 4b: Overlay FTIR spectrum of dapagliflozin + metformin (10 mg/1000 mg) tablets.

Validation of FTIR method - Chemometric Models

Validation confirms that the developed FTIR-chemometric method is reliable, reproducible, and suitable for pharmaceutical quality analysis. For FTIR coupled with PCA/PLS, validation includes: Spectroscopic method validation and Chemometric model validation.

The FTIR method was validated according to International Council for Harmonization (ICH) Q2(R^2) principles adapted for spectroscopic fingerprinting and chemometric authentication and results were summarized in the Table 1.

PCA Model Validation results were shown in Table 2.

HCA Model Validation results were summarized in Table 3.

PLS-DA Model Validation results were shown in Table 4.

DISCUSSION

The ATR crystal surface was cleaned with methanol and lint-free tissue before each measurement. A background spectrum was recorded before sample analysis. Approximately 10-20 mg of powdered tablet sample was placed on the ATR crystal. Uniform pressure was applied using the pressure arm. Spectra were recorded in transmittance mode. Three replicate spectra were obtained for each formulation to ensure reproducibility. The average spectrum was used for subsequent chemometric analysis. After each measurement, the ATR crystal was cleaned thoroughly to prevent cross-contamination. The acquired spectra were exported in suitable digital formats for preprocessing and chemometric evaluation.

The stacked and overlay ATR-FTIR spectra of marketed dapagliflozin-metformin, teneligliptin-metformin, and vildagliptin-metformin tablet formulations demonstrated highly comparable spectral profiles, indicating the presence of common active pharmaceutical ingredients and similar excipient matrices across both branded and generic products. Characteristic absorption bands were consistently observed in the region of 3380-3150 cm^{-1}, corresponding to O-H and N-H stretching vibrations associated with metformin and hydrogen-bonded functional groups. Distinct bands around 1620-1560 cm^{-1} were attributed to C=N and N-H bending vibrations of the guanidine moiety of metformin, while peaks in the 1470-1410 cm^{-1} region corresponded to C-N stretching and C-H bending vibrations. Strong absorptions observed between 1165 and 1000 cm^{-1} were assigned to C-O-C, C-N, and C-O stretching vibrations characteristic of dapagliflozin, teneligliptin, vildagliptin, and formulation excipients. The fingerprint region (1000-400 cm^{-1}) exhibited minor variations in peak intensity and position among different marketed products, reflecting differences in excipient composition, manufacturing processes, and formulation characteristics. Overlay spectra revealed substantial spectral overlap with preservation of all major functional group absorptions, confirming the chemical integrity and compatibility of the active ingredients within each formulation category. The stacked spectra further highlighted subtle inter-brand variations, particularly in the fingerprint region, which provide discriminatory information suitable for multivariate chemometric analysis. Overall, the FTIR spectral patterns confirmed the presence of the intended drug combinations and demonstrated sufficient spectral similarity for pharmaceutical equivalence assessment while retaining subtle differences that can be exploited for authentication and quality evaluation of branded and generic antidiabetic tablets.

Principal Component Analysis

PCA of the ATR-FTIR spectral datasets effectively differentiated branded and generic formulations of dapagliflozin-metformin,

teneligliptin-metformin, and vildagliptin-metformin tablets while preserving the underlying chemical similarities among products. For dapagliflozin-metformin formulations, Principal Components (PC) 1 and PC2 accounted for 87.8% and 6.5% of the total variance, respectively (94.3% cumulative variance), indicating that most spectral variability was captured by the first principal component. The PCA score plot revealed partial clustering of branded and generic products, with separation primarily driven by variations in the fingerprint region (1500-400 cm^{-1}), as confirmed by the loading plots. In teneligliptin-metformin formulations, PC1 and PC2 explained 46.2% and 33.3% of the variance, respectively (79.5% cumulative variance), demonstrating a more distributed variance structure and clearer discrimination between branded and generic products along PC1. The corresponding loading plots indicated that spectral differences in the regions associated with N-H/O-H stretching (3400-3000 cm^{-1}) and fingerprint vibrations (1600-400 cm^{-1}) contributed significantly to sample separation. For vildagliptin-metformin formulations, PC1 and PC2 accounted for 73.2% and 20.3% of the variance, respectively (93.5% cumulative variance), producing distinct clustering of branded and generic formulations with minimal overlap. The loading plots highlighted the fingerprint region, particularly 1200-400 cm^{-1}, as the major source of discrimination. Overall, the PCA models demonstrated excellent dimensionality reduction and revealed formulation-dependent spectral variations attributable to differences in excipient composition, manufacturing processes, and tablet matrix characteristics, while confirming the overall

Table 1: Summary of Method Validation Parameter Results.

Parameter	Teneligliptin + Metformin	Vildagliptin + Metformin	Dapagliflozin + Metformin	Acceptance Criteria	Interpretation
Spectral Range	4000-400 cm^{-1}	4000-400 cm^{-1}	4000-400 cm^{-1}	Fixed	Complied
Specificity	Distinct FTIR fingerprints observed	Distinct FTIR fingerprints observed	Distinct FTIR fingerprints observed	No spectral interference	Passed
Repeatability (%RSD)	0.82	0.75	0.80	<2%	Passed
Intermediate Precision (%RSD)	1.15	1.10	1.20	<3%	Passed
Instrument Precision (%RSD)	0.65	0.62	0.68	<2%	Passed
Spectral Correlation Coefficient (r)	0.9994	0.9990	0.9998	>0.99	Passed
Robustness	No significant variation observed	No significant variation observed	No significant variation observed	Consistent spectra	Passed
Fingerprint Reproducibility	Excellent	Excellent	Excellent	Consistent spectral pattern	Passed

Table 2: PCA Model Validation Results.

Parameter	Vildagliptin - Metformin formulations	Teneligliptin -Metformin formulations	Dapagliflozin - Metformin formulations
PC1 Variance (%)	73.2	46.2	87.8
PC2 Variance (%)	20.3	33.3	6.5
Total Variance (PC1+PC2) (%)	93.5	79.5	94.3
Variance Used for HCA (%)	97.1	96.2	97.7
Cluster Separation	Complete separation of branded and generic formulations	Complete separation of branded and generic formulations	Complete separation of branded and generic formulations
Outliers Detected	No significant outlier	No significant outlier	DAPANORM-M5 identified as outlier
PCA Model Robustness	Excellent	Very Good	Excellent
Authentication Capability	Excellent	Very Good	Good

Table 3: HCA Model Validation Parameters.

Validation Parameter	Vildagliptin-Metformin formulations	Teneligliptin-Metformin formulations	Dapagliflozin-Metformin formulations
Variance Used for HCA	First 3 PCs = 97.1% cumulative variance	First 4 PCs = 96.2% cumulative variance	First 3 PCs = 97.7% cumulative variance
Number of Major Clusters	2 major clusters	2 major clusters	2 major clusters + 1 outlier
Brand-Generic Separation	Complete separation; branded and generic merge only at Ward distance ≈12	Strong separation; final merge at Ward distance ≈11.5	Weak-to-moderate separation; mixed cluster observed
Within-Group Similarity	High within-group similarity (distance 2.15-3.13)	Moderate-to-high similarity (distance 2.24-5.53)	Generic group highly similar (distance ≈1.57); branded group variable
Between-Group Dissimilarity	High (distance 7.18-9.69)	High (distance 7.42-11.10)	Moderate (distance varies widely due to mixed clustering)
Cluster Stability	Excellent	Very Good	Moderate
Branded Cluster Formation	Strong (Distinct compact cluster (Intaglip-M, Jalra-M, Vildagard M))	Strong (Distinct cluster (Dynaglipt-M, Teniva-M) with Tenlimac-M showing overlap)	Weak (No compact branded cluster; high dispersion among branded products)
Generic Cluster Formation	Strong (Distinct cluster (Vildaze-M, Vildader-Met, Vildanorm))	Strong (Predominantly generic cluster (Tenebetic-M, Tenepla-M, Tinleofin-M))	Very Strong Compact generic cluster (DAPAFORD-10M, Dapadax-M, Glyfloz-M)

Table 4: PLS-DA Model Validation Results.

