ISSN: 2231-3656

International Journal of Pharmacy and Industrial Research (IJPIR)

IJPIR |Vol.15 | Issue 1 | Jan - Mar-2026

www.ijpir.com

DOI: https://doi.org/10.61096/ijpir.v15.iss1.2026.289-295

AI-Assisted Detection and Quantification of Remdesivir Impurities Using RP-HPLC

S.Venuharini[a*], S. Hameedul Fahim[a], S. Jai surya[a], P. Shek abdullah[a], S. Varshini[a], Dr. Ravisankar Mathesan[a], Dr. Nataraj Palaniyappan[b]

[a*,a] Srinivasan College of Pharmaceutical Sciences, Trichy

[b] Scientist, Novitium Pharma LLC, New jersey, USA.

*Corresponding Author: S. Venuharini

Mail Id: venuharini254@gmail.com

Published by: 20.03.2026

Abstract: Ensuring the purity of pharmaceutical drugs is essential to maintain their safety, quality, and therapeutic effectiveness. Remdesivir is a nucleotide analogue antiviral drug used for the treatment of severe viral infections. During synthesis, formulation, and storage, several process-related impurities, degradation products, and metabolites may form. Impurities such as GS-441524 and GS-704277 must be carefully monitored because their presence may influence the safety, stability, and efficacy of the drug product. Conventional analytical methods such as reverse-phase high-performance liquid chromatography (RP-HPLC) are widely used for impurity profiling; however, these methods typically rely on manual peak detection and integration, which can be time-consuming and susceptible to human error, particularly at low impurity levels. This work proposes an artificial intelligence (AI)–assisted approach for the detection and quantification of remdesivir impurities using RP-HPLC. In the proposed framework, chromatographic data obtained from impurity and degradation studies are processed using machine learning algorithms capable of automated peak detection, baseline correction, and peak integration. By analysing chromatographic features such as retention time, peak area, and spectral characteristics, AI models can estimate impurity concentrations with improved efficiency. This conceptual approach demonstrates the potential of AI to enhance impurity profiling, reduce manual data processing, and support advanced pharmaceutical quality control.

Keywords: Remdesivir, Artificial Intelligence, RP-HPLC, Impurity Profiling, Machine Learning

