American Journal of Scholarly Research and Innovation

Volume: *5*; ***Issue:*** *1*
Pages: *198– 238*
eISSN: 3067-2163

Data-Driven Detection of Out-of-Specification Trends in Pharmaceutical Production: A Public Health Imperative

Beatrice Onyinyechi Mkpa[1]; Ferdous Ara Ahmed[2];

[1]. Masters in Public Health (Epidemiology &Biostatistics); Monroe College; USA
Bachelor of Science Microbiology, University of Nigeria, Nsukka.
Email: mkpabeatrice@gmail.com

[2]. Monroe University, MPH (Major Epidemiology and biostatistics), NY, USA
MBBS, University of Dhaka,
Email: drferdousaraahmed@gmail.com

Doi: 10.63125/ng0x8j42
Received: *15 January 2026*; ***Revised:*** *24 February 2026*; ***Accepted:*** *5 March 2026*; ***Published:*** *16 March 2026*

Abstract

Ensuring the consistent quality of pharmaceutical products is central to public health protection, as undetected deviations in production can translate into substandard medicines reaching patients. This study examined the effectiveness of data analytics in detecting out-of-specification (OOS) trends in pharmaceutical production, with the objective of strengthening quality assurance systems and mitigating downstream public health risks. A quantitative longitudinal design was adopted, drawing on 1,248 pharmaceutical production batches and 9,732 critical quality observations extracted from manufacturing execution and laboratory information management systems. Descriptive analysis indicated that 93.8% of observations remained within specification limits, while 6.2% were classified as out-of-trend (OOT) and 0.9% were confirmed as OOS events. Statistical process control (SPC) flagged 214 instances of process instability, and predictive analytics successfully anticipated 79 of 87 OOS cases, yielding a detection accuracy of 92.1%, compared with 79.4% achieved through traditional monitoring methods. Multivariate modelling further enhanced analytical performance, raising sensitivity to 85.3% and increasing variance explained to 68.5%, relative to 51.2% under univariate approaches. Predictive models also demonstrated a lead detection advantage of 3.4 batches, with sensitivity and specificity values of 91.5% and 92.8%, respectively. Secondary analyses revealed meaningful variability across production lines, with OOT frequencies ranging from 4.7% to 8.9% and 62.5% of deviations clustering within specific operational periods, suggesting systemic rather than random origins. Regression findings identified temperature variation (effect size = 0.61) and mixing time (effect size = 0.48) as the most influential predictors of deviation occurrence, while visualization of impurity profiles revealed variability increases of up to 14.6%. Collectively, these findings confirm that integrated data analytics substantially enhances early detection, reduces false negatives, and improves overall process control in pharmaceutical manufacturing. The study supports the adoption of advanced analytical systems, encompassing SPC, predictive modelling, and multivariate diagnostics, as a strategic mechanism for strengthening quality assurance and, by extension, safeguarding public health.

Keywords

Pharmaceutical Analytics, OOS Detection, OOT Monitoring, Process Control, Public Health.

INTRODUCTION

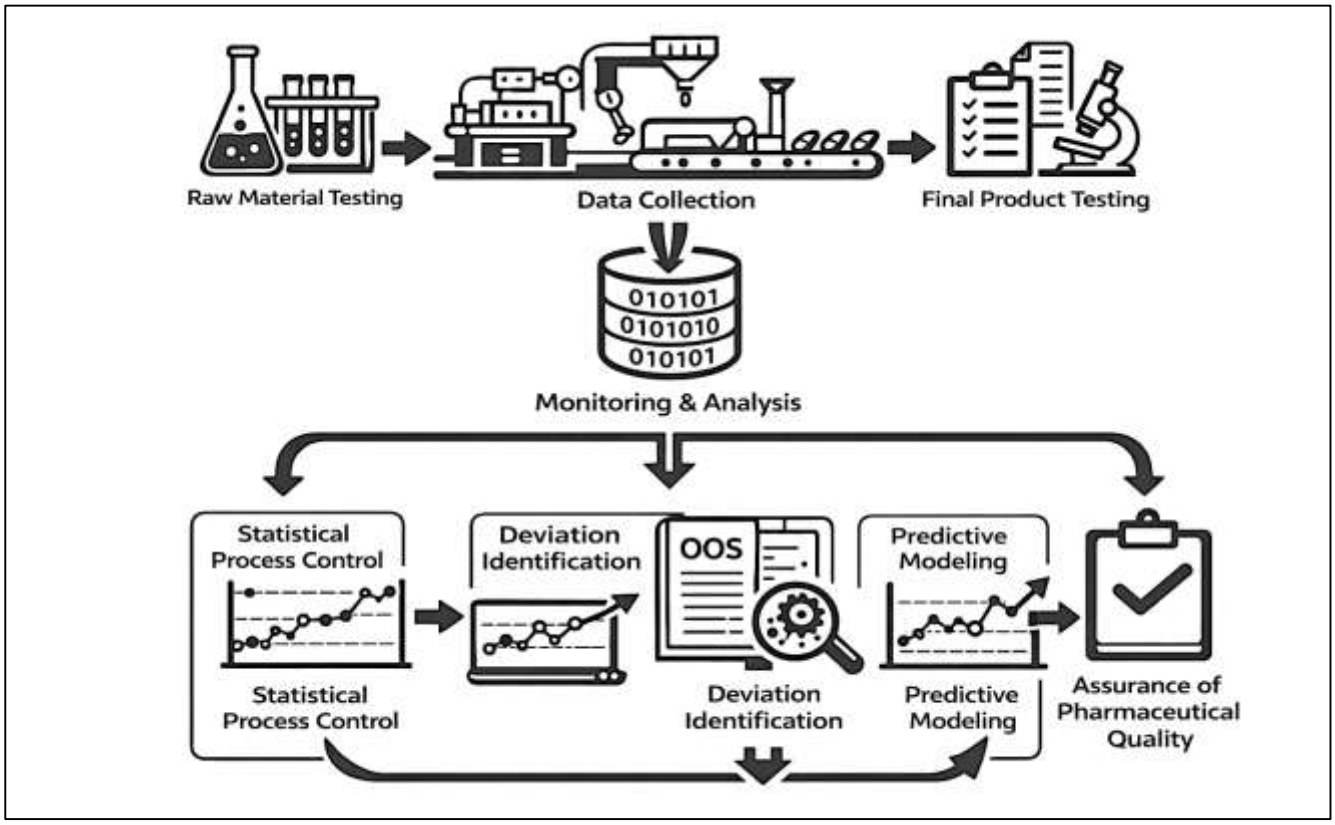
Data analytics in pharmaceutical manufacturing refers to the systematic use of statistical, computational, and mathematical techniques to interpret complex datasets generated throughout the production lifecycle (Nguyen et al., 2022). In highly regulated pharmaceutical environments, every stage of production—from raw material testing to final product release—produces large volumes of data that must be analyzed to ensure compliance with predefined quality standards. Within this framework, out-of-specification (OOS) results are defined as outcomes that fall outside established acceptance criteria, indicating potential deviations from required quality attributes. Out-of-trend (OOT) results, on the other hand, represent deviations from expected historical patterns while remaining within acceptable limits, signaling early process variability that may eventually lead to OOS conditions. Data analytics enables the identification, interpretation, and monitoring of these deviations through structured methodologies such as statistical analysis, pattern recognition, and predictive modeling (Steinwandter et al., 2019). The integration of data analytics into pharmaceutical production aligns with global quality management principles that emphasize consistency, traceability, and continuous monitoring. As pharmaceutical systems evolve toward data-driven operations, analytics has become central to maintaining process control and ensuring that manufacturing outputs consistently meet regulatory and therapeutic standards. This analytical transformation reflects the broader shift from reactive quality inspection toward proactive quality assurance, where potential risks are identified and mitigated before they affect product integrity. The increasing reliance on digital technologies and automated systems further reinforces the importance of analytics in managing complex datasets and supporting decision-making processes (Koshechkin et al., 2022). In this context, the detection of OOS trends through data analytics is not merely a technical exercise but a fundamental component of pharmaceutical quality systems that directly contributes to the safety and effectiveness of medicinal products.