Validation Parameter	Vildagliptin - Metformin formulations	Teneligliptin - Metformin formulations	Dapagliflozin - Metformin formulations
LOOCV Accuracy (%)	100.0	66.7	33.3
Sensitivity (%)	100.0	66.7	33.3
Specificity (%)	100.0	66.7	33.3
Precision (%)	100.0	66.7	33.3
F1 Score	1.00	0.67	0.33
Misclassification Rate (%)	0.0	33.3	66.7
Correctly Classified Samples	6/6	4/6	2/6
Incorrectly Classified Samples	0/6	2/6	4/6
Classification Performance	Excellent	Moderate	Poor

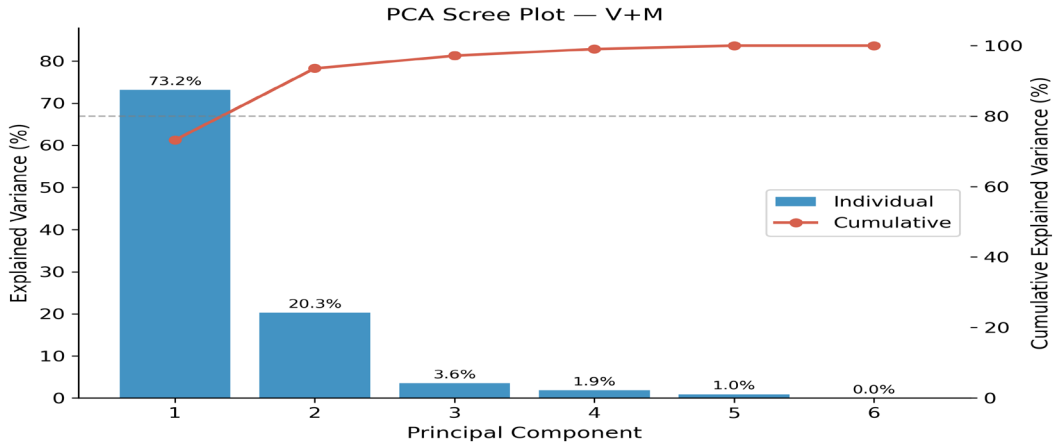


Figure 5a: PCA scree plot of vildagliptin and metformin formulations.

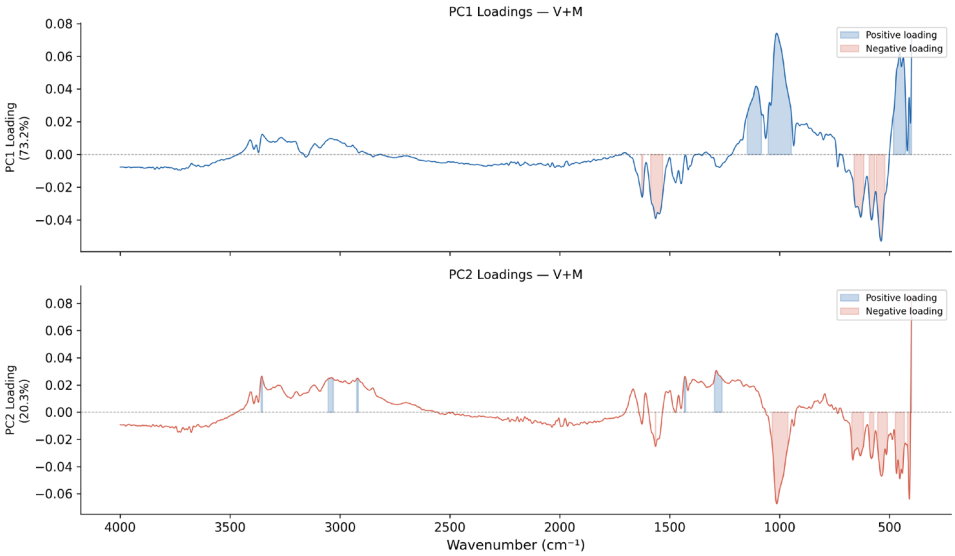


Figure 5b: PCA loading plot of vildagliptin and metformin formulations.

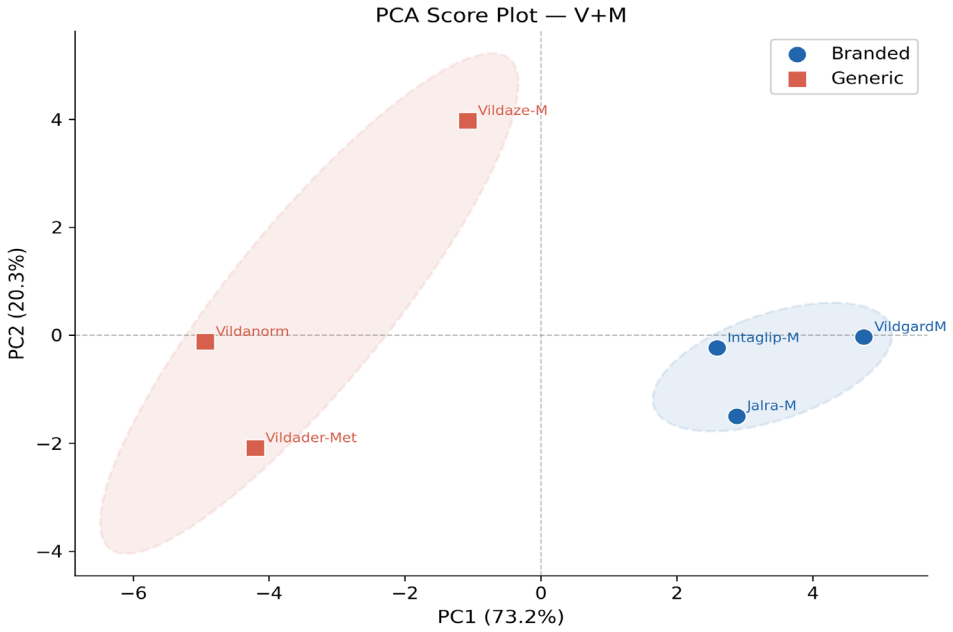


Figure 5c: PCA score plot of vildagliptin and metformin formulations.

compositional similarity expected among products containing the same active pharmaceutical ingredients. These findings support the suitability of FTIR-PCA as a rapid, non-destructive chemometric tool for authentication, classification, and quality evaluation of marketed antidiabetic tablet formulations.

Hierarchical Cluster Analysis

HCA of the ATR-FTIR spectral datasets further confirmed the discriminatory capability of FTIR spectroscopy for authentication and comparative evaluation of marketed antidiabetic formulations. For the dapagliflozin-metformin tablets, the dendrogram generated using Ward's linkage and principal component scores explaining 97.7% of the total variance revealed two major clusters, with Dapadax-M and Dapaford-10M exhibiting the highest similarity (Euclidean distance = 1.57), while Dapanorm-M5 formed a distinctly separated cluster, indicating greater spectral divergence from the remaining formulations. The heatmap analysis corroborated these findings, showing substantially higher inter-sample distances between Dapanorm-M5 and the other products (27.12-34.95). In the teneligliptin-metformin group, HCA based on four principal components explaining 96.2% of the variance identified clear clustering patterns, where Tenlimac-M and Tinleofin-M demonstrated the greatest similarity (distance = 2.24), while Dynaglipt-M and Teniva-M formed a separate branded cluster. The relatively high distances observed between branded and generic formulations (7.42-11.10) suggest notable differences in excipient composition and manufacturing characteristics despite

containing identical active pharmaceutical ingredients. Similarly, for the vildagliptin-metformin formulations, HCA performed on three principal components accounting for 97.1% of the variance produced two well-defined clusters corresponding predominantly to branded and generic products. Jalra-M and Vildgard-M exhibited strong similarity (distance = 2.39), while Vildader-Met and Vildanorm clustered closely together (distance = 2.15). The largest distances were observed between Vildgard-M and generic formulations (up to 9.69), indicating distinct spectral differences in formulation matrices. Overall, the HCA dendrograms and Euclidean distance heatmaps demonstrated high within-group similarity and substantial between-group dissimilarity, effectively differentiating branded and generic formulations. These findings confirm that ATR-FTIR spectral fingerprints combined with HCA provide a robust and reliable approach for pharmaceutical authentication, formulation classification, and quality evaluation of marketed antidiabetic tablets.