INTRODUCTION

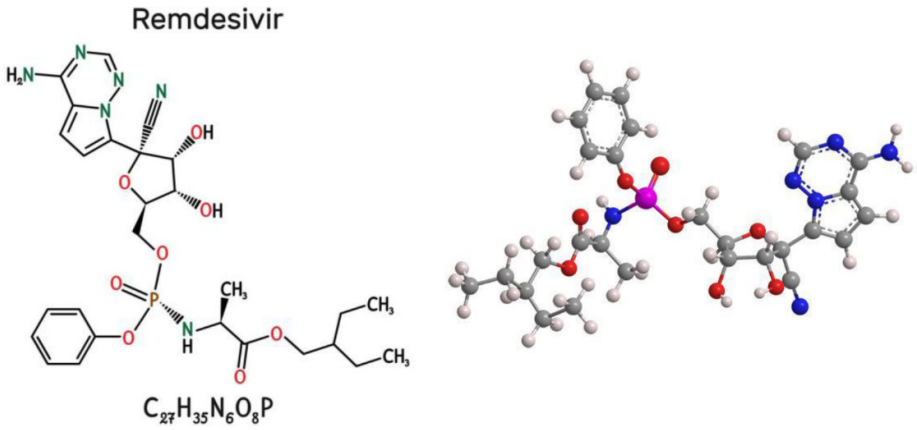

Fig 1: Structure of Remdesivir

Impurity profiling is a critical aspect of pharmaceutical quality assurance because impurities present in drug substances and drug products may affect their safety and therapeutic efficacy. During synthesis, formulation, and storage, drugs may form process-related impurities, degradation products, intermediates, and residual contaminants.[1] Regulatory authorities require the identification and control of these impurities to ensure drug safety and quality. Remdesivir has gained global attention as an antiviral agent used in the treatment of severe viral infections.[2] Due to its complex chemical structure and synthetic pathway, several potential impurities may arise during manufacturing and storage. Therefore, sensitive and reliable analytical techniques are required for the detection and quantification of these impurities.[3] Chromatographic techniques such as reverse-phase high-performance liquid chromatography (RP-HPLC) are widely used in pharmaceutical analysis because of their high resolution, reproducibility, and sensitivity. However, conventional chromatographic data processing relies heavily on manual peak integration and interpretation, which may lead to errors in impurity detection, particularly at low concentrations.[4] Recent advancements in artificial intelligence (AI) and machine learning have enabled automated analysis of complex analytical data. AI algorithms can identify patterns, detect peaks, and quantify analytes more efficiently than manual approaches. By integrating AI with chromatographic analysis, it is possible to enhance impurity detection accuracy and improve analytical efficiency. The objective of the present study is to develop and validate an AI-assisted RP-HPLC method for the detection and quantification of impurities in remdesivir, thereby providing an advanced analytical approach for pharmaceutical quality control.[5]

CLASSIFICATION OF PROCESS, DEGRADATION, AND METABOLITE-RELATED IMPURITIES IN REMDESIVIR

During the synthesis, formulation, and storage of remdesivir, several process-related impurities, intermediates, degradation products, and metabolites may be present. These impurities must be controlled according to guidelines such as ICH Q3A and ICH Q3B.[6]

Below are the major impurities reported for remdesivir.

1. Process-Related Impurities

These arise during chemical synthesis of remdesivir.

Table 1: Process Related Impurities[7]

Impurity	Description
GS-441524	Nucleoside analogue formed by hydrolysis of Remdesivir
GS-704277	Alanine metabolite intermediate
Phosphoramidate intermediate impurity	Related to phosphoramidate prodrug moiety
Protected nucleoside intermediate	Residual intermediate from synthesis
Unreacted starting nucleoside	Remaining starting material in synthesis

2. Degradation Impurities

These form during stress conditions or storage.

Table 2: Degradation Impurities[8]

Impurity	Cause
Hydrolysis impurity	Acidic or basic conditions
Oxidative impurity	Oxidation with peroxide or oxygen
Photolytic impurity	Exposure to UV or light
Thermal degradation impurity	High temperature

3. Metabolite-Related Impurities

Table 3: Metabolite-Related Impurities

Metabolite	Description
GS-441524	Major active metabolite
GS-704277	Intermediate metabolite

These compounds may also appear during stability studies or degradation experiments.

4. Potential Elemental Impurities

Elemental impurities may originate from catalysts used during synthesis.

Examples include:

- Palladium (Pd)
- Platinum (Pt)
- Copper (Cu)
- Iron (Fe)

These are controlled according to ICH Q3D.

In impurity profiling, AI uses the physicochemical properties of each impurity (such as polarity, molecular weight, UV absorption, fragmentation pattern, etc.) to detect peaks, identify compounds, and calculate their concentration from chromatographic data.[9]

QUANTITATION USING AI

1. Input Properties of Each Impurity

AI models learn from the chemical and analytical properties of impurities.[10]

Example impurities in Remdesivir

- GS-441524
- GS-704277
- Phosphoramidate intermediate
- Hydrolysis impurity
- Oxidative impurity

Table 4: Key properties used by AI

Property	Influence in chromatography
Molecular weight	Determines MS signal
Polarity / logP	Controls retention time
pKa	Determines ionization
UV absorption wavelength	Peak detection in HPLC
Fragmentation pattern	Identification in LC-MS

2. Data Collection for AI Quantitation

Experimental chromatographic data are used to train the AI model.[11]

Table 5: Typical dataset for AI Model

Impurity	Concentration (µg/mL)	Retention time	Peak area	m/z
GS-441524	0.2	5.1	345000	292
GS-704277	0.2	7.4	289000	443
Hydrolysis impurity	0.1	6.5	152000	310

AI learns the relationship between:

Peak signal → Impurity concentration

3. AI Algorithms Used for Quantitation

a) Artificial Neural Network (ANN)

ANN models predict impurity concentration from chromatographic features.