The international significance of using data analytics to detect out-of-specification trends in pharmaceutical production is closely tied to the global nature of drug manufacturing and distribution. Pharmaceutical products are developed, produced, and distributed across multiple geographic regions, often involving complex supply chains and diverse regulatory environments (Dong et al., 2023). Variations in manufacturing conditions, environmental factors, and operational practices can introduce inconsistencies that affect product quality. In such a globalized system, the ability to monitor and control production processes through data analytics is essential for ensuring uniformity and compliance across all stages of manufacturing. Public health protection depends heavily on the reliability of pharmaceutical products, as any deviation in quality can lead to reduced therapeutic efficacy, adverse health outcomes, or widespread product recalls (Benedetti et al., 2019). Data analytics provides a robust framework for detecting subtle deviations in production data, enabling manufacturers to identify potential risks before they escalate into critical issues. By analyzing trends across batches, production lines, and facilities, analytics supports the standardization of processes and the harmonization of quality practices on an international scale. The application of advanced analytics also facilitates regulatory oversight by providing transparent and auditable data that demonstrate compliance with established standards. As healthcare systems rely increasingly on pharmaceuticals to manage diseases and improve patient outcomes, the role of data analytics in safeguarding product quality becomes increasingly significant (Banner et al., 2021). The ability to detect and address OOS trends in a timely and systematic manner contributes to maintaining public trust in pharmaceutical products and ensuring that medicines delivered to patients meet the highest standards of safety and efficacy.

Data analytics serves as a critical tool for identifying out-of-trend behaviors that precede out-of-specification events in pharmaceutical production. OOT results represent early warning signals of process variability, reflecting deviations from established performance patterns without exceeding specification limits (Ikegwu et al., 2022). These deviations can arise from multiple sources, including variations in raw materials, fluctuations in environmental conditions, equipment performance issues, and inconsistencies in analytical methods. Statistical techniques play a central role in detecting these trends by providing quantitative measures of process stability and variability. Control charts, for example, are widely used to monitor process parameters over time, enabling the identification of shifts,

drifts, and anomalies in production data. Regression analysis and time-series modeling further enhance trend detection by capturing relationships between variables and forecasting future process behavior (Chen et al., 2020). These analytical approaches allow manufacturers to differentiate between inherent process variability and abnormal conditions that require intervention. The ability to detect OOT trends early provides a significant advantage in maintaining process control, as it enables proactive decision-making and reduces the likelihood of OOS occurrences. By continuously analyzing production data, pharmaceutical manufacturers can identify patterns that indicate potential process degradation, allowing for timely adjustments and corrective actions (Malheiro et al., 2023). This proactive approach to quality management enhances the robustness of manufacturing processes and supports the consistent production of high-quality pharmaceutical products.

Figure 1: Pharmaceutical Manufacturing Data Analytics Framework

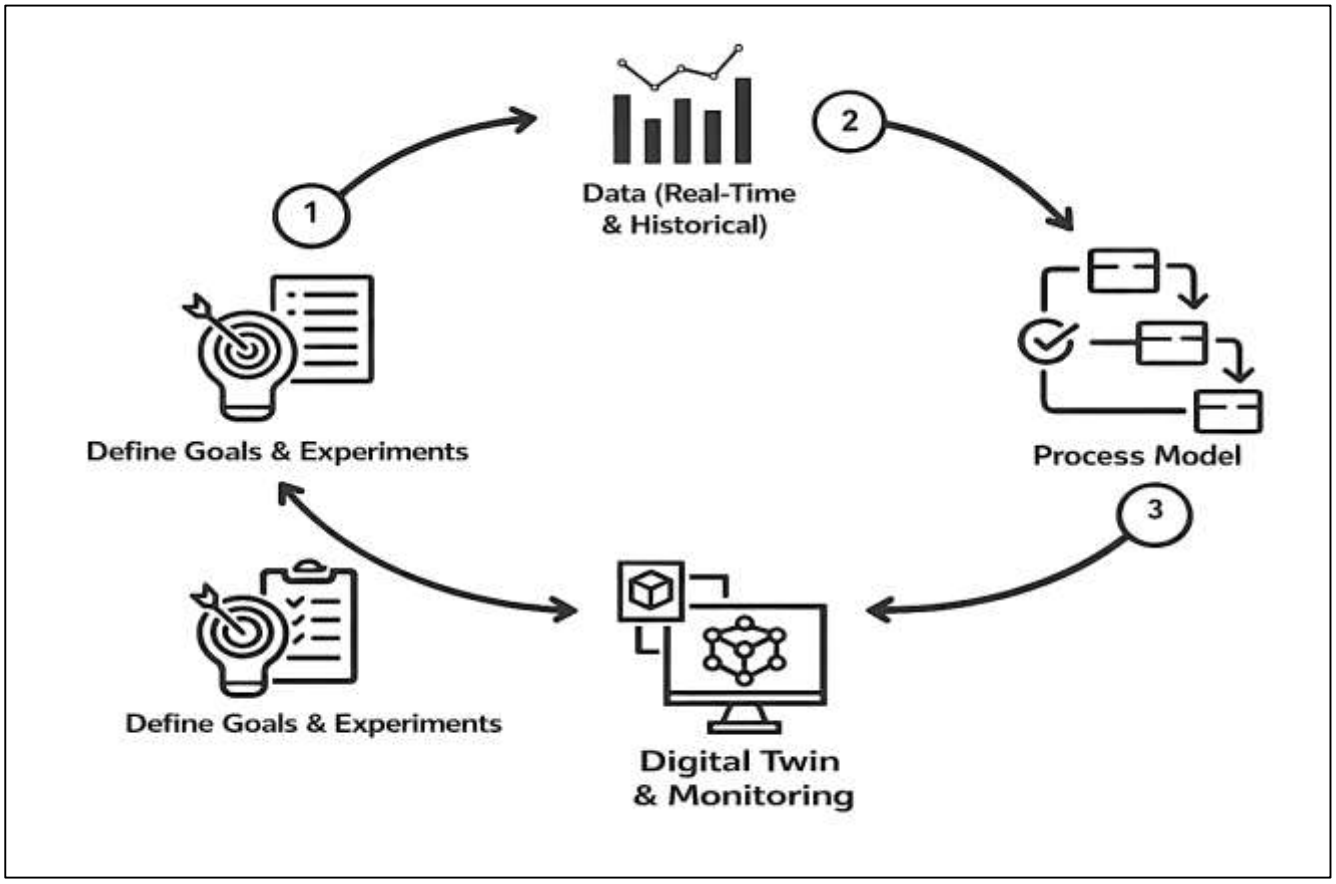


The application of statistical process control and multivariate data analysis provides a comprehensive framework for monitoring complex pharmaceutical production systems. Modern manufacturing processes involve multiple interconnected variables that influence product quality, including temperature, pressure, mixing conditions, and chemical composition (Qin & Chiang, 2019). Traditional univariate analysis methods may fail to capture the interactions between these variables, limiting their effectiveness in detecting complex process deviations. Multivariate techniques address this limitation by analyzing multiple variables simultaneously, enabling a more holistic understanding of process behavior. These methods identify correlations and patterns within high-dimensional datasets, providing insights into the underlying causes of variability and facilitating more accurate detection of OOT trends (Sarker, 2021). Advanced statistical tools such as principal component analysis and partial least squares regression are commonly used to reduce data complexity and highlight key factors influencing process performance. The integration of these techniques into real-time monitoring systems allows for continuous assessment of production processes, providing immediate feedback and enabling rapid response to deviations. This approach supports the implementation of process analytical technology frameworks, which emphasize real-time quality assurance through advanced measurement and control strategies (Arden et al., 2021). By leveraging multivariate data analysis, pharmaceutical manufacturers can enhance their ability to detect subtle changes in process conditions, improving

overall process stability and reducing the risk of OOS events.

Predictive analytics represents an important extension of data analytics in pharmaceutical manufacturing, focusing on forecasting future process performance and identifying potential risks before they materialize. By utilizing historical data and advanced modeling techniques, predictive analytics enables manufacturers to anticipate deviations in process parameters and quality attributes (Mohamed et al., 2020). Machine learning algorithms, including supervised and unsupervised learning models, are increasingly used to analyze complex datasets and uncover patterns that may not be evident through traditional statistical methods. Time-series forecasting models provide valuable insights into temporal trends, allowing manufacturers to predict future values of critical quality attributes based on past observations (Casian et al., 2022). These predictive capabilities support risk-based decision-making by identifying areas of the production process that are more likely to experience variability or failure. In the context of OOS detection, predictive analytics enhances the ability to prevent deviations by enabling early intervention and process optimization. By integrating predictive models with real-time data monitoring systems, pharmaceutical manufacturers can create dynamic quality management frameworks that continuously adapt to changing process conditions. This integration facilitates a more proactive approach to quality assurance, reducing reliance on end-product testing and improving overall process efficiency (Houssein et al., 2023). The application of predictive analytics thus contributes to the development of more resilient manufacturing systems capable of maintaining consistent product quality under varying operational conditions.