Partial Least Squares-Discriminant Analysis

PLS-DA was employed to evaluate the classification performance of ATR-FTIR spectral data for differentiating branded and generic antidiabetic tablet formulations. For the dapagliflozin-metformin products, the PLS-DA score plot demonstrated partial separation between branded and generic formulations along the latent variables; however, the LOOCV confusion matrix yielded a classification accuracy of only 33.3%, indicating considerable overlap in spectral characteristics and limited predictive discrimination. The VIP analysis identified the spectral

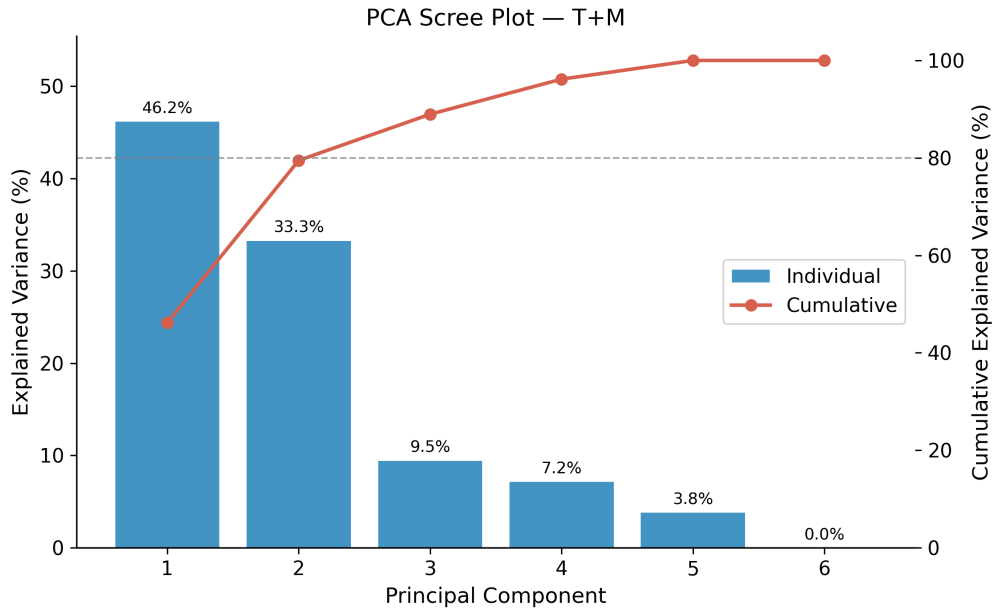


Figure 6a: PCA scree plot of teneligliptin and metformin formulations.

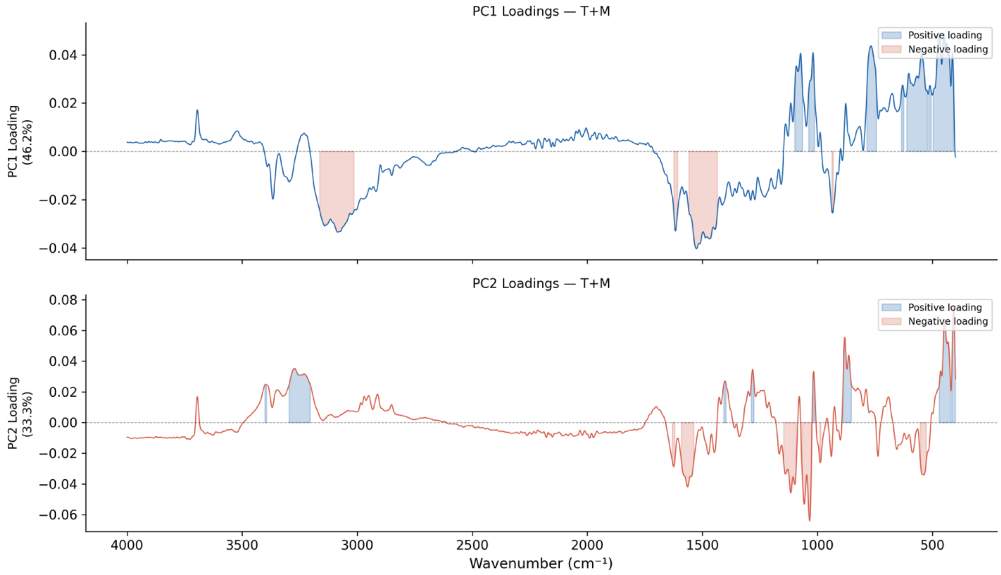


Figure 6b: PCA loading plot of teneligliptin and metformin formulations.

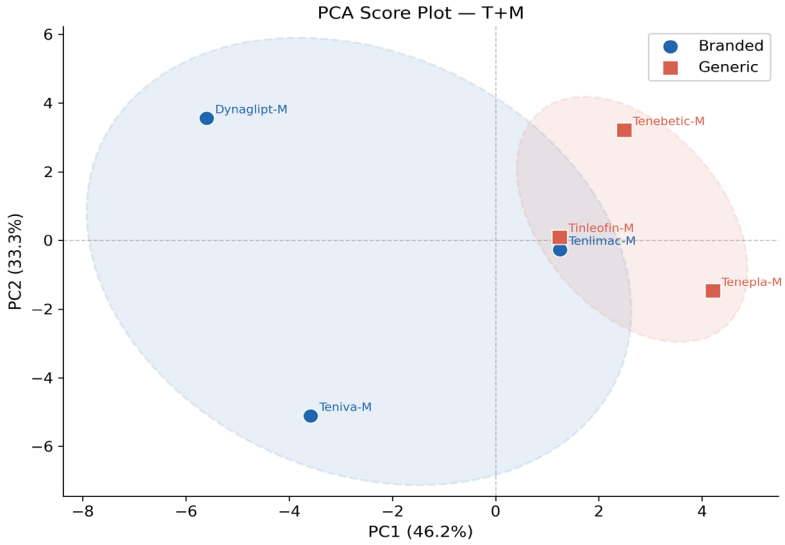


Figure 6c: PCA score plot of teneligliptin and metformin formulations.

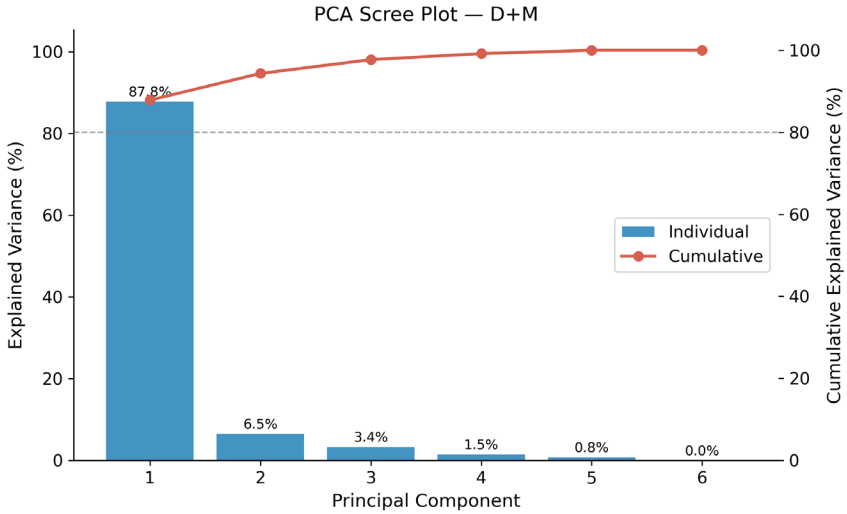


Figure 7a: PCA scree plot of dapagliflozin and metformin formulations.

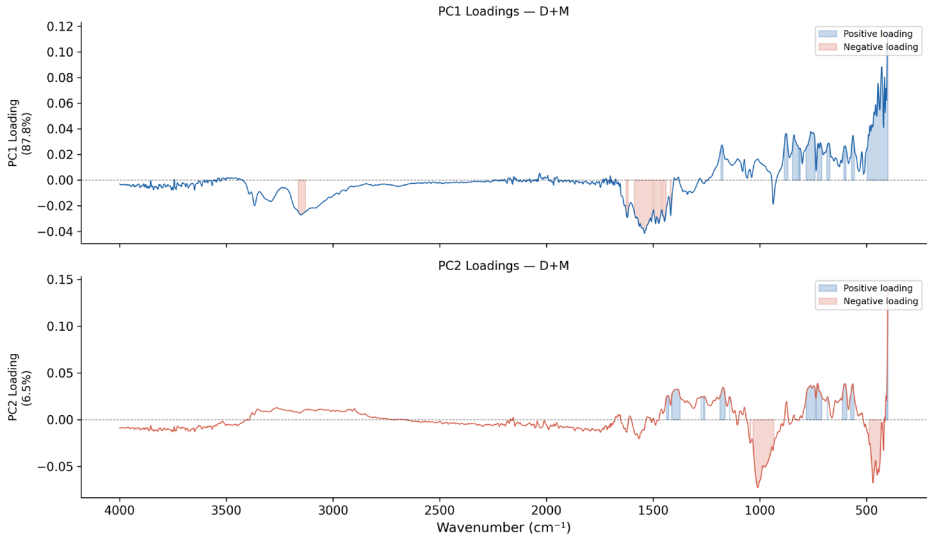


Figure 7b: PCA loading plot of dapagliflozin and metformin formulations.

