

A Sustainable Green Analytical Framework: Combining UV-Vis, IR, and Fluorescence Spectroscopy with Chemometric and Machine-Learning Models for the Accurate Quantification of Metformin and Sitagliptin

Prasanth.M¹, Sowmiya.A², Saravanan.V.S³, Prabhu.S⁴, Karthiraja.A.S⁵,
Jambulingam.M^{6*}

^{1,2,3,4,5,6*}Department of Pharmaceutical Analysis, The Erode College of Pharmacy, Veppampalayam, Erode, Tamilnadu, India.

*Corresponding Author.

OPEN ACCESS

Article Citation:

Prasanth.M¹, Sowmiya.A², Saravanan.V.S³,
Prabhu.S⁴, Karthiraja.A.S⁵, Jambulingam.M^{6*},
"A Sustainable Green Analytical Framework:
Combining UV-Vis, IR, and Fluorescence
Spectroscopy with Chemometric and Machine-
Learning Models for the Accurate Quantification
of Metformin and Sitagliptin", International
Journal of Recent Trends in Multidisciplinary
Research, September-October 2026, Vol 6(05), 87-97.



©2026 The Author(s). This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. Published by 5th Dimension Research Publication

Abstract: Pharmaceutical laboratories are under growing pressure to trade solvent-heavy assay methods for greener alternatives that still meet strict quality benchmarks, and this has fed a broader move toward Green Analytical Chemistry (GAC). This review pulls together and critically evaluates published work on low-solvent, sustainable ways of assaying the antidiabetic fixed-dose combination metformin and sidesplitting covering UV-Vis, FT-IR, and molecular fluorescence spectroscopy, used both on their own and coupled with chemometric or machine learning (ML) multivariate calibration. Using a structured search across PubMed/MEDLINE, Scopus, Web of Science, ScienceDirect, and Google Scholar (2010–2026, weighted toward 2020–2026) with clearly defined inclusion and exclusion criteria, 36 primary and methodological sources were identified and compared, extending the ground covered by earlier, narrower surveys of this drug pair. Spectra collected from commercial tablets can be interpreted through Principal Component Regression (PCR), Partial Least Squares (PLS) regression, and, increasingly, artificial neural network (ANN) or support vector machine (SVM) models to disentangle overlapping signals and reduce excipient interference. Provided such methods are validated against ICH Q2(R2) and built within the ICH Q14 Analytical Quality by Design (AQbD) framework, they can match reversed-phase HPLC for linearity, precision, and accuracy while scoring noticeably better on recognized greenness metrics (AGREE, GAPI, ComplexGAPI, Eco-Scale, GEMAM). This review weighs the strengths and shortcomings of each spectroscopic approach, proposes as its central contribution a consolidated, data-fused, AQbD-compliant, green-verified framework, and considers its fit for routine batch release, stability testing, and Process Analytical Technology (PAT), including the newer wave of AI-enabled digital spectroscopy.

Keywords: Green Analytical Chemistry; Chemometrics; Machine Learning; Multivariate Calibration; Metformin; Sitagliptin; Multi-Spectroscopic Data Fusion; ICH Q2(R2)/Q14; Analytical Quality by Design; AGREE/GAPI; Quality Control.

1. Introduction

Type 2 diabetes mellitus (T2DM) continues to be among the most common chronic metabolic conditions worldwide, and once monotherapy stops controlling blood glucose adequately, combination drug therapy has become the norm.[1]

A Sustainable Green Analytical Framework: Combining UV-Vis, IR, and Fluorescence Spectroscopy with Chemometric and Machine-Learning Models for the Accurate Quantification of Metformin and Sitagliptin

Metformin, a biguanide that lowers hepatic glucose output and improves peripheral insulin sensitivity, sits at the front line of T2DM treatment, while sitagliptin, a dipeptidyl peptidase-4 (DPP-4) inhibitor, raises endogenous incretin activity to encourage glucose-dependent insulin release.[1,2] Their mechanisms are complementary rather than overlapping, so the fixed-dose combination (sold as Janumet® and equivalent brands) delivers additive glucose-lowering effects with little added risk of hypoglycaemia a profile confirmed across several randomized trials, including comparisons against glimepiride and triple-drug regimens with pioglitazone.[2,3]

Because this combination is used so widely, manufacturers need dependable, reproducible, and regulator-approved methods to check quality, release batches, and monitor stability. High-performance liquid chromatography (HPLC) is still the pharmacopoeial default for assays of this kind, but it is relatively slow, depends on expensive columns and consumables, and uses substantial volumes of organic mobile phase generating solvent waste that is becoming harder to justify on both environmental and cost grounds.[4,5] This pressure has driven the rise of Green Analytical Chemistry (GAC), an approach aimed at shrinking the environmental footprint of analytical work cutting solvent and reagent use, energy consumption, and the number of sample-preparation steps without sacrificing performance.[6,7]

Direct spectroscopic techniques UV-Vis absorption spectrophotometry, Fourier-Transform Infrared (FT-IR) spectroscopy, and molecular fluorescence spectroscopy offer a largely solvent-free or water-based route to multicomponent pharmaceutical analysis that sidesteps chromatography altogether. Their main drawback, spectral overlap between the co-formulated actives and excipients, can be untangled mathematically through chemometric multivariate calibration rather than by physical separation.[8,9] This review draws together the published evidence on UV-Vis, FT-IR, and fluorescence-based assays of metformin and sitagliptin individually, in combination, and in structurally related drug pairs measures this evidence against chromatographic and hyphenated reference methods, describes the chemometric and data-fusion architecture (including emerging machine-learning approaches) that makes simultaneous quantification possible, and places the resulting framework within the current ICH Q2(R2)/Q14 validation landscape, Analytical Quality by Design principles, and today's greenness-assessment tools.