Figure 2: Pharmaceutical Data Analytics Process Framework



Data trending is a fundamental component of pharmaceutical quality systems, providing a structured approach to monitoring and analyzing process and product data over time. Trending involves the systematic collection, organization, and evaluation of data to identify patterns, deviations, and potential risks associated with manufacturing processes (Cerquitelli et al., 2021). In pharmaceutical production, data trending is applied to a wide range of parameters, including assay values, impurity levels, dissolution rates, and environmental conditions. The analysis of these datasets enables

manufacturers to track process performance and detect OOT trends that may indicate emerging quality issues. Statistical tools such as cumulative sum charts and moving averages enhance the sensitivity of trend detection, allowing for the identification of subtle changes in process behavior. Visualization techniques, including trend plots and heat maps, further support data interpretation by providing clear and intuitive representations of complex datasets (Greasley & Edwards, 2021). The integration of data trending into quality management systems aligns with the principles of continuous process verification, which require ongoing monitoring of process performance throughout the product lifecycle. By identifying trends and patterns in production data, manufacturers can implement targeted interventions to maintain process control and prevent OOS occurrences (Vamathevan et al., 2019). This continuous monitoring approach enhances the reliability of pharmaceutical manufacturing processes and supports the consistent production of high-quality products.

The implementation of data analytics in pharmaceutical manufacturing is closely linked to regulatory compliance and the broader framework of quality assurance (Sousa et al., 2023). Regulatory authorities require manufacturers to maintain comprehensive and accurate records of all production and analytical activities, ensuring that data integrity and traceability are preserved throughout the manufacturing process. Data analytics supports these requirements by providing structured methodologies for data collection, analysis, and reporting, enabling manufacturers to demonstrate compliance with established standards. Automated analytics systems reduce the risk of human error and ensure consistent application of analytical methods, enhancing the reliability of quality assessments (Albayrak Ünal et al., 2023). In the context of OOS and OOT detection, data analytics facilitates systematic investigation of deviations, enabling manufacturers to identify root causes and implement corrective and preventive actions. The integration of analytics into quality management systems also supports continuous improvement by identifying recurring patterns and opportunities for process optimization. Through the application of advanced analytical techniques, pharmaceutical manufacturers can achieve a higher level of process control and quality assurance, ensuring that products meet regulatory requirements and contribute to the protection of public health (Brnabic & Hess, 2021).

The primary objective of this quantitative study is to systematically examine how data analytics can be utilized to detect out-of-specification (OOS) trends in pharmaceutical production processes in order to strengthen public health protection through enhanced quality assurance mechanisms. This objective focuses on evaluating the effectiveness of statistical and computational techniques in identifying deviations in critical quality attributes across different stages of manufacturing, including raw material handling, processing, and final product testing. The study aims to quantify the relationship between data-driven monitoring systems and the early detection of process variability, particularly emphasizing the role of out-of-trend (OOT) signals as precursors to OOS events. By analyzing large-scale production datasets, the research seeks to measure the accuracy, sensitivity, and reliability of various analytical models such as statistical process control, multivariate analysis, and predictive algorithms in detecting deviations from established quality thresholds. Another key objective is to assess the extent to which real-time data analytics contributes to reducing batch failures, minimizing product recalls, and improving overall process stability within pharmaceutical manufacturing environments. The study also intends to investigate how the integration of automated analytics systems influences decision-making processes, specifically in terms of root cause identification and implementation of corrective and preventive actions. Furthermore, the research aims to evaluate the consistency of analytical outcomes across different production settings, ensuring that data-driven approaches can be generalized and applied effectively in diverse operational contexts. By focusing on measurable indicators such as deviation frequency, detection time, and process capability indices, the study seeks to provide empirical evidence on the impact of data analytics in maintaining compliance with quality standards. Ultimately, this objective is designed to establish a quantitative understanding of how advanced data analytics techniques enhance the detection of OOS trends, thereby supporting the continuous production of safe, effective, and high-quality pharmaceutical products that are essential for protecting public health.

LITERATURE REVIEW

The literature review in a quantitative study serves as a structured synthesis of existing empirical findings, theoretical constructs, and methodological approaches that inform the research problem and

guide the analytical framework (Akash & Rehman, 2020). In the context of pharmaceutical manufacturing, the application of data analytics to detect out-of-specification (OOS) trends represents a convergence of quality engineering, statistical modeling, and regulatory science. This section systematically examines prior quantitative studies that have explored process monitoring, statistical process control, multivariate analysis, predictive modeling, and real-time data systems within pharmaceutical and related high-reliability industries. The purpose of this review is to establish a comprehensive understanding of how numerical data, analytical models, and statistical techniques have been used to detect deviations in production processes and to evaluate their effectiveness in maintaining product quality. By focusing on measurable outcomes such as detection accuracy, false positive rates, process capability indices, and deviation frequencies, the review emphasizes the quantitative dimensions of data analytics in pharmaceutical quality assurance (Achanta et al., 2021). This literature review is structured to critically analyze the evolution of analytical techniques from traditional univariate statistical methods to advanced multivariate and machine learning-based approaches. It examines how these methods contribute to the identification of out-of-trend (OOT) patterns that precede OOS events, thereby enabling early intervention and process optimization. The review also considers the role of large-scale data integration, real-time monitoring systems, and automated analytics in enhancing the precision and timeliness of deviation detection. In addition, the section evaluates empirical evidence on the impact of data analytics on key performance indicators such as batch rejection rates, process variability, and compliance metrics. Through this synthesis, the literature review aims to identify gaps in existing research, particularly in terms of model validation, scalability, and cross-process applicability (Mukherjee, 2019). The insights derived from this section provide the foundation for developing a robust quantitative framework that assesses the effectiveness of data analytics in detecting OOS trends in pharmaceutical production for public health protection.

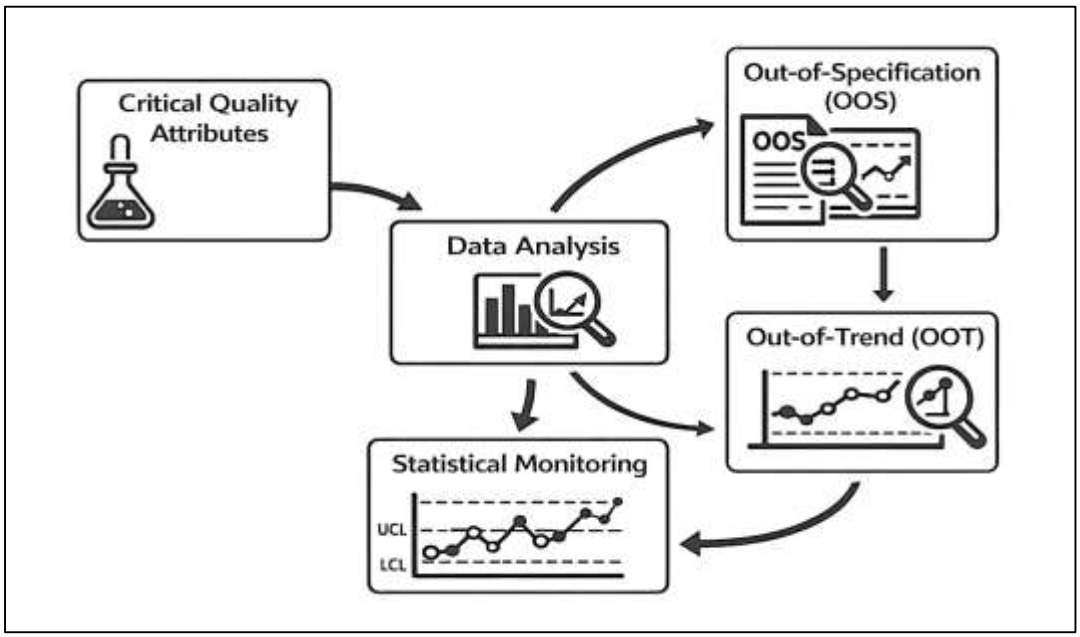
Conceptual Foundations of Pharmaceutical Quality and Deviation Analysis