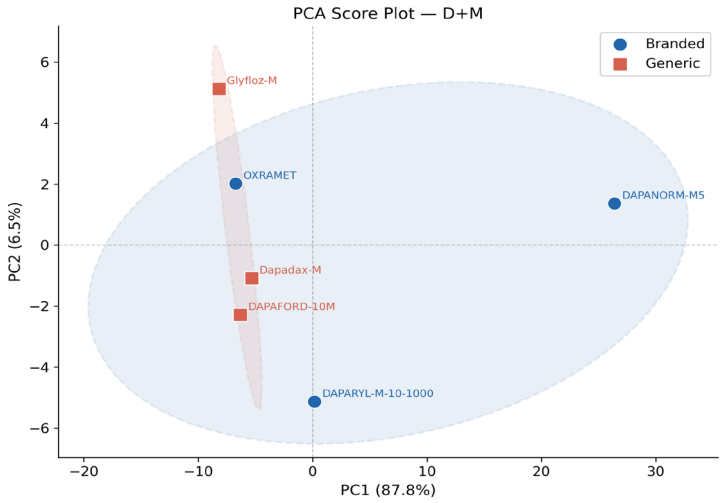


Figure 7c: PCA score plot of dapagliflozin and metformin formulations.

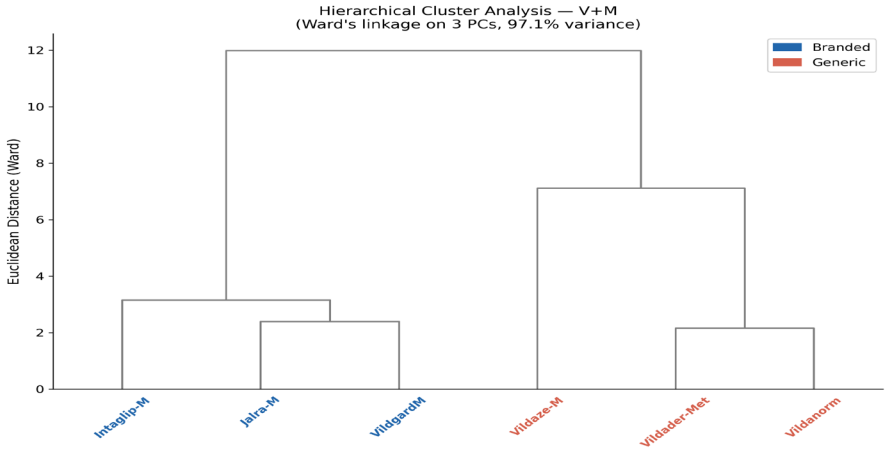


Figure 8a: HCA dendrogram of vildagliptin and metformin formulations.

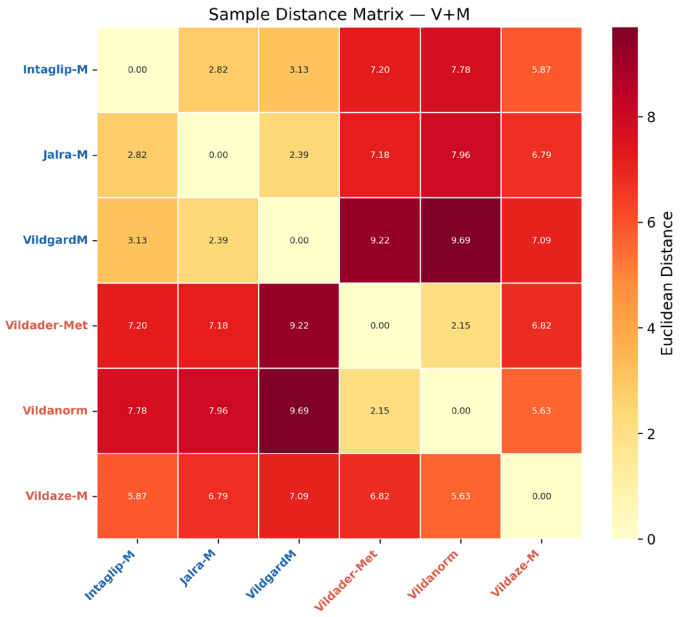


Figure 8b: HCA Heatmap Analysis of vildagliptin and metformin formulations.

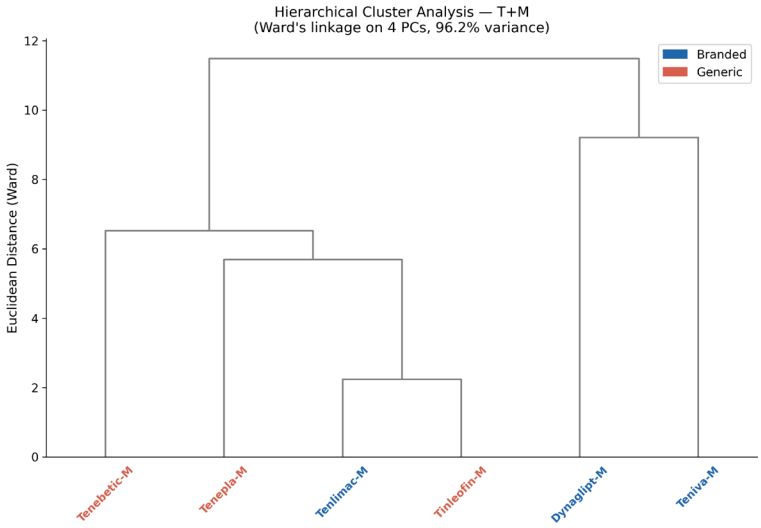


Figure 9a: HCA dendrogram of teneligliptin and metformin formulations.

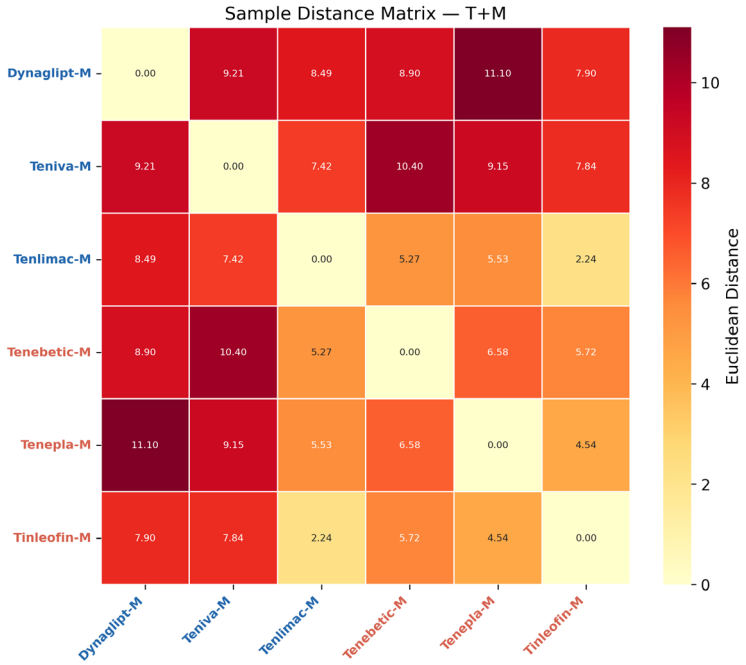


Figure 9b: HCA Heatmap Analysis of teneligliptin and metformin formulations.

regions around 400-454 cm^{-1} and 1000-1100 cm^{-1} as the most influential variables contributing to class differentiation. In contrast, the teneligliptin-metformin dataset exhibited improved discrimination, with the score plot showing clearer clustering of branded and generic formulations and a LOOCV classification accuracy of 66.7%. The VIP profile highlighted significant discriminatory contributions from the regions around 1074-1082 cm^{-1}, 411-476 cm^{-1}, 1500-1600 cm^{-1}, and 2900-3200 cm^{-1}, suggesting differences in excipient composition and molecular interactions among formulations. The vildagliptin-metformin formulations displayed the most distinct separation, with branded samples clustered on the positive side of LV1 and generic samples on the negative side, resulting in complete class discrimination and a LOOCV accuracy of 100%. The corresponding confusion matrix showed zero misclassifications, confirming excellent model robustness. VIP analysis revealed that the spectral region around 1002-1020 cm^{-1}, together with bands near 1500-1600 cm^{-1} and 400-500 cm^{-1}, contributed most significantly to formulation differentiation. Overall, the PLS-DA results demonstrated a progressive increase in classification performance from dapagliflozin-metformin to teneligliptin-metformin and vildagliptin-metformin formulations, confirming that ATR-FTIR spectroscopy coupled with chemometric modeling can effectively discriminate branded and generic pharmaceutical products, with the highest predictive capability observed for vildagliptin-metformin tablets.