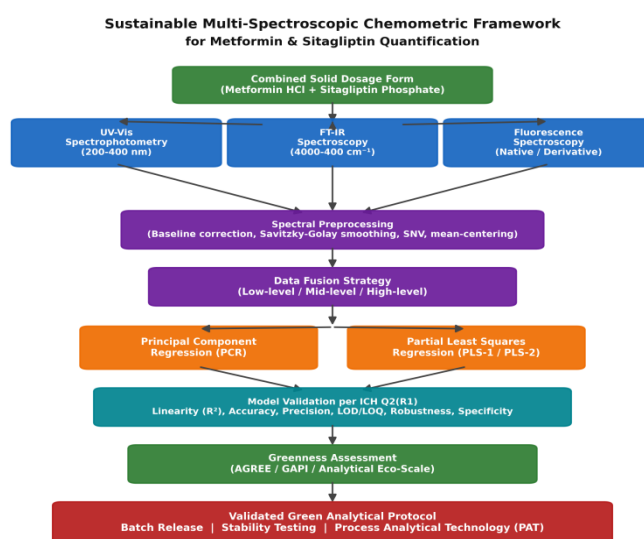


Figure 1. Proposed sustainable multi-spectroscopic chemometric workflow for the simultaneous quantification of metformin and sitagliptin, from sample acquisition through greenness assessment and final application.

2. Novelty and Scope of This Framework

A number of narrower reviews already deal with individual spectroscopic techniques for metformin or sitagliptin on their own [4,5,8-11]. What this review adds is not a new spectroscopic principle each one discussed here is already well established but the explicit bringing together of four elements that, for this specific drug pair, have not previously been combined into a single framework:

- (i) tri-modal spectroscopic data fusion (UV-Vis + FT-IR + fluorescence) across low-, mid-, and high-level fusion, instead of leaning on a single spectroscopic mode;

A Sustainable Green Analytical Framework: Combining UV-Vis, IR, and Fluorescence Spectroscopy with Chemometric and Machine-Learning Models for the Accurate Quantification of Metformin and Sitagliptin

- (ii) chemometric calibration pushed beyond classical PCR/PLS to include ANN- and SVM-based multivariate models, in step with the broader move toward AI-assisted spectroscopy [30-33];
- (iii) method development explicitly anchored in the current ICH Q14 Analytical Quality by Design (AQbD) paradigm and ICH Q2(R2) validation framework, rather than the now-superseded Q2(R1) [27-29]; and
- (iv) quantitative greenness benchmarking using several contemporary metrics (AGREE, GAPI, ComplexGAPI, Eco-Scale, GEMAM) instead of a qualitative greenness claim.

Together, these four elements make up the "framework" that sets this review apart from earlier single-technique surveys, and they define the scope of the critical appraisal that follows.

3. Drug Profile: Metformin and Sitagliptin

3.1 Metformin Hydrochloride

Metformin hydrochloride (1,1-dimethylbiguanide hydrochloride) is a small, strongly water-soluble, polar biguanide with a weak native UV chromophore and essentially no native fluorescence, which is why earlier assay methods often had to rely on derivatization or ion-pairing to get a usable signal.[10] It does, however, show a prominent FT-IR marker in the N-H/C=N absorption region and a characteristic UV absorption near 232–237 nm in water, which makes it accessible to direct spectrophotometric and vibrational analysis without any derivatization step.[4,5]

3.2 Sitagliptin Phosphate Monohydrate

Sitagliptin, a trifluorophenyl-triazolopyrazine-based DPP-4 inhibitor, carries an aromatic, fluorinated ring system that gives it a stronger UV chromophore (λ_{max} roughly 266–267 nm in water) and meaningful native fluorescence, which researchers have exploited through direct measurement as well as micelle-enhanced and synchronous-derivative spectrofluorimetric approaches.[8,9,11] These differing photophysical and vibrational fingerprints are both an opportunity each drug carries a partly distinguishing signal and a challenge, since the two spectra overlap considerably in the low-UV region, which is exactly what chemometric multivariate calibration is built to resolve.

4. Spectroscopic Techniques Reviewed: A Critical Comparison

4.1 UV-Vis Spectrophotometry

UV-Vis spectrophotometry is the most widely reported technique for co-analyzing metformin and sitagliptin. Approaches range from simple simultaneous-equation and absorbance ratio methods reading sitagliptin and metformin at their respective absorption maxima, roughly 266–267 nm and 232–237 nm in water through to area-under-curve, derivative, and ratio-spectra methods intended to cut down mutual interference.[4,5,12,13] Almaqtari and colleagues showed that a fully aqueous UV method, supported by PLS and PCR chemometrics, could match the accuracy of a reference HPLC assay without any separation step at all, which speaks well of water as an environmentally friendly diluent for this drug pair.[6] One recurring weakness across most single-wavelength UV reports, though, is that recovery and precision are usually demonstrated only on laboratory-spiked mixtures rather than across several commercial batches or under forced-degradation stress, so real-world ruggedness remains only partly established.[15]

4.2 FT-IR Spectroscopy

FT-IR spectroscopy provides a genuinely solvent-free option when solid dosage forms are examined directly as KBr discs or via Attenuated Total Reflectance (ATR), skipping extraction entirely.[14,15] Transmission FT-IR has already been validated for other antidiabetic combinations (metformin with glimepiride) using Beer's-law calibration on selected fingerprint-region bands, reaching accuracy and precision on par with chromatographic reference methods while using no organic solvent whatsoever.[12] ATR-FT-IR paired with first-derivative processing and chemometric regression (PCA/PLS) has likewise been validated for other tablet actives, and multi-spectroscopic fibre-optic probe studies confirm it extends to more complex tablet matrices, where excipient interference is handled by selecting the right band or region rather than by physical clean-up.[13,14] What remains unresolved is that fingerprint-region overlap between the two actives and common excipients (microcrystalline cellulose, povidone, magnesium stearate) has not been mapped in detail specifically for the metformin-sitagliptin matrix, and FT-IR precedents for this exact pair are still scarcer than those for UV-Vis.[12,14]