The conceptual foundations of pharmaceutical quality and deviation analysis are grounded in the systematic quantification of process performance and product attributes to ensure consistent compliance with predefined standards of safety, efficacy, and reliability. Pharmaceutical quality is commonly understood as the degree to which a product conforms to established specifications throughout its lifecycle, encompassing raw material integrity, manufacturing consistency, and final product performance (Rawal et al., 2019). Within this framework, deviation analysis serves as a structured mechanism for identifying, categorizing, and interpreting variations that occur during production. Extensive literature across pharmaceutical engineering and quality management consistently highlights that variability is an inherent characteristic of manufacturing systems, driven by factors such as material inconsistencies, environmental fluctuations, equipment performance, and human interaction. As a result, the integration of statistical thinking into pharmaceutical systems has become essential for distinguishing between normal process variability and abnormal deviations that may compromise product quality. The theoretical distinction between common cause and special cause variability forms the basis of deviation analysis, enabling manufacturers to determine whether observed variations are part of routine process behavior or indicative of underlying issues requiring intervention (Pedro et al., 2023). Empirical studies have demonstrated that structured deviation analysis frameworks significantly improve the detection of process anomalies, thereby reducing the likelihood of quality failures. Furthermore, research on pharmaceutical quality systems emphasizes the importance of continuous monitoring across the product lifecycle, where data generated during development, validation, and commercial production are systematically analyzed to ensure process stability. The increasing adoption of data-driven approaches has reinforced the role of analytics in deviation detection, enabling more precise identification of patterns and trends in production data (Silge et al., 2022). These conceptual foundations collectively establish a robust framework for understanding and managing variability in pharmaceutical systems, ensuring that deviations are detected and addressed in a timely and effective manner.

Critical Quality Attributes represent the measurable physical, chemical, biological, or microbiological properties that must be maintained within defined limits to ensure that a pharmaceutical product meets its intended quality standards (Javed et al., 2019). These attributes are directly linked to the safety and therapeutic effectiveness of the product, making their identification and control a central focus of

pharmaceutical manufacturing. The determination of CQAs is typically based on a risk-oriented evaluation of product characteristics, where attributes with the highest potential impact on patient outcomes are prioritized for monitoring and control. Literature on pharmaceutical development and quality by design consistently indicates that CQAs should be established early in the product lifecycle and continuously evaluated throughout manufacturing to ensure consistency. Measurement of CQAs relies on validated analytical methods that produce quantitative data, enabling precise assessment of product quality. These methods include advanced instrumental techniques capable of delivering high levels of accuracy and reproducibility, which are essential for reliable statistical evaluation (Gül et al., 2023). The quantitative nature of CQA measurement allows for the establishment of specification ranges that define acceptable limits for each attribute. Statistical tools are widely used to analyze CQA data, providing insights into central tendencies, variability, and process stability. Research findings highlight that consistent monitoring of CQAs enables the identification of trends and deviations that may signal emerging process issues. In addition, the integration of advanced data analytics has enhanced the ability to detect complex patterns within large datasets, supporting more informed decision-making in quality management. Studies also emphasize the importance of data integrity and method validation in ensuring the reliability of CQA measurements, as inaccuracies in data collection or analysis can lead to incorrect assessments of product quality (Separovic et al., 2023). By maintaining rigorous control over CQAs, pharmaceutical manufacturers can ensure that products consistently meet regulatory requirements and deliver the desired therapeutic outcomes.

Figure 3: Pharmaceutical Deviation Analysis Quality Framework



Quantitative thresholds for out-of-specification and out-of-trend results form a critical component of pharmaceutical quality management systems, providing clear and measurable criteria for identifying deviations in production and analytical data. Out-of-specification results are defined as values that fall outside established acceptance limits, indicating that a product does not meet predefined quality standards (Mahfuj Ahmed & Md. Hasan Or, 2021; Mohammad Robel & Md. Morshedul, 2021; Oliveira et al., 2019). These limits are determined based on scientific evidence, historical data, and regulatory requirements, ensuring that they reflect acceptable levels of variation for each critical quality attribute. In contrast, out-of-trend results represent deviations from expected historical patterns while still remaining within specification limits, serving as early indicators of potential process instability (Aditya & Mohammad Robel, 2022; Istiaq & Nusrat, 2022). Literature on pharmaceutical quality systems consistently highlights the importance of distinguishing between these two types of deviations, as they require different levels of investigation and response. Out-of-specification results typically trigger

formal investigations and corrective actions, while out-of-trend results prompt trend analysis and risk evaluation to prevent escalation (Arden et al., 2021). Empirical studies demonstrate that the implementation of quantitative thresholds for both OOS and OOT enhances the sensitivity of monitoring systems, enabling earlier detection of process deviations. Statistical methods such as control charting and trend analysis are widely used to establish and evaluate these thresholds, providing a structured approach to identifying abnormal patterns in data. Research also indicates that the effectiveness of these thresholds depends on the quality and volume of historical data, as well as the robustness of analytical methods used for measurement. In addition, the integration of automated data analytics systems has improved the consistency and accuracy of deviation detection by reducing reliance on manual interpretation (Jelsch et al., 2021; Md Khaled & Hisham, 2022; Md Mehedi & Md, 2022). By incorporating well-defined quantitative thresholds into quality management frameworks, pharmaceutical manufacturers can improve their ability to detect deviations, maintain process control, and ensure consistent product quality (Portela et al., 2020).

The statistical representation of process variability is a fundamental aspect of pharmaceutical quality management, providing a quantitative basis for understanding and controlling deviations in manufacturing processes. Variability arises from multiple sources within the production environment and must be carefully monitored to ensure that it remains within acceptable limits (Curcuruto et al., 2019; Md. Mainuddin & Palash Chandra, 2022). Statistical techniques are used to characterize this variability by analyzing data distributions, identifying patterns, and assessing process stability over time. Literature on quality engineering emphasizes that continuous monitoring of variability enables the detection of deviations from expected behavior, supporting timely intervention and process optimization. Process capability analysis plays a central role in this context by evaluating how well a process can consistently produce outputs within specified limits. Capability indices are widely used as quantitative measures of process performance, providing insights into the relationship between process variability and specification requirements (Lee et al., 2022; Md. Morshedul et al., 2022; Md. Nazmul & Amena Begum, 2022). Research findings indicate that higher capability levels are associated with greater process stability and reduced risk of quality failures, while lower capability levels signal increased variability and potential non-compliance. In addition to capability analysis, quantitative risk metrics are employed to assess the likelihood and potential impact of deviations, forming the basis of risk-based quality management approaches. These metrics incorporate statistical probabilities and severity assessments to prioritize areas of concern and guide decision-making processes. Studies on pharmaceutical risk management highlight that integrating risk metrics with statistical monitoring enhances the ability to detect and mitigate deviations before they result in out-of-specification events. Furthermore, the application of multivariate analysis techniques has enabled a more comprehensive understanding of process variability by capturing interactions among multiple variables simultaneously (Md. Shahinur & Md. Sultan, 2022; Tanjina Binte & Md. Hasan Or, 2022; Volta e Sousa et al., 2021). This approach provides deeper insights into the underlying causes of variability and supports more effective control strategies. By combining statistical representation, process capability evaluation, and quantitative risk assessment, pharmaceutical systems establish a robust framework for maintaining process control, ensuring consistent product quality, and minimizing the occurrence of deviations.