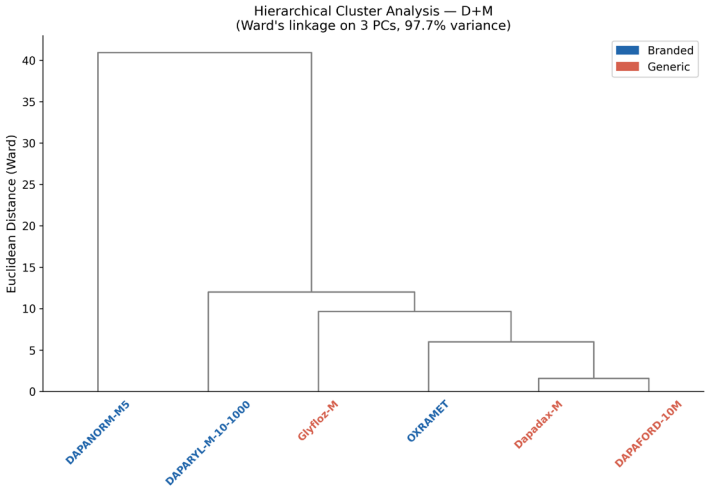


Figure 10a: HCA dendrogram of dapagliflozin and metformin formulations.

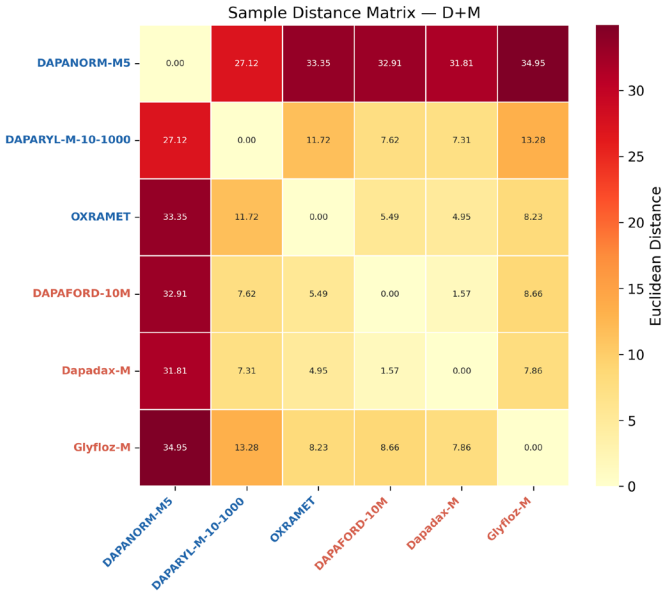


Figure 10b: HCA Heatmap Analysis of dapagliflozin and metformin formulations.

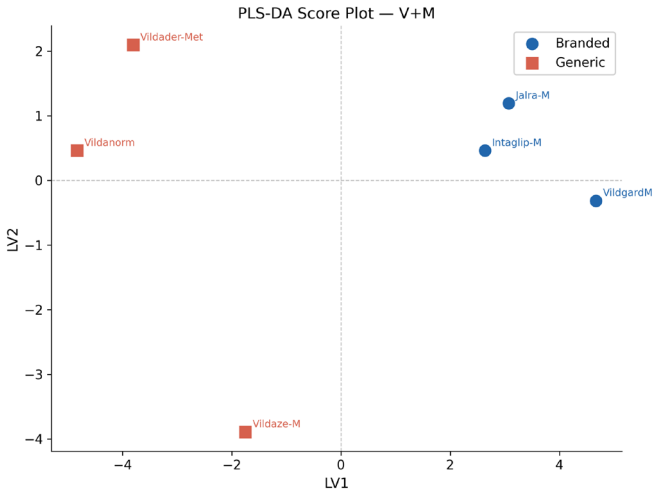


Figure 11a: PLS-DA Score Plot of vildagliptin and metformin tablet formulations.

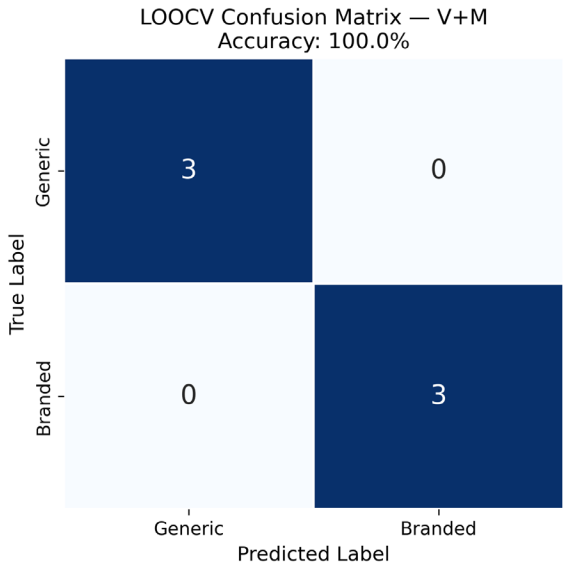


Figure 11b: LOOCV confusion matrix of vildagliptin and metformin tablet formulations.

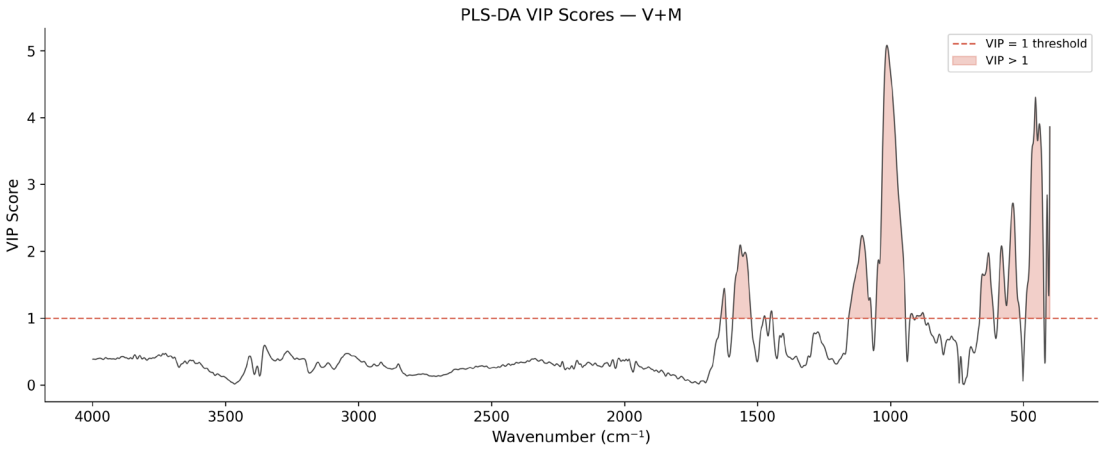


Figure 11c: PLS-DA VIP Score spectrum of vildagliptin and metformin tablet formulations.

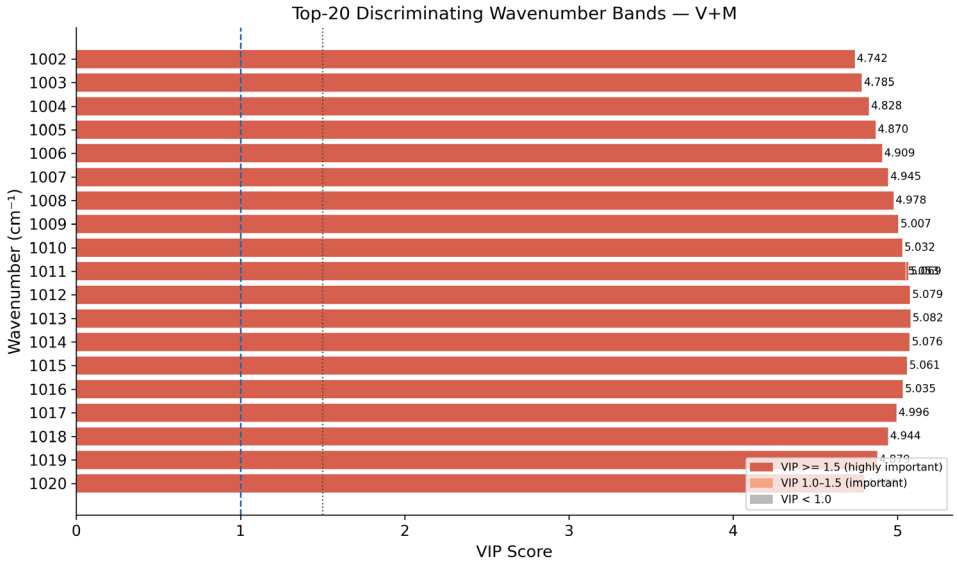


Figure 11d: Top discriminating wavenumber bands of vildagliptin and metformin tablet formulations.