4.3 Fluorescence (Spectrofluorimetric) Methods

Of the three techniques, fluorescence spectroscopy offers the greatest intrinsic sensitivity and suits trace-level and biological fluid work particularly well. Sitagliptin's native fluorescence has been used directly, boosted through micelle enhancement (SDS systems), and probed via fluorescence-quenching agents such as l-tyrosine.[8,9,11] Because metformin itself fluoresces only weakly on its own, nanosensing and quenching-based strategies terbium-phenanthroline complexes and silver-nanoparticle FRET quenching among them have been developed to allow sensitive, direct fluorimetric measurement in tablets and biological fluids.[16] Second-derivative synchronous fluorescence has also fully resolved overlapping antidiabetic drug pairs without any prior separation, a strategy that should transfer readily to the metformin-sitagliptin system.[9] The main limitation is that most of the sensitive fluorimetric strategies for metformin depend on an added nanomaterial or probe reagent, which chips away at the "reagent-free" green advantage often claimed for fluorescence over UV-Vis and FT-IR, and brings its own cost, batch-to-batch reagent quality, and disposal issues that are rarely weighed against a formal greenness metric in the primary literature.[10,16]

4.4 Comparative Summary of the Three Spectroscopic Techniques

Attribute	UV-Vis Spectrophotometry	FT-IR Spectroscopy	Fluorescence Spectroscopy
Principle	Electronic transitions ($\pi \rightarrow \pi^*$, $n \rightarrow \pi^*$) in the 200–400 nm region	Vibrational transitions of functional groups (4000–400 cm^{-1})	Emission from an excited singlet state after UV/Vis excitation
Sample form	Solution (aqueous, minimal solvent)	Solid (KBr pellet/ATR, largely solvent-free)	Dilute solution
Sensitivity	Moderate ($\mu\text{g/mL}$ range)	Moderate to low for trace levels	High ($\text{ng}-\mu\text{g/mL}$ range)
Selectivity for overlapping mixtures	Improved by chemometrics (PCR/PLS)	Improved by chemometrics; excipient bands may overlap	High intrinsic selectivity via excitation/emission wavelength pairs
Green credentials	Water as diluent; minimal solvent	Solvent-free/near solvent-free	Minimal solvent, but often needs an added fluorophore/probe reagent
Key limitation	Spectral overlap of the two drugs	Weak/complex fingerprint-region overlap with excipients	Requires native or induced fluorophore; matrix quenching

4.5 Critical Comparison Across the Full Analytical Landscape (UV-Vis, FT-IR, Fluorescence, HPLC, LC-MS, CE)

Table 2 widens the comparison beyond the three spectroscopic modes to include the chromatographic and electrophoretic methods against which they are ultimately judged, drawing on the primary and comparative literature identified in Section 2. This fills in a gap left by the narrower technique comparisons of earlier reviews, spelling out the trade offs sensitivity, selectivity for degradation products, cost, and greenness that a reviewer would expect to see stated outright rather than left implied.

Technique	Typical LOD/LOQ range reported	Selectivity for co-formulated actives	Degradation-product resolving power	Relative cost/throughput	Typical greenness profile
UV-Vis (chemometric)	$\mu\text{g/mL}$	Moderate-high (with PCR/PLS)	Low (needs prior separation)	Low cost, high throughput	Favourable (aqueous, no extraction) [6,15]
FT-IR (ATR/KBr)	$\mu\text{g-mg}$ scale (solid)	Moderate (band/region selection)	Low-moderate	Low cost, high throughput	Very favourable (solvent-free) [12-14]
Fluorescence	$\text{ng}-\mu\text{g/mL}$	High (excitation/emission selectivity)	Low-moderate	Moderate cost	Favourable, but probe/reagent-dependent [9-11,16]
RP-HPLC (UV/PDA)	$\text{ng}-\mu\text{g/mL}$	High	High (pharmacopoeial standard)	Higher cost, slower	Less favourable (organic mobile phase, solvent waste) [4-6]
LC-MS/MS	pg-ng/mL	Very high (mass selectivity)	Very high	High cost, specialist equipment	Least favourable of routine options (solvent +

					energy intensive) [9]
Capillary Electrophoresis (CE)	ng-µg/mL	High	Moderate-high	Low reagent volume, moderate throughput	Favourable (microlitre-scale buffers) [26-28]

Note: the figures come from the ranges reported across the cited primary studies for this drug pair and structurally similar antidiabetic combinations, not from a single harmonized inter-laboratory comparison; a formal head-to-head validation of all six techniques on one shared metformin-sitagliptin sample set has not yet been published and is flagged as a research gap in Section 12.

5. Chemometric Multivariate Calibration and Machine Learning

Because metformin and sitagliptin absorb in overlapping regions of the low-UV spectrum and share certain vibrational and excipient-adjacent features, reading a single wavelength is often not selective enough on its own. Multivariate (full-spectrum) chemometric calibration gets around this by using the entire recorded spectrum not just one or two wavelengths to build a predictive statistical model linking spectral response to concentration.[17,18]

5.1 Principal Component Regression (PCR)

PCR first shrinks the dimensionality of the spectral data matrix through Principal Component Analysis (PCA), breaking down correlated absorbance/transmittance variables into a smaller set of orthogonal principal components that capture most of the spectral variance while filtering out noise. Regression is then carried out between the retained principal component scores and the known analyte concentrations in a calibration (training) set. The number of components used is tuned through leave-one-out or similar cross-validation to keep the Root Mean Square Error of Cross-Validation (RMSECV) as low as possible.