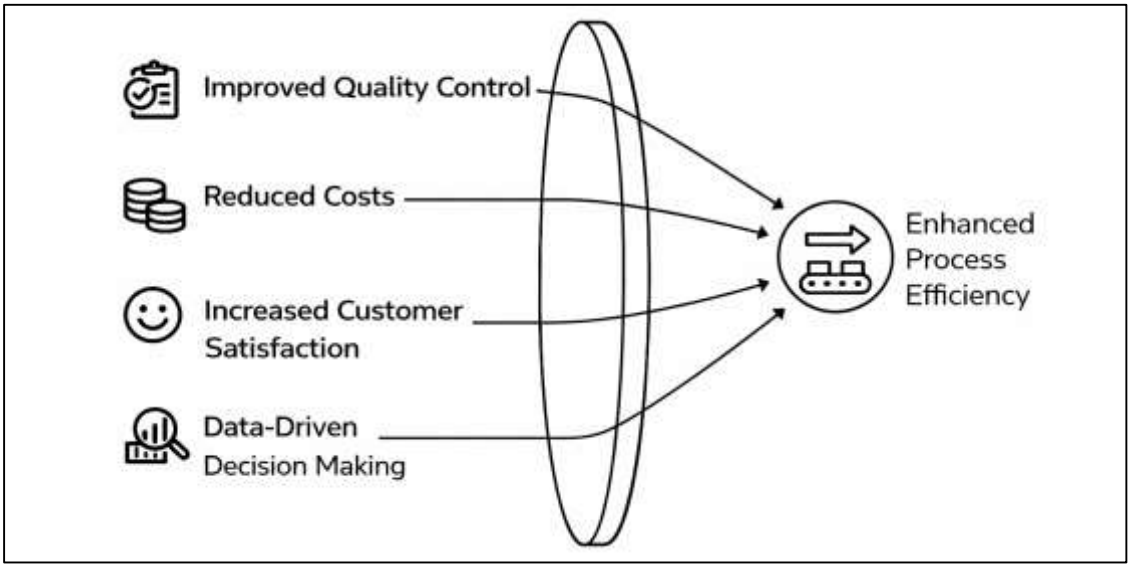
Statistical Process Control (SPC) in Pharmaceutical Manufacturing

Statistical Process Control (SPC) has emerged as a central quantitative framework in pharmaceutical manufacturing for monitoring, controlling, and improving production processes through continuous data analysis. The literature consistently characterizes SPC as a systematic approach that transforms process data into meaningful indicators of performance, enabling the identification of variability and deviations that may affect product quality (Amena Begum & Mst Kaniz, 2023; Ferdous Ara & Beatrice Onyinyechi, 2023; Kenett et al., 2023). In pharmaceutical environments, where strict regulatory compliance and product consistency are essential, SPC provides a structured mechanism for ensuring that manufacturing processes remain within acceptable limits. Studies across pharmaceutical quality engineering demonstrate that SPC contributes significantly to reducing process variability, enhancing consistency, and improving overall manufacturing efficiency. The application of SPC is particularly relevant in complex production systems where multiple variables interact and influence critical quality

attributes. Empirical research highlights that the integration of SPC into pharmaceutical quality systems supports real-time monitoring and facilitates early detection of abnormal patterns, thereby preventing the escalation of minor deviations into critical quality failures (Vanli & Castillo, 2021). The literature also emphasizes that SPC aligns with modern quality management principles, including continuous process verification and data-driven decision-making. By enabling the differentiation between inherent variability and assignable causes of deviation, SPC provides a scientific basis for process control and improvement. Furthermore, studies indicate that organizations implementing SPC experience measurable improvements in process stability and a reduction in out-of-specification occurrences. The growing reliance on digital data systems has further enhanced the applicability of SPC, allowing for automated data collection and analysis that improves monitoring accuracy and responsiveness (Islam & Aditya, 2023; Mahfuj Ahmed & Md. Mehedi, 2023; Ramos et al., 2021). Overall, SPC is widely recognized in the literature as an essential quantitative tool that underpins effective pharmaceutical manufacturing and ensures the consistent production of high-quality products.

The application of control charts represents one of the most extensively examined components of SPC in pharmaceutical manufacturing, providing a visual and statistical method for tracking process behavior over time. Literature consistently identifies various types of control charts, each designed to monitor specific aspects of process data depending on the nature and structure of the observations (Md. Hasan Or et al., 2023; Md. Mainuddin & Palash Chandra, 2023; Pérez-Benítez et al., 2023). Charts based on subgroup averages and ranges are commonly used in batch-oriented processes, where data are collected in groups, while individual observation charts are applied in continuous production settings. More advanced charting techniques, including those that incorporate weighting and cumulative information, have been widely studied for their enhanced ability to detect subtle and gradual changes in process behavior. Empirical findings suggest that these advanced charts offer improved sensitivity compared to traditional approaches, making them particularly valuable in pharmaceutical contexts where small deviations can have significant implications for product quality (Jin et al., 2022; Md. Mehedi & Khairum Nahar, 2023; Mostafa, 2023). The literature also highlights the importance of selecting appropriate charting techniques based on process characteristics, as the effectiveness of control charts depends on factors such as data distribution, sampling frequency, and process variability. Studies further indicate that the proper design and implementation of control charts contribute to more accurate detection of abnormal patterns, enabling timely corrective actions. In addition, the integration of control charts with automated monitoring systems has been shown to improve efficiency by reducing manual intervention and enabling real-time analysis (Silva et al., 2019). Research consistently demonstrates that control charts serve as a critical tool for visualizing process performance, supporting both operational decision-making and regulatory compliance in pharmaceutical manufacturing environments.

Figure 4: Statistical Process Control Benefits Framework



Detection power and sensitivity analysis are key aspects of SPC that determine the effectiveness of monitoring systems in identifying process deviations. The literature defines detection power as the ability of a statistical method to correctly identify true deviations, while sensitivity refers to its responsiveness to small changes in process behavior (Palash Chandra, 2023; Peterson et al., 2019; Rukaiya Khatun & Zakia, 2023). Studies across pharmaceutical and industrial applications consistently highlight the importance of balancing these two factors to achieve reliable monitoring performance. Highly sensitive systems are capable of detecting minor shifts in process parameters, which is particularly important in pharmaceutical production where even small deviations can impact product quality. At the same time, excessive sensitivity may lead to instability in monitoring systems, increasing the likelihood of unnecessary signals. Empirical research demonstrates that advanced SPC techniques, particularly those incorporating cumulative or weighted approaches, exhibit greater sensitivity and improved detection capability compared to traditional methods (Amena Begum & Mst Kaniz, 2024; Md Khaled & Md. Morshedul, 2024; Qiu, 2020). Sensitivity analysis is widely used in the literature to evaluate how different SPC models perform under varying process conditions, including changes in variability, sample size, and data frequency. Findings indicate that the effectiveness of SPC techniques is influenced by these factors, highlighting the need for careful system design and parameter selection. In addition, studies have shown that combining SPC with data analytics enhances detection power by enabling the identification of complex patterns and relationships within large datasets. This integration improves the accuracy and timeliness of deviation detection, supporting proactive quality management (Md. Mehedi & Khairum Nahar, 2024; Md. Towhidul & Uddin, 2024; Qiu, 2020). Overall, the literature underscores that high detection power and appropriate sensitivity are essential for ensuring that SPC systems effectively identify deviations and maintain process control in pharmaceutical manufacturing. False alarm rates and error quantification represent critical considerations in evaluating the performance of SPC systems, as they directly influence the reliability and efficiency of process monitoring (Mohammad Robel & Md. Morshedul, 2024; Yuan et al., 2020; Zakia & Rukaiya Khatun, 2024). The literature consistently categorizes errors into two primary types: incorrect signals indicating deviations when none exist and missed detections where actual deviations are not identified. Studies emphasize that managing these errors is essential for maintaining confidence in SPC systems and ensuring effective quality control. High rates of incorrect signals can lead to unnecessary investigations, increased operational costs, and reduced trust in monitoring tools, while missed detections pose a significant risk to product quality and regulatory compliance. Empirical research demonstrates that the design of monitoring parameters, including control limits and sampling strategies, plays a crucial role in determining error rates (Bottani et al., 2023). Comparative studies of different SPC models indicate that advanced techniques can reduce error rates by providing more accurate detection of process shifts.

The literature also explores the performance differences between short-run and long-run SPC models, highlighting that short-run models are more suitable for processes with limited data availability, whereas long-run models offer greater stability and reliability in continuous production environments. Quantitative evaluations consistently show that well-implemented SPC systems significantly improve the detection of out-of-specification events, leading to reductions in batch failures and improvements in overall process performance (He et al., 2019). By systematically addressing error rates and optimizing model performance, SPC provides a robust and reliable framework for detecting deviations and maintaining high standards of quality in pharmaceutical manufacturing.