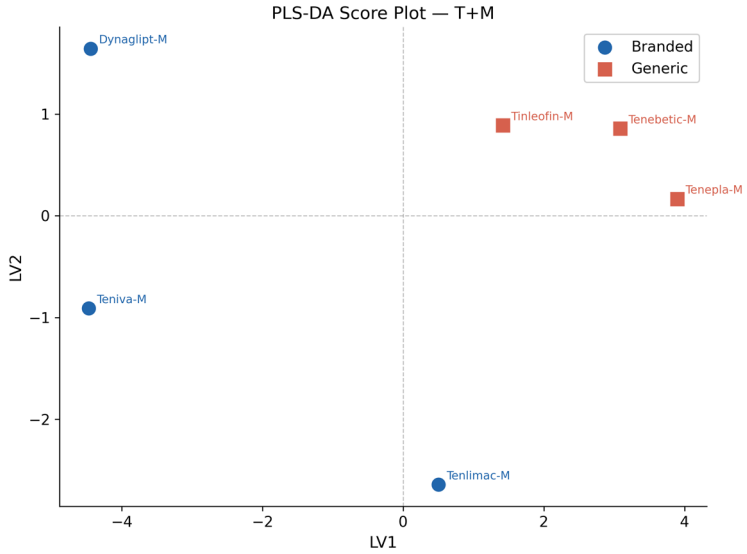


Figure 12a: PLS-DA Score Plot of teneligliptin and metformin tablet formulations.

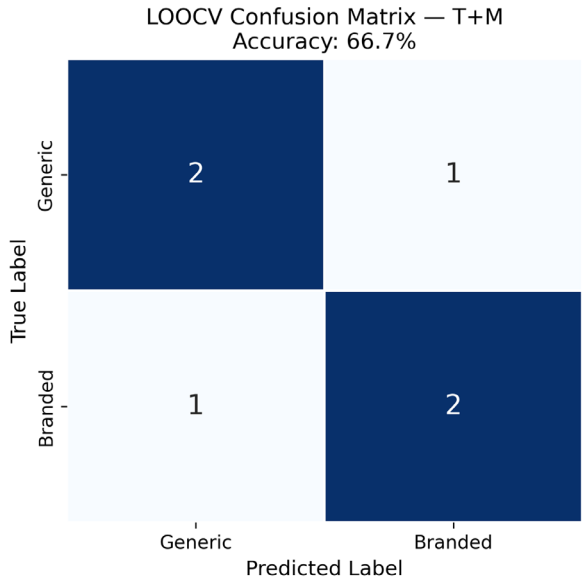


Figure 12b: LOOCV confusion matrix of teneligliptin and metformin tablet formulations.

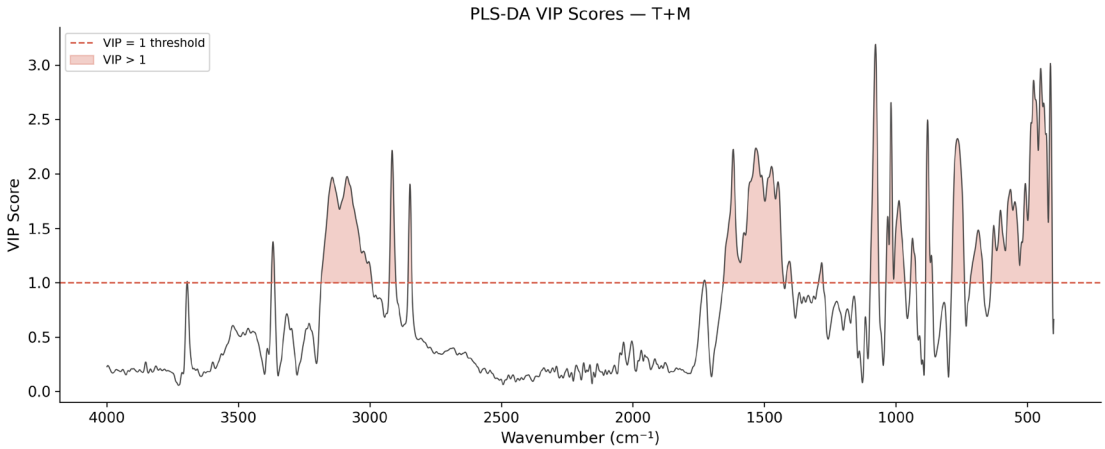


Figure 12c: PLS-DA VIP Score spectrum of teneligliptin and metformin tablet formulations.

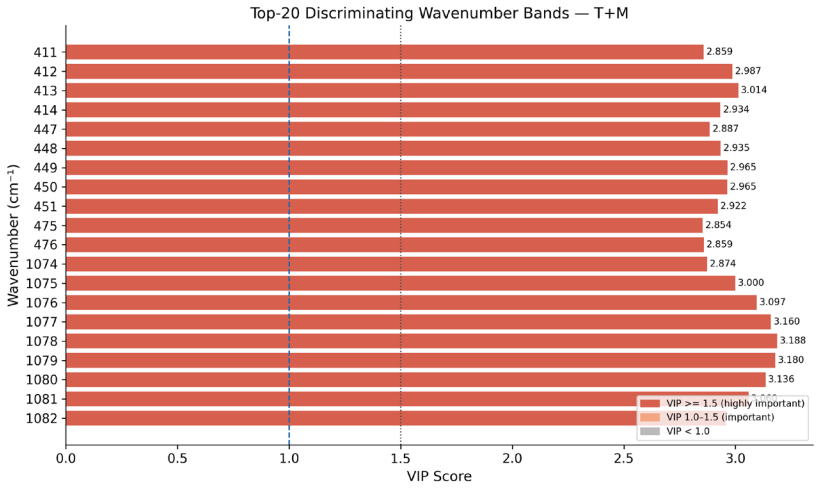


Figure 12d: Top discriminating wavenumber bands of teneligliptin and metformin tablet formulations.

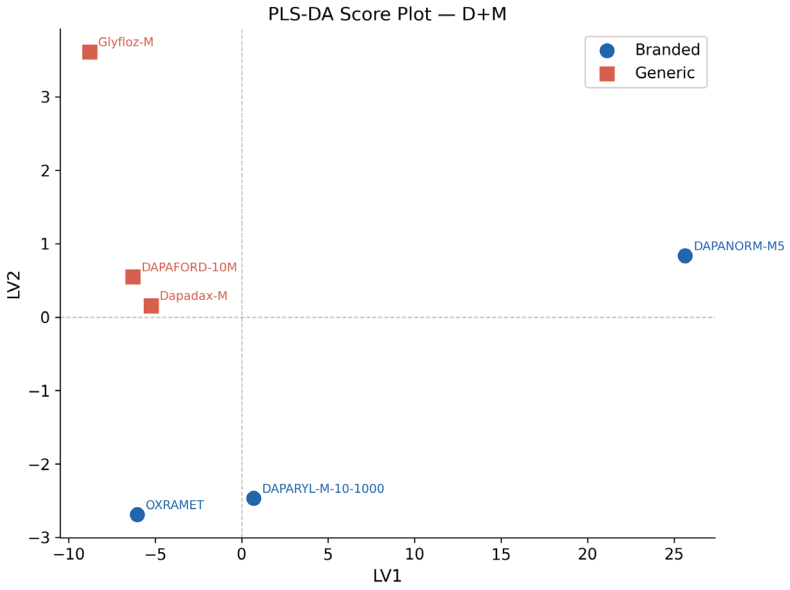


Figure 13a: PLS-DA Score Plot of dapagliflozin and metformin tablet formulations.

Validation of FTIR method - Chemometric Models

The Vildagliptin - Metformin PCA model showed the strongest discriminative performance with complete brand-generic separation and 93.5% variance explained by PC1 and PC2. The Teneligliptin + Metformin model demonstrated moderate separation due to overlap between one branded and one generic formulation. The Dapagliflozin - Metformin model achieved the highest PC1 variance (87.8%) and overall variance capture (94.3%), but exhibited weaker class discrimination because of greater chemical similarity among marketed formulations and the presence of one outlier sample (DAPANORM-M5).