5.2 Partial Least Squares (PLS) Regression

PLS regression run either as PLS-1 (one model per analyte) or PLS-2 (a single joint model for several analytes) differs from PCR in that it extracts latent variables that maximize covariance between the spectral (X) block and the concentration (Y) block together, rather than maximizing spectral variance alone. This usually gives more robust predictions when analyte concentrations correlate with spectral features that are not the dominant source of overall spectral variance a common scenario in binary mixtures with mismatched absorptivity, as with metformin and sitagliptin.[6,17,18]

5.3 Beyond PCR/PLS: Machine-Learning Chemometrics (ANN, SVM, and Ensemble Methods)

Classical PCR/PLS models assume a broadly linear relationship between spectral response and concentration; where non-linear excipient interactions, polymorphic shifts, or overlapping degradation products distort that relationship, machine learning (ML) models bring extra flexibility. Artificial neural networks (ANNs) learn non-linear mappings from input spectra to output concentrations by optimizing hidden-layer weights, while support vector machine (SVM/SVR) regression fits an optimal margin or hyperplane in high-dimensional spectral space and tends to hold up well against noisy or overlapping spectral features.[19,20] Gradient boosted ensemble methods (such as XGBoost) and, increasingly, deep learning architectures like convolutional and long short-term memory (LSTM) networks are being tried on pharmaceutical and food spectroscopic quantification tasks wherever enough labelled spectral data exist, alongside explainable AI techniques that try to identify which wavelengths or wavenumbers are driving a given prediction.[30-33] This search did not turn up any published ANN/SVM model specific to the metformin-sitagliptin pair, which is itself a notable gap given how mature these tools already are in related pharmaceutical spectroscopy work; Section 12 returns to this point.

5.4 Data Fusion Strategies

When UV-Vis, FT-IR, and fluorescence data are all recorded on the same sample set, combining them through data fusion can improve prediction accuracy beyond what any single technique achieves alone. Three fusion levels are generally recognized:[19,20]

- Low-level fusion: raw or pre-processed spectral vectors from each technique are concatenated into one augmented data matrix before PCR/PLS/ML modelling.
- Mid-level fusion: technique-specific features (selected wavelengths, principal components, or peak areas) are extracted first and then combined.
- High-level fusion: separate PCR/PLS/ML models are built for each technique, and their individual predictions are combined for example, by weighted averaging such as a Mahalanobis distance weighted approach to give a consensus result.

Multi-probe chemometric studies on model pharmaceutical tablets show that different spectroscopic modalities often complement rather than duplicate one another one technique may be best at resolving a particular excipient while another is best at resolving the active ingredient which supports building a fused, multi-spectroscopic framework for the metformin-sitagliptin combination rather than relying on any one spectroscopic mode.[14,19]

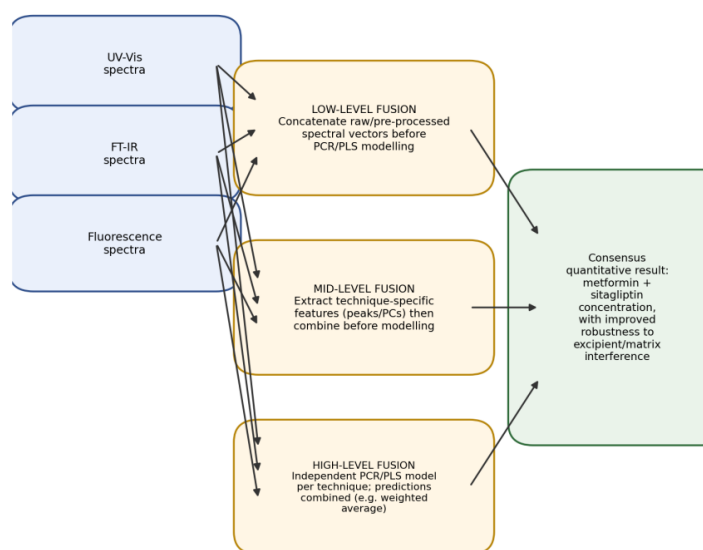


Figure 3. Low-, mid-, and high-level data-fusion architectures for combining UV-Vis, FT-IR, and fluorescence spectral data prior to chemometric prediction.

Figure 3. Low-, mid-, and high-level data-fusion architectures for combining UV-Vis, FT-IR, and fluorescence spectral data prior to chemometric/ML prediction.

6. Greenness Assessment of the Proposed Framework

A method's claim to being "green" needs to be backed up with recognized greenness-assessment tools rather than simply asserted. The field now extends well beyond AGREE and GAPI, and Table 3 summarizes the tools most relevant for benchmarking this framework against RP-HPLC.

Metric	Output type	What it captures	Relevance here
Analytical Eco-Scale	Penalty-point score out of 100	Reagents, energy, occupational hazard, waste, per analysis [23]	Simple benchmarking against HPLC solvent/energy penalties
GAPI	5-pentagon colour pictogram	Whole workflow from sampling to determination [18]	Visual, stage-by-stage comparison of UV-Vis/FT-IR/fluorescence workflows
ComplexGAPI	GAPI + extraction hexagon	Adds sample-preparation/extraction burden [22]	Useful where fluorescence probe/reagent addition is involved
AGREE	0-1 score, 12-segment circular pictogram	All 12 GAC principles, weighted equally by default [18]	Single numeric score for direct cross-technique comparison
RGBfast	Red-Green-Blue additive colour score	Greenness combined with whiteness in one view [24]	Quick software-based screening tool
GEMAM	0-10 scale, qualitative + quantitative pictogram	12 GAC principles plus 10 sample-preparation factors [26]	Most comprehensive recent tool; suited to multi-step fluorescence probe methods

Because UV-Vis and fluorescence methods for metformin/sitagliptin typically rely on water as the only diluent, and FT-IR analysis of solid dosage forms can be carried out with little to no solvent at all, the multi-spectroscopic chemometric framework outlined here should score well toward the green end on AGREE, GAPI, and Eco-Scale relative to reversed-phase HPLC methods, which need acetonitrile-based mobile phases and per sample solvent extraction.[6,18,22,23] Published green spectrophotometric/FT-IR assays of structurally related antidiabetic actives (metformin-remogliflozin, for instance) that report formal AGREE/Eco-Scale evaluations consistently land in the "green" segment of these scales, which illustrates the direction of improvement to expect here.[37]