Inputs:

- Retention time
- Peak area
- Peak width
- UV spectra

Output:

- Predicted concentration.

Ex: Peak area + retention time → ANN → impurity concentration.[12]

b) Random Forest Regression

Used for accurate impurity quantitation.

Features used:

- Peak area
- Peak height
- Spectral intensity
- Mass fragmentation data.

Output: Concentration prediction.

Advantage: Handles complex chromatograms.[13]

c) Support Vector Regression (SVR)

Used for low-level impurity quantitation. Best for: Impurities below 0.1% level.[14]

4. AI Peak Detection and Integration

Before quantitation, AI identifies chromatographic peaks.

Deep learning model (CNN)

Tasks:

- Detect peaks automatically
- Separate overlapping peaks
- Remove noise from chromatogram.

Example:

Traditional integration error

↓

AI identifies real peak boundaries

↓

Accurate peak area calculation.

5. Calibration Model Using AI

- Traditional method:
- Linear regression calibration curve.
- AI method:
- Machine learning calibration model.

Table 6: Example dataset

Conc. (µg/mL)	Peak Area
0.1	150000
0.2	300000
0.3	450000

AI learns nonlinear relationships between signal and concentration.

6. AI Quantitation Workflow

Step 1

Collect HPLC or LC-MS data

Step 2

Preprocess data

- Noise removal
- Baseline correction.

Step 3

AI peak detection

Step 4

Peak classification: Identify impurity.

Step 5

Quantitation model: AI predicts concentration using trained regression model.

7. Example AI Quantitation Output [15-20]

Table 7: AI Quantitation Output

Impurity	Predicted concentration (%)
GS-441524	0.15
GS-704277	0.08
Hydrolysis impurity	0.05

AI-BASED CHROMATOGRAPHIC DATA PROCESSING

Artificial intelligence (AI) has significantly improved the processing and interpretation of chromatographic data obtained from analytical techniques such as RP-HPLC and LC-MS. In impurity profiling studies of drugs like Remdesivir and its related substances including GS-441524 and GS-704277, large volumes of chromatographic data are generated.[21] Conventional data processing relies on manual peak detection and integration, which may introduce variability and errors, especially when detecting trace impurities. AI-based chromatographic data processing enables automated, accurate, and efficient analysis of chromatographic signals.[22]

1. Data Acquisition

The first step in AI-based chromatographic processing involves acquiring raw chromatographic data from the RP-HPLC system. The chromatogram typically contains information such as retention time, detector response, and peak shape. This raw data forms the input for AI-based analysis.[23]

2. Data Preprocessing

Before applying machine learning algorithms, the chromatographic data undergo preprocessing to improve data quality. This step includes noise filtering, baseline correction, and signal smoothing.[24] Techniques such as digital filtering and polynomial baseline correction are applied to remove background noise and improve signal clarity. Proper preprocessing ensures that

the AI model can accurately identify chromatographic peaks.[25]

3. Automated Peak Detection

One of the most important steps in chromatographic data processing is peak detection. AI algorithms, particularly deep learning models such as Convolutional Neural Networks (CNN), are capable of automatically detecting peaks in chromatograms.[26] These algorithms analyze signal patterns to identify peak boundaries, even when peaks are small or partially overlapping. Automated peak detection significantly reduces manual interpretation and improves detection sensitivity.[27]

4. Peak Integration

After peak detection, AI algorithms perform peak integration to calculate the area under each peak. The peak area is directly related to the concentration of the analyte or impurity. AI-based peak integration identifies accurate start and end points of peaks, resulting in more reliable quantitation compared to manual integration methods.[28]