These HCA validation results demonstrate that Vildagliptin - Metformin formulations exhibited the strongest brand-generic discrimination, followed by Teneligliptin - Metformin, whereas Dapagliflozin - Metformin formulations showed greater spectral similarity and pharmaceutical equivalence among marketed products, resulting in weaker cluster separation.

The PLS-DA validation results demonstrated varying classification performance among the three antidiabetic tablet combinations. Vildagliptin - metformin exhibited excellent discrimination with 100% accuracy, sensitivity, specificity, precision, and F1 score, with all samples correctly classified and no misclassifications. This indicates clear spectral differences between branded and generic formulations, enabling reliable classification. Teneligliptin - metformin showed moderate performance, achieving 66.7% accuracy and an F1 score of 0.67, with four out of six samples correctly classified. The presence of some misclassified samples suggests partial overlap between branded and generic spectral profiles. In contrast, Dapagliflozin - metformin demonstrated poor classification performance, with only 33.3% accuracy and an F1 score of 0.33. The high misclassification rate (66.7%)

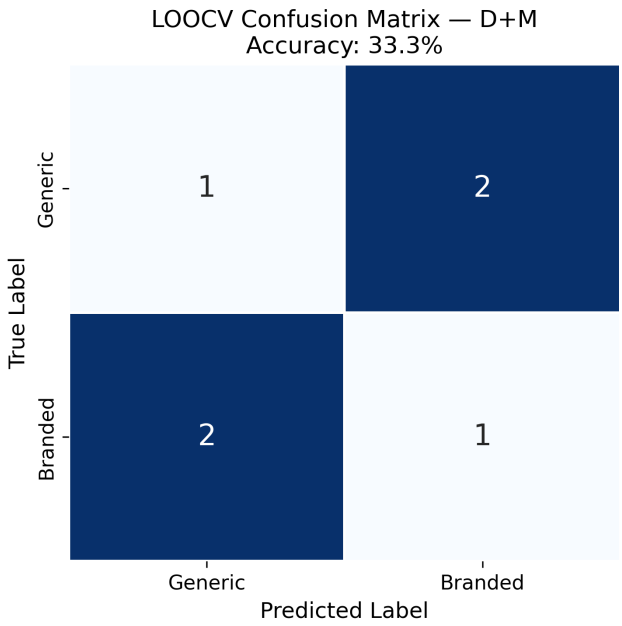


Figure 13b: LOOCV confusion matrix of dapagliflozin and metformin tablet formulations.

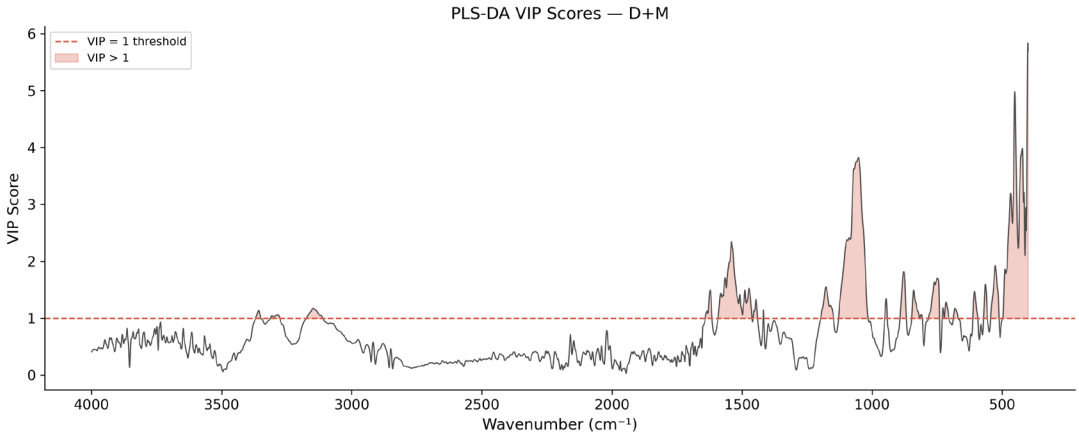


Figure 13c: PLS-DA VIP Score spectrum of dapagliflozin and metformin tablet formulations.

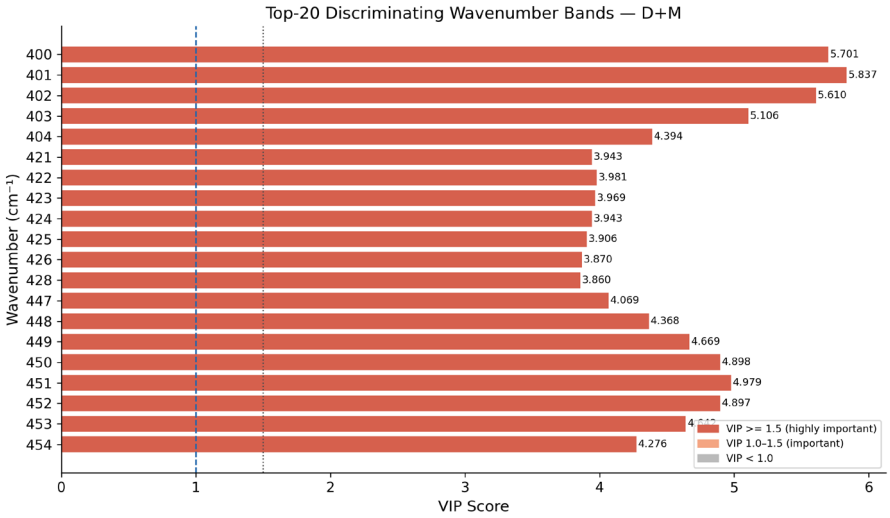


Figure 13d: Top discriminating wavenumber bands of dapagliflozin and metformin tablet formulations.

indicates substantial spectral similarity between branded and generic products, limiting the discriminative ability of the model.

Overall, the classification performance followed the order vildagliptin - metformin > teneligliptin - metformin > dapagliflozin - metformin, highlighting that the effectiveness of ATR-FTIR coupled with PLS-DA depends on the extent of formulation-related spectral differences present in each dosage form.

CONCLUSION

The present study demonstrated the successful application of ATR-FTIR spectroscopy coupled with chemometric techniques for rapid authentication and quality evaluation of marketed antidiabetic tablet formulations. PCA and HCA effectively differentiated formulations based on spectral characteristics, while PLS-DA provided robust classification and prediction performance. The developed methodology offers a rapid, non-destructive, environmentally sustainable, and cost-effective alternative to conventional chromatographic techniques. This approach has considerable potential for routine pharmaceutical quality control, counterfeit drug detection, and regulatory surveillance programs.

ACKNOWLEDGEMENT

The authors are thankful to the Principal and Management of Institute for providing laboratory facilities and research support to conduct this study. This work contains results obtained from the experiments performed at the SATHI-CISCoM at IIT Hyderabad, which was provided via the DST project SR/SATHI/2022/247.

ABBREVIATIONS

ATR-FTIR: Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy; **API:** Active Pharmaceutical Ingredient; **DM:** Diabetes Mellitus; **DPP-4:** Dipeptidyl Peptidase-4; **FTIR:** Fourier Transform Infrared Spectroscopy; **GC**: Gas Chromatography; **HCA:** Hierarchical Cluster Analysis; **HPLC:** High Performance Liquid Chromatography; **ICH:** International Council for Harmonization; **LC-MS:** Liquid Chromatography-Mass Spectrometry; **LOOCV:** Leave-One-Out Cross Validation; **PC:** Principal Components; **PCA:** Principal Component Analysis; **PLS-DA:** Partial Least Squares Discriminant Analysis; **PLSR:** Partial Least Squares Regression; **SGLT2:** Sodium Glucose Cotransporter-2; **T2DM:** Type 2 Diabetes Mellitus; **VIP:** Variable Importance in Projection.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHORS CONTRIBUTIONS

Pranay Kumar Ravula: Collection and/or assembly of data, Data analysis and interpretation. **Lakshmi Manogna:** Collection and/or assembly of data, Data analysis and interpretation, Writing the article. **Pravalika:** Collection and/or assembly of data. **Jahnavi Bandla:** Research concept and design, Collection and/or assembly of data, Data analysis and interpretation, Writing the article, Critical revision of the article and Final approval of the article.

REFERENCES

Ahmad, M. K. (2025). Application of chemometric-assisted FTIR spectroscopy for rapid authentication of adulterated herbal pharmaceuticals. *Journal of Pharmaceutical Analysis and Chemistry, 1*(1), 41-51.