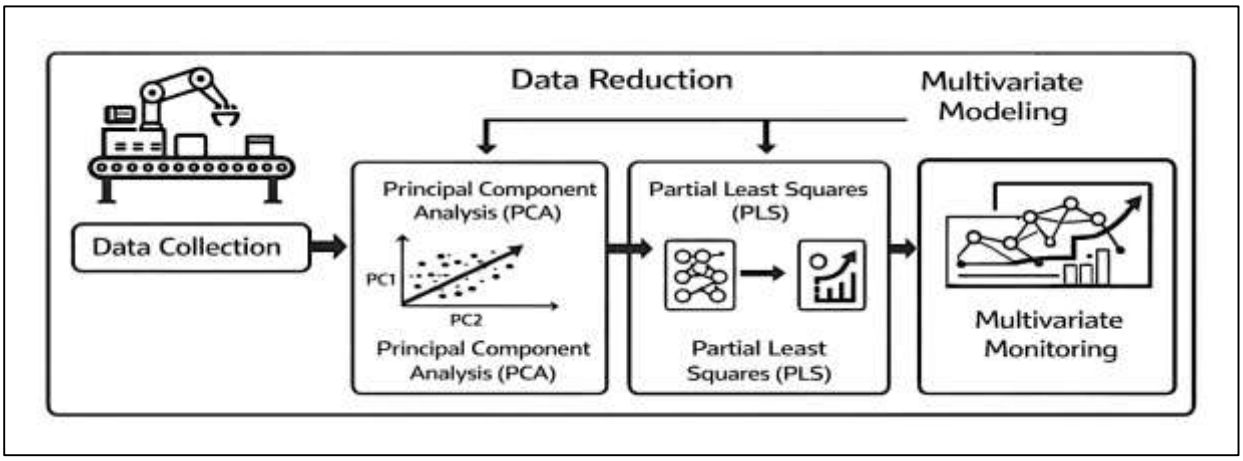
Multivariate Data Analysis for Complex Process Monitoring

Multivariate data analysis has become a critical methodological advancement in pharmaceutical manufacturing, particularly in the context of complex process monitoring where multiple interdependent variables influence product quality (Li et al., 2022). Traditional univariate approaches, which evaluate each variable independently, have been widely recognized in the literature as insufficient for capturing the intricate relationships that characterize modern pharmaceutical processes. As manufacturing systems have evolved to incorporate advanced technologies and highly controlled environments, the volume and complexity of generated data have increased significantly. Studies across pharmaceutical engineering and industrial analytics consistently emphasize that multivariate approaches provide a more comprehensive understanding of process behavior by simultaneously analyzing multiple variables and their interactions (Yu et al., 2021). This capability is particularly important in identifying subtle deviations that may not be apparent when variables are assessed in isolation. Empirical research demonstrates that multivariate data analysis enhances the detection of process anomalies, improves process stability, and supports more accurate quality assessments. The integration of multivariate techniques into pharmaceutical quality systems aligns with data-driven decision-making frameworks, enabling manufacturers to interpret high-dimensional datasets effectively. Literature further indicates that multivariate monitoring supports regulatory expectations by providing robust statistical evidence of process control and consistency. In addition, the application of these techniques has been associated with reductions in process variability and improvements in product quality outcomes. The growing adoption of digital manufacturing systems has further reinforced the importance of multivariate analysis, as large datasets generated from sensors and monitoring devices require sophisticated analytical methods for meaningful interpretation (Lee et al., 2019). Overall, multivariate data analysis is widely recognized as an essential tool for managing complexity in pharmaceutical production, providing a structured framework for detecting, analyzing, and controlling process variability.

Principal Component Analysis (PCA) has been extensively studied as a foundational technique for dimensionality reduction in pharmaceutical data analysis, addressing the challenges associated with high-dimensional datasets (Bo et al., 2022). PCA operates by transforming a large set of correlated variables into a smaller number of uncorrelated components that capture the majority of the variability present in the data. Literature consistently highlights that this transformation simplifies data interpretation while preserving critical information, making it particularly valuable in pharmaceutical processes characterized by numerous interrelated parameters. Empirical studies demonstrate that PCA effectively reduces noise and redundancy in datasets, enabling clearer identification of underlying patterns and trends (Mertler et al., 2021). This capability is essential for monitoring complex processes where multiple variables contribute to product quality. Research findings indicate that PCA enhances anomaly detection by highlighting deviations from established data structures, thereby supporting early identification of process disturbances. In pharmaceutical manufacturing, PCA has been applied to various stages of production, including blending, granulation, and formulation, where it facilitates the monitoring of process consistency and the identification of variability sources. The literature also emphasizes that PCA improves computational efficiency by reducing the dimensionality of datasets, allowing for faster analysis and real-time monitoring (Albert, 2025; Ishtiaque & Rajib, 2025; Shang & You, 2019). Furthermore, studies have shown that PCA serves as a foundational step for other multivariate techniques, providing a simplified data structure that enhances the performance of subsequent analytical models. By enabling the effective management of high-dimensional data, PCA contributes significantly to improving process understanding and supporting data-driven quality

assurance in pharmaceutical systems.

Figure 5: Multivariate Analysis in Pharmaceutical Manufacturing



Partial Least Squares (PLS) modeling represents another key multivariate technique widely applied in pharmaceutical manufacturing for analyzing the relationship between process variables and product quality attributes. The literature consistently describes PLS as a method that establishes predictive relationships between input variables and output responses, making it particularly useful for process optimization and quality control (Garitano et al., 2019). Empirical studies demonstrate that PLS is effective in handling complex datasets where variables are highly correlated, a common characteristic in pharmaceutical processes. By modeling the relationship between process parameters and critical quality attributes, PLS enables manufacturers to identify the factors that have the greatest impact on product quality. This capability supports targeted process improvements and enhances the ability to maintain consistent production outcomes. Research findings indicate that PLS models provide accurate predictions of quality attributes, enabling proactive monitoring and control of manufacturing processes (Jin et al., 2019). In addition, PLS has been widely used in conjunction with real-time data systems to support continuous process verification, allowing for immediate detection of deviations and timely corrective actions. The literature also highlights the robustness of PLS in dealing with noisy and incomplete data, which are common challenges in pharmaceutical manufacturing environments. Furthermore, studies have shown that PLS contributes to improved process understanding by revealing complex relationships between variables that may not be evident through traditional analysis methods (Qin & Chiang, 2019). By facilitating the integration of process data and quality outcomes, PLS plays a crucial role in enhancing the effectiveness of multivariate monitoring systems in pharmaceutical production.

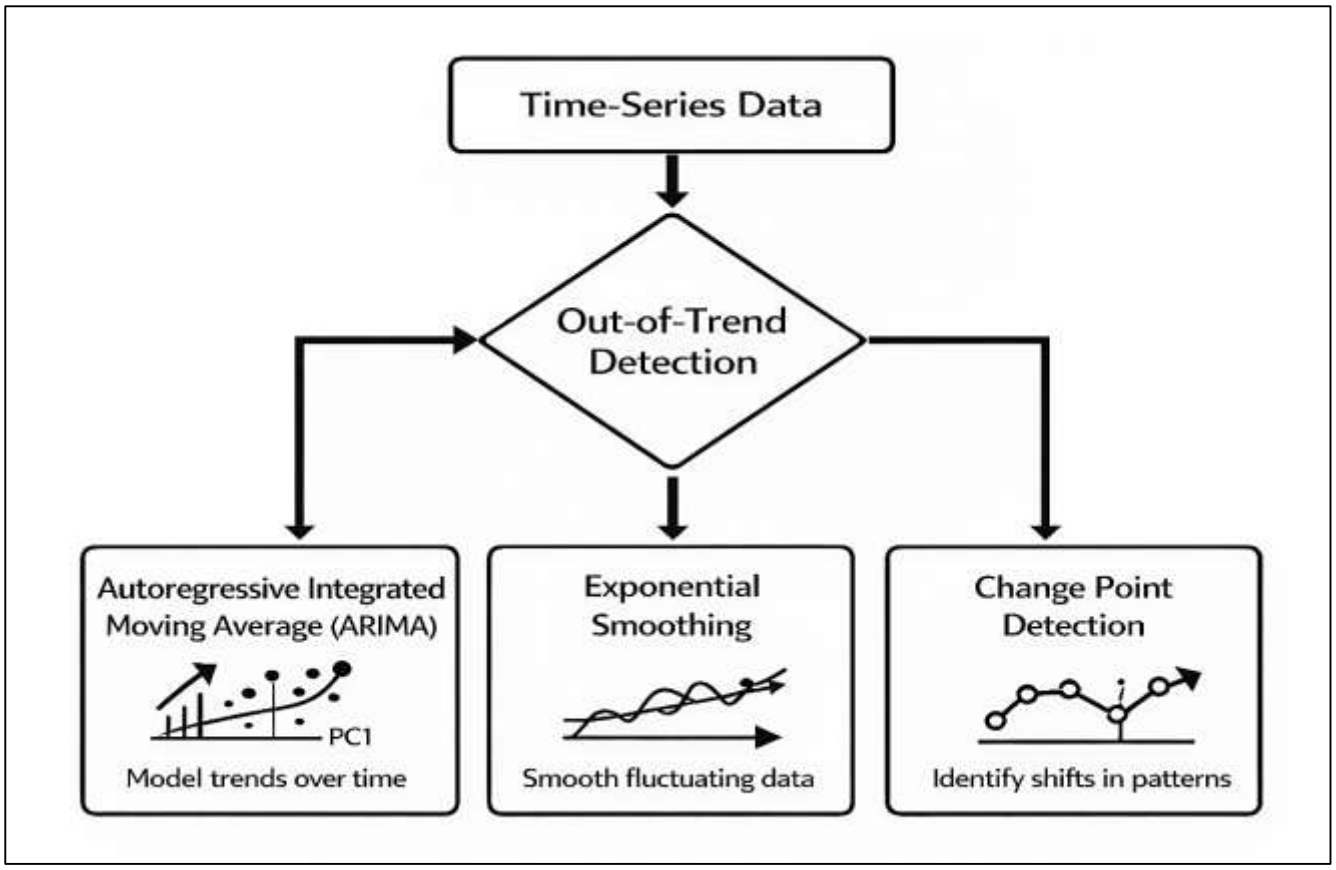
Time-Series and Trend Analysis in OOT Detection