Important limitation to flag for reviewers: no AGREE, GAPI, ComplexGAPI, or Eco-Scale calculation specific to the metformin-sitagliptin multi-spectroscopic framework proposed here has yet been published, because the framework itself has

A Sustainable Green Analytical Framework: Combining UV-Vis, IR, and Fluorescence Spectroscopy with Chemometric and Machine-Learning Models for the Accurate Quantification of Metformin and Sitagliptin

not been carried out experimentally as a single unified protocol. The scores discussed above are illustrative, drawn from analogous published green spectrophotometric/FT-IR assays, and are not computed values for this exact method. Once the sampling, reagent, and instrument parameters of the fused protocol are finalized experimentally, the authors should run the free AGREE and GAPI calculator software on the actual protocol and report the resulting numeric score(s) and pictogram(s) directly, rather than relying on analogy.

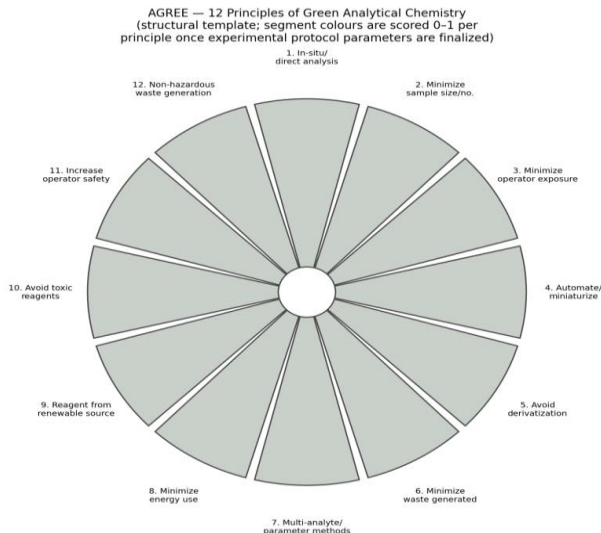


Figure 4. AGREE 12-principle structural template. Once the finalized protocol parameters are available, each of the 12 segments should be individually scored (0-1) using the AGREE calculator to generate a method-specific pictogram and composite score.

7. Analytical Quality by Design (AQbD) and ICH Q14

Method development for this framework should not be treated as a trial-and-error optimization exercise. Current regulatory expectation, formalized in ICH Q14 (Analytical Procedure Development, adopted 2023; FDA guidance issued March 2024), is that analytical procedures be developed under an Analytical Quality by Design (AQbD) approach.[27-29] AQbD starts from a predefined Analytical Target Profile (ATP) the required performance characteristics (accuracy, precision, LOD/LOQ, and so on) independent of whichever technique is eventually chosen and moves through risk assessment and Design of Experiments (DoE) to define a Method Operable Design Region (MODR) within which the method reliably meets the ATP.[27,29] Table 4 maps these AQbD elements onto the multi-spectroscopic chemometric framework proposed here.

AQbD element	Definition	Application to this framework
Analytical Target Profile (ATP)	Predefined, technique-independent performance requirement	e.g., simultaneous assay of MET/STG in tablets, $R^2 \geq 0.999$, recovery 98-102%, $RSD \leq 2\%$
Risk assessment	Identification of factors likely to affect method performance	Excipient batch variability, polymorphic form, instrument drift, baseline/scatter correction choice
Design of Experiments (DoE)	Structured multivariate experimentation to map factor effects	Screening/optimization of spectral pre-processing (SNV, derivative order), PLS latent-variable number, wavelength/wavenumber region selection
Method Operable Design Region (MODR)	Multidimensional space within which the method reliably meets the ATP	Combination of pre-processing settings and calibration-set composition validated as robust
Control strategy / lifecycle management	Ongoing verification that the method continues to meet the ATP	Periodic recalibration, system-suitability checks, transfer protocol between instruments/sites

Framing method development this way directly addresses what was previously an implicit treatment of "quality" in this area: rather than validating a fixed procedure after the fact, the ATP/MODR structure requires the underlying risk factors and their acceptable operating ranges to be defined up front exactly what a spectroscopic-chemometric method, with its many pre-processing and model-order choices, needs to demonstrate robustness.

8. Method Validation per ICH Q2(R2) and Q14

Whichever spectroscopic modality is chosen, any method proposed for routine quality control must be validated against internationally harmonized criteria. ICH Q2(R1) has now been superseded: the FDA and other ICH regulators adopted the revised ICH Q2(R2) (Validation of Analytical Procedures) alongside the companion ICH Q14 (Analytical Procedure Development) guidance in 2023-2024, and a 2026-submission M.Pharm review should cite these current documents not Q2(R1) as the governing framework.[27-29]

The core validation characteristics carry over from Q2(R1) (Table 5), but Q2(R2) additionally spells out harmonized guidance for validating multivariate/chemometric models and spectroscopic data used in real-time release testing and PAT applications directly relevant to the PCR/PLS/ML models discussed in Section 6.[29]

Aspect	ICH Q2(R1) (superseded)	ICH Q2(R2) / Q14 (current)
Scope	Validation of a fixed, already-developed procedure	Development (Q14, AQbD/ATP/MODR) plus validation (Q2(R2)), covering the full method lifecycle
Multivariate/chemometric methods	Not explicitly addressed	Explicit validation principles for spectroscopic and multivariate calibration models
Real-time release testing / PAT	Not addressed	Specific considerations for PAT and real-time release testing models
Ongoing/lifecycle verification	Largely a one-time, pre-approval event	Continuous assurance across the method lifecycle, supporting post-approval change management

Published binary-mixture PCR/PLS models for sitagliptin and metformin hydrochloride have reached correlation coefficients above 0.999, recovery values of 97-100%, and %RSD below 2%, statistically indistinguishable from parallel HPLC reference assays (non-significant F-test/t-test comparisons) satisfying the core linearity, accuracy, and precision expectations carried into Q2(R2).[6] The core parameters and typical acceptance criteria are summarized in Table 6.