Attia, K. A. M., Serag, A., Eid, S. M., & Abbas, A. E. F. (2022). A new chemometrically assisted UV spectrophotometric method for simultaneous determination of tamsulosin and dutasteride in their pharmaceutical mixture. *Journal of AOAC International, 105*(6), 1755-1761.

Beć, K. B., Grabska, J., & Huck, C. W. (2021). Principles and applications of infrared spectroscopy in pharmaceutical analysis: Recent advances. *Molecules, 26*(15), 4552. https://doi.org/10.3390/molecules26154552.

Beebe, K. R., Pell, R. J., & Seasholtz, M. B. (1998). Chemometrics: A practical guide. Wiley-Interscience, New York, 348.

Brereton, R. G. (2018). Chemometrics for pattern recognition (2nd ed.). John Wiley & Sons.

Brereton, R. G., & Lloyd, G. R. (2014). Partial least squares discriminant analysis: Taking the magic away. *Journal of Chemometrics, 28*(4), 213-225. https://doi.org/10.1002/cem.2609

Bro, R., & Smilde, A. K. (2014). Principal component analysis. *Analytical Methods, 6*(9), 2812-2831. https://doi.org/10.1039/C3AY41907J.

Bunaciu, A.A., Aboul-Enein, H.Y., & Fleschin, S. (2010). Application of Fourier Transform Infrared Spectrophotometry in Pharmaceutical Drugs Analysis. *Applied Spectroscopy Reviews, 45*(3), 206-219. https://doi.org/10.1080/00387011003601044.

Chaudhury, A., Duvoor, C., Reddy Dendi, V. S., Kraleti, S., Chada, A., Ravilla, R., Marco, A., Shekhawat, N. S., Montales, M. T., Kuriakose, K., Sasapu, A., Beebe, A., Patil, N., Musham, C. K., Lohani, G. P., & Mirza, W. (2017). Clinical Review of Antidiabetic Drugs: Implications for Type 2 Diabetes Mellitus Management. Frontiers in Endocrinology, 8(6). https://doi.org/10.3389/fendo.2017.00006

Darbandi, A., & Sohrabi, M. R. (2024). Chemometrics assisted UV spectrophotometry for the quantitative determination of metformin and sitagliptin in commercial sample compare to HPLC analysis. Iranian Journal of Chemistry and Chemical Engineering, 43(7), 2649-2659.

Deidda, R., Sacre, P. Y., Clavaud, M., Coic, L., Avohou, H., Hubert, P., & Ziemons, E. (2019). Vibrational spectroscopy in analysis of pharmaceuticals: Critical review of innovative portable and handheld NIR and Raman spectrophotometers. TrAC Trends in Analytical Chemistry, 114, 251-259. https://doi.org/10.1016/j.trac.2019.02.035

Eriksson, L., Byrne, T., Johansson, E., Trygg, J., & Vikström, C. (2013). *Multi- and megavariate data analysis: Basic principles and applications* (3rd ed.). Umetrics Academy.

Fatmarahmi, D. C., Susidarti, R. A., Swasono, R. T., & Rohman, A. (2022). A development method of FTIR spectroscopy coupled with chemometrics for detection of synthetic drug adulterants of herbal products in quaternary mixture. *Journal of Applied Pharmaceutical Science, 12*(3), 191-201. https://doi.org/10.7324/JAPS.2022.120322.

Galicia-Garcia, U., Benito-Vicente, A., Jebari, S., Larrea-Sebal, A., Siddiqi, H., Uribe, K. B., Ostolaza, H., & Martín, C. (2020). *Pathophysiology of type 2 diabetes mellitus*. *International Journal of Molecular Sciences, 21*(17), 6275. https://doi.org/10.3390/ijms21176275

Griffiths, P. R., & de Haseth, J. A. (2007). Fourier transform infrared spectrometry (2nd ed.). John Wiley & Sons. https://doi.org/10.1002/047010631X

Madbouly, E. A., El-Shanawani, A. A., El-Adl, S. M., & Abdelkhalek, A. S. (2024). Green chemometric-assisted UV-spectrophotometric methods for the determination of favipiravir, cefixime and moxifloxacin hydrochloride as an effective therapeutic combination for COVID-19: Application in pharmaceutical form and spiked human plasma. *BMC Chemistry,* 18(1), 65. https://doi.org/10.1186/s13065-024-01191-5.

Magalhães, S., Goodfellow, B. J., & Nunes, A. (2021). FT-IR spectroscopy in biomedical research: How to get the most out of its potential. *Applied Spectroscopy Reviews, 56*(8-10), 869-908. https://doi.org/10.1080/05704928.2021.1946822

Miller, J. N., & Miller, J. C. (2018). *Statistics and chemometrics for analytical chemistry* (7th ed.). Pearson Education.

Otto, M. (2017). *Chemometrics: Statistics and computer application in analytical chemistry* (3rd ed.). Wiley-VCH.

Palur, K., Archakam, S. C., & Koganti, B. (2020). Chemometric-assisted UV spectrophotometric and RP-HPLC methods for simultaneous determination of paracetamol, diphenhydramine, caffeine, and phenylephrine in tablet dosage form.

Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 243, 118801. https://doi.org/10.1016/j.saa.2020.118801

Pasquini, C. (2018). Near infrared spectroscopy: A mature analytical technique with new perspectives—A review. Analytica Chimica Acta, 1026, 8-36. https://doi.org/10.1016/j.aca.2018.04.004

Pruksapha, P., Khongkaew, P., Suwanvecho, C., Nuchtavorn, N., Phechkrajang, C., & Suntornsuk, L. (2023). Chemometrics-assisted spectroscopic methods for rapid analysis of combined anti-malarial tablets. *Journal of Food and Drug Analysis, 31*(2), 338-357.

Rohman, A., & Che Man, Y. B. (2012). The use of FTIR spectroscopy in combination with chemometrics for pharmaceutical and biomedical analysis. *Applied Spectroscopy Reviews, 47*(1), 1-20. https://doi.org/10.1080/05704928.2011.631094.

Rohman, A., Che Man, Y. B., Ismail, A., & Hashim, P. (2017). FTIR spectroscopy coupled with chemometrics of multivariate calibration and discriminant analysis for authentication of extra virgin olive oil. *International Journal of Food Properties, 20*(Suppl. 1), S1173-S1181. https://doi.org/10.1080/10942912.2017.1336718.

Sherif, M. E., Shymaa, S. S., Mohamed, R. E., & Ola, M. A. (2020). ATR-FTIR coupled with chemometrics for quantification of vildagliptin and metformin in pharmaceutical combinations having diverged concentration ranges. Vibrational Spectroscopy, 106, 102995. https://doi.org/10.1016/j.vibspec.2019.102995

Skoog, D. A., Holler, F. J., & Crouch, S. R. (2017). Principles of Instrumental Analysis (7th ed.). Sunder College Publisher, New York.

Smith, B. C. (2011). Fundamentals of Fourier transform infrared spectroscopy (2nd ed.). CRC Press. https://doi.org/10.1201/b10777

Stuart, B. H. (2004). Infrared spectroscopy: Fundamentals and applications. John Wiley & Sons. https://doi.org/10.1002/0470011149

Tiernan, H., Byrne, B., & Kazarian, S. G. (2020). ATR-FTIR spectroscopy and spectroscopic imaging for the analysis of biopharmaceuticals. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 241, 118636. https://doi.org/10.1016/j.saa.2020.118636

U.S. Food and Drug Administration. (2021). *Generic drug facts*. U.S. Department of Health and Human Services.

Van Haaren, C., De Bock, M., & Kazarian, S. G. (2023). Advances in ATR-FTIR spectroscopic imaging for the analysis of tablet dissolution and drug release. *Molecules, 28*(12), 4705. https://doi.org/10.3390/molecules28124705.

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems, 58*(2), 109-130. https://doi.org/10.1016/S0169-7439(01)00155-1.

World Health Organization. (2024a). *Diabetes*.

World Health Organization. (2024b). *Quality assurance of pharmaceuticals: A compendium of guidelines and related materials* (Vol. 1, 10th ed.).

Yuliani, F., Riyanto, S., & Rohman, A. (2018). Application of FTIR spectra combined with chemometrics for analysis of candlenut oil adulteration. *International Journal of Applied Pharmaceutics, 10*(5), 54-59. https://doi.org/10.22159/ijap.2018v10i5.27972.

Cite this article: Ravula PK, Parupudi LM, Pottigari P, Bandla J. Chemometric-Assisted ATR-FTIR Spectroscopy for Rapid Authentication and Quality Evaluation of Marketed Antidiabetic Tablet Formulations. Int. J. Pharm. Investigation. 2027;17(1):381-400.