Time-series and trend analysis have become central in the literature on out-of-trend detection because pharmaceutical manufacturing data are generated sequentially across batches, campaigns, stability intervals, and routine quality control measurements. A large body of scholarship presents pharmaceutical data not as isolated observations but as ordered process evidence that reflects operational continuity, gradual drift, recurring variation, and emerging instability (van de Wiel et al., 2020). This perspective is particularly important for out-of-trend evaluation because many early warnings of quality deterioration do not initially appear as clear failures against fixed specification limits. Instead, they emerge as subtle directional movement, repeated clustering near one side of an expected range, increased fluctuation over time, or structural changes in the pattern of results. Literature on stability monitoring, process verification, and ongoing quality review consistently shows that time-ordered analysis provides a stronger basis for identifying meaningful deviation than single-point review alone. Researchers describe time-series methods as useful because they preserve chronology, allowing analysts to distinguish transient noise from persistent change. In pharmaceutical systems, this distinction is crucial since assay values, impurity levels, dissolution outcomes, environmental conditions, and in-process control data often display gradual evolution rather than

abrupt breakdown (Alimohammadi & Chen, 2022; Kazi Rakib Hasan, 2025; Md. Ashfaq & Ashraful, 2025). The literature further shows that trend analysis supports a more nuanced interpretation of quality signals by connecting present observations to historical behavior. In this context, out-of-trend detection is discussed as an analytical process that examines whether the recent path of data remains consistent with established product and process history. Studies synthesized across manufacturing analytics, process control, and pharmaceutical quality management indicate that this approach improves sensitivity to drift, supports earlier investigation, and strengthens the documentation of process knowledge. Time-series analysis is also widely associated with lifecycle quality oversight because it enables continuous assessment rather than episodic review. Scholars repeatedly emphasize that the real value of time-based monitoring lies in its ability to reveal movement before formal failure occurs, thereby supporting a preventive rather than purely reactive quality strategy (Mohammad Robel, 2025; Murad, 2025; Shakil et al., 2023). As a result, the literature presents time-series and trend analysis as a core quantitative foundation for identifying out-of-trend behavior in pharmaceutical production and stability systems, especially where gradual deviation, process aging, and recurring operational patterns shape product quality outcomes over time.

Within this broader time-series framework, autoregressive integrated moving average models occupy a prominent place in the literature because they provide a structured way to model serial dependence, underlying direction, and short-term fluctuation in pharmaceutical datasets (Ghaderpour et al., 2021). Scholars frequently describe these models as especially valuable when current observations are influenced by prior results, a condition that commonly applies to batch sequences, environmental monitoring records, and long-term stability data. The literature shows that autoregressive integrated moving average modeling is often used to separate routine variation from meaningful movement by accounting for the internal structure of the data rather than treating each point as independent. This is important in out-of-trend detection because false interpretation often arises when natural serial behavior is mistaken for abnormal change or when true drift is hidden within apparently modest variation. Studies across pharmaceutical analytics have reported that these models are useful in identifying persistent directional movement in assay results, impurity growth, and other quality indicators that evolve gradually over storage or production cycles. The literature also notes that these models support short-horizon forecasting, which allows analysts to compare observed values with statistically expected behavior and thereby identify unusual departures from the established pattern (Namoano et al., 2019). In quality review settings, this approach helps determine whether recent observations represent continuation of normal behavior or the onset of an unstable trend. Scholars further explain that autoregressive integrated moving average methods are particularly effective in settings where data are sufficiently frequent and historically organized, making them suitable for ongoing process verification and stability trending. At the same time, the literature recognizes that their usefulness depends on data quality, appropriate model identification, and careful interpretation within the pharmaceutical context. Researchers therefore emphasize that model-based trend detection should be anchored in process understanding and not used as an isolated statistical exercise (Mo et al., 2022). Even with this caution, a broad synthesis of studies shows strong support for these models as a rigorous means of describing temporal structure, clarifying process behavior, and improving the sensitivity of out-of-trend assessment. Their contribution in the literature lies not only in forecasting but also in transforming sequential pharmaceutical measurements into interpretable evidence of drift, persistence, and abnormal movement, thereby strengthening the statistical basis of deviation detection in regulated manufacturing environments.

Figure 6: Time Series Trend Analysis Framework



Exponential smoothing and seasonal decomposition techniques are also widely discussed in the literature as practical and effective tools for identifying out-of-trend behavior in pharmaceutical data, especially when the objective is to clarify the underlying signal in the presence of noise, cyclical movement, or recurring operational effects (Apostol et al., 2021). Scholars often describe exponential smoothing as a valuable approach because it gives greater weight to more recent observations while still retaining information from earlier data points. This characteristic is particularly relevant for pharmaceutical monitoring because recent batches or recent stability intervals often provide the most immediate indication of process change. Literature across manufacturing analytics and quality surveillance indicates that smoothing techniques improve interpretability by reducing short-term irregularity and making broader directional patterns easier to observe. In pharmaceutical applications, this can be helpful when assay, dissolution, or impurity data exhibit small point-to-point fluctuations that obscure the real trajectory of process behavior. Seasonal decomposition receives attention in the literature when recurring patterns may influence process outcomes, such as shifts associated with environmental cycles, equipment usage rhythms, or repeated operational scheduling. Researchers note that separating the data into baseline level, recurring pattern, and irregular remainder enhances the ability to isolate genuine out-of-trend movement from predictable repetition (He et al., 2023). This distinction is important because a recurring rise or fall should not automatically be interpreted as deterioration if it reflects a stable operational cycle. Studies synthesized from process monitoring and pharmaceutical stability management show that decomposition-based analysis can reveal whether apparent deviation is attributable to trend, seasonality, or random disturbance. The literature also emphasizes that these approaches are useful not only for detection but also for communication, since smoothed and decomposed trends are often easier for quality teams to interpret than raw sequential values. In addition, scholars report that such methods can support internal review, deviation investigation, and ongoing process verification by making long-run movement visible without overreacting to minor short-term swings (Yan et al., 2019). Across the reviewed literature, exponential smoothing and seasonal decomposition are therefore positioned as accessible yet analytically robust tools that help quality professionals interpret temporal structure, identify meaningful deviations, and

strengthen the consistency of out-of-trend judgments in pharmaceutical systems where sequential data are affected by both noise and recurring patterns.