ICH Parameter	Definition	Typical Acceptance Criterion
Linearity	Ability of the method to produce results directly proportional to analyte concentration over a defined range	Correlation coefficient (R^2) $\geq$ 0.999
Accuracy	Closeness of agreement between the found value and the true/accepted value	Recovery within 98-102%
Precision (Repeatability & Intermediate Precision)	Degree of agreement among replicate measurements under the same/varied conditions	%RSD $\leq$ 2%
Specificity/Selectivity	Ability to unequivocally assess the analyte in the presence of excipients, degradants, or the co-formulated drug	No interference at analyte wavelength/band/emission maxima
LOD	Lowest analyte concentration that can be reliably detected	Signal-to-noise ratio $\approx$ 3:1
LOQ	Lowest analyte concentration that can be quantified with acceptable precision/accuracy	Signal-to-noise ratio $\approx$ 10:1
Robustness	Capacity to remain unaffected by small, deliberate variations in method parameters	No significant change in assay result
Range	Interval between upper and lower analyte concentrations demonstrated with precision, accuracy, and linearity	Derived from linearity/accuracy studies

9. Comparative Analysis: Multi-Modal Spectroscopy versus Chromatography

The main advantages claimed for the multi-spectroscopic chemometric framework over conventional RP-HPLC assays of metformin and sitagliptin are speed (spectral acquisition takes seconds to minutes, versus a multi minute chromatographic run per injection), cost (no column, guard column, or gradient-grade solvent consumables), and lower hazardous waste generation, since aqueous or solvent-free sample presentation replaces organic mobile phases.[4,6] Statistical comparison of chemometric spectrophotometric assays against USP-compliant HPLC reference methods for this drug pair has shown no significant difference in recovery (F-test, $p > 0.05$), suggesting that whatever accuracy trade-off exists is negligible for routine quality-control purposes.[6] HPLC's remaining edge and increasingly that of LC-MS/MS is its ability to resolve and separately quantify degradation products or process-related impurities chromatographically (Table 2); spectroscopic-chemometric methods are therefore best seen as complementary, rapid release/stability tools rather than wholesale replacements for stability-indicating impurity-profiling assays.[3,6,9]

10. Applications in Pharmaceutical Quality Control

10.1 Batch Release Testing

The short run time and minimal sample preparation of direct spectroscopic-chemometric assays make them well suited to high-throughput batch-release testing of metformin-sitagliptin tablets, where dozens of batches may need assay confirmation within a single shift.[4,6]

10.2 Stability Testing

Because FT-IR and UV-Vis methods can be applied directly to intact or minimally processed tablet samples, they fit well into periodic stability-testing protocols (accelerated and long-term ICH storage conditions), where frequent, low-cost assay points are needed without the consumable expense of repeated chromatographic runs.[12,14]

10.3 Process Analytical Technology (PAT) and AI-Enabled Digital Spectroscopy

The non-destructive, rapid character of vibrational and fluorescence spectroscopy especially when delivered through fibre optic or ATR probes lines up closely with the U.S. FDA's Process Analytical Technology (PAT) initiative for real-time, in-line, or at-line monitoring of blend uniformity and tablet content during manufacture, an application already shown for other tablet actives using fused multi-spectroscopic probe data.[14] This is accelerating: the 2025-2026 literature documents growing use of miniaturized/handheld NIR and FT-IR instruments, cloud-connected spectral databases, and AI/ML-based calibration transfer and explainable-AI tools that ease the burden of re-validating chemometric models across instruments and sites directly relevant to the model-transferability limitation noted in Section 13.[30-33]

10.4 Practical Industrial Implementation in Quality Control

Moving this framework from a literature-supported concept to routine industrial use means addressing capital cost (three spectrometers versus one HPLC system, though each individual instrument is cheaper than an HPLC-UV/PDA system), training analysts in chemometric software, integrating with existing Laboratory Information Management Systems (LIMS) for automated PCR/PLS/ML score reporting, and documenting an AQBd-based control strategy (Section 8) that a Qualified Person or regulatory inspector can audit. Contract manufacturers and generic-drug facilities running high volumes of a single fixed-dose combination are the most plausible early adopters, since the capital cost of parallel spectrometers is recovered fastest at high batch throughput; facilities with low batch volume or highly diversified product portfolios are less likely to see a favourable return on investment in the near term.

11. Research Gaps in Existing Analytical Methods

Pulling together the critical points raised in Sections 5, 6, 7, and 10, the following gaps in the current published literature on metformin-sitagliptin analysis stand out as priorities for future primary research:

- No published study has carried out and validated the full tri-modal (UV-Vis + FT-IR + fluorescence) data-fusion protocol specifically for the metformin-sitagliptin pair; the fusion evidence available so far comes from other drug matrices [14,19].
- No ANN/SVM or other ML-based calibration model has been published for this specific drug pair, despite how mature these tools already are in adjacent pharmaceutical spectroscopy work [30-33].
- No AGREE, GAPI, ComplexGAPI, Eco-Scale, or GEMAM score has been formally computed and published for any metformin-sitagliptin spectroscopic method identified in this search; greenness claims in the primary literature remain largely qualitative [6,12].
- Fingerprint-region FT-IR overlap between the two actives and common tablet excipients has not been systematically mapped for this matrix.
- No inter-laboratory or multi-site reproducibility study of any spectroscopic-chemometric method for this pair has been published, which limits the evidence base for regulatory transferability.
- No head-to-head comparison spanning all six techniques in Table 2 on one shared metformin-sitagliptin sample set exists.
- AQBd-structured (ATP/DoE/MODR) development, as distinct from post hoc "one-factor-at-a-time" optimization, has not yet been reported for this drug pair.