5. Peak Classification and Identification

AI models can also classify detected peaks based on their characteristics such as retention time, UV spectra, and mass spectral data. Machine learning algorithms such as Random Forest and Support Vector Machines can be trained using known impurity data to identify specific compounds. This enables automated identification of process-related impurities, degradation products, and metabolites.[29]

6. Data Output and Quantitative Analysis

Once peaks are detected and classified, the AI system generates quantitative outputs including peak area, retention time, and predicted concentration of impurities. These results are used for impurity quantification and quality assessment of the drug substance.[30] AI-assisted analysis improves reproducibility and reduces the risk of human error in chromatographic data interpretation.[31]

7. Advantages of AI-Based Chromatographic Processing

AI-based chromatographic data processing offers several advantages in pharmaceutical analysis. It enables rapid analysis of complex chromatograms, improves the detection of trace impurities, and reduces manual intervention.[32] Additionally, AI algorithms can learn from large datasets and continuously improve their performance, making them valuable tools for advanced analytical method development and pharmaceutical quality control.[33-36]

CONCLUSION

This proposal highlights the potential application of artificial intelligence (AI) in the detection and quantification of impurities in Remdesivir using RP-HPLC. Impurity profiling is an essential requirement in pharmaceutical quality control, as trace levels of process-related impurities, degradation products, and metabolites such as GS-441524 and GS-704277 may affect the safety, stability, and therapeutic efficacy of the drug substance. Conventional chromatographic analysis often relies on manual peak detection and integration, which can be time-consuming and subject to human variability. The proposed approach integrates machine learning algorithms with RP-HPLC chromatographic analysis to enable automated peak detection, baseline correction, peak integration, and impurity quantification. By utilizing chromatographic features such as retention time, peak area, and spectral information, AI models such as Artificial Neural Networks, Random Forest Regression, and Support Vector Regression can be trained to identify and quantify impurities more efficiently than traditional manual methods. The integration of AI with chromatographic data processing offers several potential advantages, including improved sensitivity in detecting trace impurities, enhanced reproducibility, reduced analysis time, and minimized human error. Although the present work outlines a conceptual and methodological framework rather than experimental validation, it demonstrates how AI-assisted analytical techniques could significantly improve impurity profiling of remdesivir.

Overall, the proposed AI-assisted RP-HPLC approach represents a promising strategy for modern pharmaceutical analysis. With further experimental validation and model training using real chromatographic datasets, this approach could support more efficient impurity monitoring, stability assessment, and

quality control of remdesivir and related pharmaceutical compounds.