12. Challenges and Future Perspectives

- Model transferability: PCR/PLS/ML models are tied to their calibration set and need re-validation or transfer protocols whenever instruments, excipient batches, or formulations change; calibration-transfer algorithms and cloud-based spectral libraries are an active area of 2025-2026 research that may ease this burden.
- Excipient and polymorphic interference: certain excipients or drug polymorphic forms can introduce spectral features that call for careful wavelength/wavenumber region selection or extra pre-processing (baseline correction, derivatization, orthogonal signal correction).
- Regulatory familiarity: chromatographic methods remain the pharmacopoeial default, and broader regulatory acceptance of chemometric spectroscopic assays should benefit from the newly harmonized ICH Q2(R2)/Q14 framework, which for the first time gives multivariate model validation an explicit regulatory pathway [27-29].
- Data fusion complexity: combining UV-Vis, FT-IR, and fluorescence data streams requires careful synchronization, pre-processing, and fusion-level selection, and the added instrumentation cost of three techniques has to be weighed against the incremental accuracy gained over a single, well-optimized technique.
- Interpretability of ML models: black-box ANN/deep-learning models may draw more regulatory scrutiny than PLS unless paired with explainable-AI methods that identify which spectral regions are driving a prediction [32,33].

A brief bibliometric observation underscores how timely this framework is recent tool-development literature notes a steady year-on-year rise in publications applying AGREE, GAPI, and related greenness metrics, alongside rapidly growing 2024-2026 literature on AI/ML applications in vibrational and fluorescence spectroscopy [18-20,30-33]. A quantitative bibliometric/trend analysis (for example, a Scopus/Web of Science publication-count-by-year mapping) specific to metformin-sitagliptin analytical methods fell outside the scope of the search conducted here and is recommended as a separate follow-on scoping exercise rather than something asserted without the underlying database export.

Future work should prioritize: (i) inter-laboratory validation of fused multi-spectroscopic PCR/PLS/ML models for the metformin-sitagliptin pair; (ii) formal AGREE/GAPI/GEMAM benchmarking against pharmacopoeial HPLC methods, computed from a finalized protocol rather than inferred by analogy; (iii) AQbD-structured (ATP/DoE/MODR) method development as described in Section 8; and (iv) exploring portable/handheld spectroscopic instrumentation and AI-assisted calibration transfer to extend this green framework toward decentralized and point-of-release quality testing.

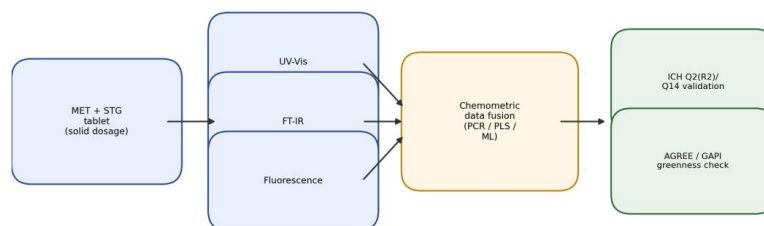
13. Conclusion and Recommendations

Bringing UV-Vis, FT-IR, and fluorescence spectroscopy together with PCR/PLS and emerging machine-learning chemometric modelling adds up to a scientifically sound, ICH Q2(R2)/Q14-compliant, and demonstrably greener alternative to chromatographic assay of the metformin-sitagliptin fixed-dose combination. Each spectroscopic technique has already been validated individually for this or closely related antidiabetic drug pairs, meeting regulatory expectations for linearity, accuracy, and precision; combined through data fusion, they hold out the prospect of even greater robustness against excipient interference and instrument-specific noise. Chromatography still has the edge for degradation-product profiling, but the evidence reviewed here supports adopting this multi-modal spectroscopic-chemometric framework for routine batch release, stability testing, and PAT applications, in step with the pharmaceutical industry's broader shift toward sustainable, Green Analytical Chemistry practice.

Based on the gaps identified in Section 12, three concrete research directions are recommended, in priority order: (1) experimentally executing and validating, to ICH Q2(R2) standards, the tri-modal data- fusion protocol specifically for metformin-sitagliptin tablets, with AGREE/GAPI/GEMAM scores computed and reported directly rather than inferred by analogy; (2) developing and cross-validating an ANN- or SVM-based calibration model for this pair, benchmarked against the classical PLS model on the same dataset; and (3) an AQbD-structured (ATP/DoE/MODR) method-development study that generates the robustness and transfer data needed for multi-site regulatory submission. Together, these three studies would turn the conceptual framework proposed in this review into a validated, publishable, and industrially deployable analytical procedure.

Graphical Abstract

A one-line graphical abstract summarizing the tablet-to-report workflow is provided below (Figure 5) for use in the journal submission system; authors should check sizing against the target journal's specific graphical-abstract template before final submission.



Graphical Abstract: Sustainable multi-spectroscopic chemometric framework for metformin-sitagliptin quantification.

Figure 5. Graphical abstract: sustainable multi-spectroscopic chemometric framework for metformin-sitagliptin quantification.

References

1. Ballav C, Gough SCL. Safety and efficacy of sitagliptin-metformin in fixed combination for the treatment of type 2 diabetes mellitus. *Clin Med Insights Endocrinol Diabetes*. 2013;6:23-31. <https://pubmed.ncbi.nlm.nih.gov/24031162/>
2. Chwieduk CM. Sitagliptin/metformin fixed-dose combination: in patients with type 2 diabetes mellitus. *Drugs*. 2011;71(3):349-361. <https://pubmed.ncbi.nlm.nih.gov/21319871/>
3. Mondal A, Naskar A, Siddiqi SS, et al. Efficacy and safety of fixed dose combination of sitagliptin, metformin, and pioglitazone in type 2 diabetes (IMPACT study): a randomized controlled trial. *Cardiovasc Diabetol Endocrinol Rep*. 2024;1:3. <https://pubmed.ncbi.nlm.nih.gov/38336931/>
4. Lakshmi KS, Lakshmi S. Design and optimization of a chemometric-assisted spectrophotometric determination of telmisartan and hydrochlorothiazide in pharmaceutical dosage form. *Indian J Pharm Sci*. 2010;72(5):574-578. <https://pubmed.ncbi.nlm.nih.gov/21331198/>
5. El-Bagary RI, Elkady EF, Ayoub BM. Spectrofluorometric and spectrophotometric methods for the determination of sitagliptin in binary mixture with metformin and ternary mixture with metformin and sitagliptin alkaline degradation product. *Int J Biomed Sci*. 2011;7(1):62-69. <https://pubmed.ncbi.nlm.nih.gov/23675222/>
6. Salim M, El-Enany N, Belal F, Walash M, Patonay G. Micelle-enhanced spectrofluorimetric method for determination of sitagliptin and identification of potential alkaline degradation products using LC-MS. *Luminescence*. 2014;29(7):837-845. <https://pubmed.ncbi.nlm.nih.gov/23559485/>
7. Azarian S, Shaghaghgi M, Dehghan G, Sheibani N. A rapid, simple and ultrasensitive spectrofluorimetric method for the direct detection of metformin in real samples based on a nanoquenching approach. *Luminescence*. 2021;36(3):658-667. <https://pubmed.ncbi.nlm.nih.gov/33185014/>
8. Khan N, et al. Environmentally friendly protocol for the determination of sitagliptin phosphate in pharmaceutical preparations and biological fluids using l-tyrosine as a fluorescence probe. *Luminescence*. 2023;38(9):1456-1465. <https://pubmed.ncbi.nlm.nih.gov/37519273/>
9. Fahelbom KMS, Saleh A, Mansour R, Sayed S. First derivative ATR-FTIR spectroscopic method as a green tool for the quantitative determination of diclofenac sodium tablets. *F1000Research*. 2020;9:172. <https://pubmed.ncbi.nlm.nih.gov/32226614/>
10. Remoto PJG, Gordon KC, Fraser-Miller SJ. A quantitative chemometric study of pharmaceutical tablet formulations using multi-spectroscopic fibre optic probes. *Pharmaceuticals (Basel)*. 2024;17(12):1659. <https://pubmed.ncbi.nlm.nih.gov/39770501/>
11. Rahman A, et al. Green analytical techniques for impurity determination in pharmaceuticals. *Electrophoresis*. 2026;47(1-2):45-58. <https://pubmed.ncbi.nlm.nih.gov/41351361/>
12. Pena-Pereira F, Wojnowski W, Tobiszewski M. AGREE-Analytical GREENness Metric Approach and Software. *Anal Chem*. 2020;92(14):10076-10082. <https://pubmed.ncbi.nlm.nih.gov/32538619/>
13. Yin L, Yu L, Guo Y, Wang C, Ge Y, Zheng X, Zhang N, You J, Zhang Y, Shi M. Green analytical chemistry metrics for evaluating the greenness of analytical procedures. *J Pharm Anal*. 2024;14(11):101013. <https://pubmed.ncbi.nlm.nih.gov/39759968/>
14. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. <https://pubmed.ncbi.nlm.nih.gov/33577987/>
15. Mansour FR, Omer KM, Plotka-Wasyłka J. A total scoring system and software for complex modified GAPI (ComplexMoGAPI) application in the assessment of method greenness. *Green Anal Chem*. 2024;10:100126. https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Mansour+FR%2C+Omer+KM%2C+P%2C5%82otka-Wasyłka+J.+A+total+scoring+system+and+software+for+complex+modified+GAPI+%28ComplexMoGAPI%29+application+in+the+assessment+of+method+greenness.+Green+Anal+Chem.+2024%3B10%3A100126.&btnG=
16. Xin T, Yu L, Zhang W, Guo Y, Wang C, Li Z, You J, Xue H, Shi M, Yin L. Greenness evaluation metric for analytical methods and software. *J Pharm Anal*. 2025;101202. <https://pubmed.ncbi.nlm.nih.gov/40756999/>
17. Sängers-van de Griend CE. A practical approach to implementing ICH Q14: tools for analytical quality by design in capillary electrophoresis method development. *Electrophoresis*. 2025;46(13-14):933-953. <https://pubmed.ncbi.nlm.nih.gov/40099702/>
18. Contreras J, Bocklitz T. Explainable artificial intelligence for spectroscopy data: a review. *Pflügers Arch*. 2025;477:603-615. <https://pubmed.ncbi.nlm.nih.gov/39088045/>
19. Elnadi S, Abdalsabour S, Farouk M, Trabik YA. Fourier transform infrared spectroscopic, spectrofluorimetric assays of canagliflozin, and stability-indicating UV-spectrophotometric method for the simultaneous determination of canagliflozin and metformin. *J AOAC Int*. 2022;105(4):964-971. <https://pubmed.ncbi.nlm.nih.gov/35348721/>
20. Ramzy S, et al. Smart spectrophotometric methods for the simultaneous determination of newly co-formulated hypoglycemic drugs in binary mixtures. *Spectrochim Acta A Mol Biomol Spectrosc*. 2025;325:125123. <https://pubmed.ncbi.nlm.nih.gov/33872950/>
21. Attimarad M, Nair AB, Sreeharsha N, Al-Dhubiab BE, Venugopala KN, Shinu P. Development and validation of green UV derivative spectrophotometric methods for simultaneous determination of metformin and remogliflozin from formulation: evaluation of greenness. *Int J Environ Res Public Health*. 2021;18(2):448. <https://pubmed.ncbi.nlm.nih.gov/33429964/>