REFERENCES

1. LeCun Y, Bengio Y and Hinton G. Deep learning. Nature. 2015;521(7553):436–444.
2. Goodfellow I, Bengio Y and Courville A. Deep learning methods for data analysis and pattern recognition. MIT Press. 2016;1(1):1–775.
3. Wishart DS. Advances in metabolite identification using artificial intelligence. Trends Anal Chem. 2019;120:115324.
4. Bro R and Smilde AK. Principal component analysis. Anal Methods. 2014;6(9):2812–2831.
5. Brown SD and Tauler R. Chemometrics in analytical chemistry. Anal Chem. 2019;91(1):1–17.
6. Gromski PS, Muhamadali H and Ellis DI. A tutorial review: Metabolomics and partial least squares-discriminant analysis. Anal Chim Acta. 2015;879:10–23.
7. Camacho J, Picó J and Ferrer A. Data understanding with PCA: Structural and variance information plots. Chemom Intell Lab Syst. 2010;100(1):48–56.
8. Berrueta LA, Alonso-Salces RM and Héberger K. Supervised pattern recognition in food analysis. J Chromatogr A. 2007;1158(1–2):196–214.
9. Xu R and Wang Q. Large-scale extraction of adverse drug events from clinical notes using natural language processing methods. J Am Med Inform Assoc. 2014;21(5):848–855.
10. Johnson SR and Lange BM. Open-access metabolomics databases for natural product research. Nat Prod Rep. 2015;32(3):447–457.
11. Debnath S and Haldar S. Application of machine learning in chromatographic peak detection. J Chromatogr A. 2020;1621:461089.
12. Kalivas JH. Chemometrics for chromatographic data analysis. J Chromatogr A. 2012;1261:1–12.
13. Hopfgartner G and Varesio E. Artificial intelligence in LC–MS data analysis. Trends Anal Chem. 2019;120:115310.
14. Patel BN, Rosenberg L and Willcox G. Human–machine partnership with artificial intelligence for clinical decision support. NPJ Digit Med. 2019;2(1):1–9.
15. Topol EJ. High-performance medicine: The convergence of human and artificial intelligence. Nat Med. 2019;25(1):44–56.
16. Boobier S, Osbourn A and Vickerstaff R. Machine learning in pharmaceutical discovery and development. Nat Rev Drug Discov. 2017;16(6):417–428.
17. Sheridan RP. Using random forest to model the domain applicability of chemical data. J Chem Inf Model. 2013;53(11):2837–2850.
18. Goh GB, Siegel C and Vishnu A. Deep learning for computational chemistry. J Comput Chem. 2017;38(16):1291–1307.
19. Wang J, Wang S and Song H. Applications of artificial intelligence in analytical chemistry. Trends Anal Chem. 2020;122:115733.
20. Kotsiantis SB. Supervised machine learning: A review of classification techniques. Informatica. 2007;31(3):249–268.
21. Cortes C and Vapnik V. Support-vector networks. Mach Learn. 1995;20(3):273–297.
22. Breiman L. Random forests. Mach Learn. 2001;45(1):5–32.
23. Bishop CM. Pattern recognition and machine learning. Springer. 2006;1(1):1–738.
24. Warren TK, Jordan R and Lo MK. Therapeutic efficacy of the small molecule GS-5734 against Ebola virus. Nature. 2016;531(7594):381–385.
25. Eastman RT, Roth JS and Brimacombe KR. Remdesivir: A review of its discovery and development. ACS Cent Sci. 2020;6(5):672–683.
26. Gordon CJ, Tchesnokov EP and Feng JY. The antiviral compound remdesivir inhibits RNA-dependent RNA polymerase. J Biol Chem. 2020;295(15):4773–4779.
27. Beigel JH, Tomashek KM and Dodd LE. Remdesivir for the treatment of COVID-19. N Engl J Med. 2020;383(19):1813–1826.
28. Avataneo V, de Nicolò A and Cusato J. Development and validation of a LC–MS/MS method for remdesivir quantification. J Antimicrob Chemother. 2020;75(7):1772–1777.
29. Kaur H and Singh B. Stability indicating RP-HPLC method for determination of remdesivir. J Pharm Biomed Anal. 2021;196:113935.

30. Mishra P and Singh R. Analytical methods for impurity profiling in pharmaceuticals. J Pharm Biomed Anal. 2018;147:357–371.
31. Chen H, Engkvist O and Wang Y. The rise of deep learning in drug discovery. Drug Discov Today. 2022;27(1):1–9.
32. Vamathevan J, Clark D and Czodrowski P. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov. 2022;21(3):175–196.
33. Rodrigues T and Bernardes GJL. Machine learning for target discovery in drug development. Nat Chem Biol. 2022;18(4):343–352.
34. Xu Y, Dai Z and Chen T. Artificial intelligence in analytical chemistry: Trends and future perspectives. Trends Anal Chem. 2022;153:116648.
35. Smith R and Jones K. Machine learning approaches for chromatographic peak detection and integration. Anal Chem. 2022;94(14):5678–5686.
36. Li Y, Zhang X and Wang J. Deep learning-based analysis of LC–MS data for metabolite identification. Anal Chim Acta. 2023;1245:340567.