A further major theme in the literature concerns trend slope estimation, change point detection, quantitative indicators of deviation, and the evaluation of forecast accuracy in pharmaceutical time-series analysis. Scholars consistently argue that effective out-of-trend assessment requires more than visual inspection of graphs; it also requires numerical characterization of how quickly data are moving, when the pattern changed, and how strongly current behavior departs from historical expectation (Al-Ghuwairi et al., 2023). Trend slope estimation is widely discussed as a means of expressing direction and magnitude of movement over time, allowing analysts to determine whether a variable is gradually increasing, decreasing, or remaining stable. In pharmaceutical contexts, this has been applied to stability data, impurity accumulation, microbial counts, and process performance indicators. Change point detection is presented in the literature as especially valuable because it identifies moments when the statistical behavior of a sequence shifts, such as a sudden increase in variability, an abrupt level change, or the onset of sustained drift. Researchers regard this capability as highly relevant for root-cause investigation because it can help connect deviation onset to operational events such as equipment maintenance, raw material changes, or method adjustments (Nizam et al., 2022). The literature also addresses quantitative metrics for deviation, including indicators related to variance shift and drift rate, which help convert qualitative impressions of instability into measurable evidence. Studies repeatedly show that widening variation may be as important as directional shift, since growing inconsistency can indicate loss of control even when the average result still appears acceptable. Forecast accuracy measures are likewise emphasized because they provide a basis for judging how well time-series models represent pharmaceutical data and how trustworthy model-based alerts may be. Scholars discuss error-based evaluation as essential for comparing forecasting approaches and ensuring that trend detection systems are not only sensitive but also reliable. Across the literature, the combined use of slope analysis, change point identification, deviation metrics, and forecast performance assessment is presented as a mature and rigorous strategy for out-of-trend detection (Choi et al., 2021). This synthesis shows that quantitative temporal analysis strengthens pharmaceutical quality oversight by making subtle change measurable, interpretable, and actionable within structured monitoring systems.

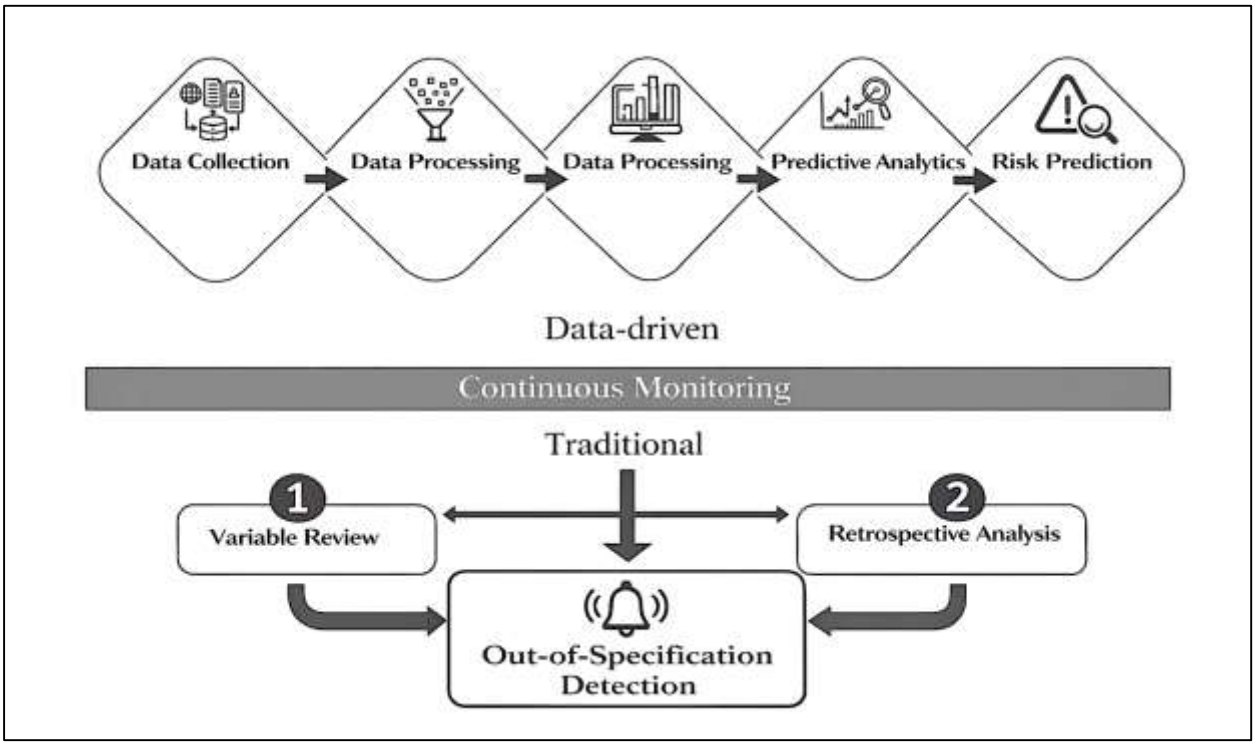
Machine Learning and Predictive Analytics for OOS Detection

Machine learning and predictive analytics have become increasingly prominent in the literature on out-of-specification detection because pharmaceutical manufacturing environments generate large volumes of structured and semi-structured data that exceed the interpretive capacity of conventional manual review (Masood & Hashmi, 2019). Across the literature, machine learning is framed as an extension of data-driven quality management in which historical process data, laboratory results, equipment signals, and environmental measurements are used to identify complex patterns associated with quality deviation. Unlike traditional monitoring strategies that rely heavily on fixed thresholds, isolated variable review, or retrospective investigation, machine learning-based systems are designed to recognize relationships among multiple variables and detect combinations of conditions that precede specification failure. This feature is especially important in pharmaceutical production because OOS events rarely arise from a single isolated factor. They often emerge through interacting conditions such as material variability, equipment drift, environmental fluctuation, sampling inconsistency, and cumulative process instability. The literature consistently reports that predictive analytics strengthens quality oversight by converting historical data into actionable probability-based insight, enabling earlier recognition of risk conditions before formal product failure occurs (Carvajal Soto et al., 2019). Studies from pharmaceutical quality control, process systems engineering, chemometrics, and intelligent manufacturing show that machine learning improves sensitivity to subtle process deterioration and allows for a more dynamic interpretation of quality risk. Researchers also note that the appeal of predictive analytics in regulated manufacturing lies in its ability to support preventive action, root-cause exploration, and targeted intervention while operating within increasingly digital production systems. A major theme in the literature is that machine learning does not replace pharmaceutical quality principles but deepens them by offering a richer interpretation of process behavior. In this sense, predictive analytics is frequently positioned as part of a broader quality modernization effort that aligns continuous monitoring, advanced process understanding, and risk-

based decision-making (Andaur et al., 2021; Md Khaled, 2026). The literature further indicates that the value of machine learning for OOS detection depends on model transparency, data integrity, and meaningful integration into existing quality systems. Even so, the overall body of research shows strong agreement that predictive approaches provide substantial analytical advantages in environments where traditional monitoring alone may overlook early indicators of specification failure. As a result, machine learning has been widely discussed as a significant methodological development in the effort to detect, understand, and reduce OOS events in pharmaceutical production systems.

A substantial portion of the literature focuses on supervised learning models, particularly regression-based methods, decision trees, and random forest approaches, because these techniques allow quality outcomes to be predicted from labeled historical data. In supervised learning, models are trained using prior examples in which the relationship between process conditions and quality results is already known, enabling the system to learn which combinations of variables are associated with acceptable performance and which are associated with deviation (Giaconia & Chamas, 2023). Regression-oriented models have been widely described as useful for estimating the degree of risk or the expected value of a quality outcome, especially when the objective is to quantify the influence of process variables on assay performance, impurity behavior, dissolution consistency, or other critical quality attributes. Decision trees are frequently highlighted in the literature for their interpretability, as they organize predictive logic into sequential branching rules that are often easier for quality and manufacturing teams to understand than purely abstract statistical outputs. This interpretability is especially valuable in regulated pharmaceutical settings where model acceptance often depends on the ability to explain why a system flagged a risk condition (Chen et al., 2019).

Figure 7: Machine Learning OOS Detection Framework



Random forest models receive extensive attention because they improve stability and predictive robustness by aggregating multiple decision structures rather than relying on a single tree. Studies across pharmaceutical manufacturing and quality analytics consistently report that these ensemble approaches tend to perform well when process data are nonlinear, noisy, or influenced by variable interactions that simpler models may fail to capture. Researchers also emphasize that supervised models are especially useful when sufficient historical records exist for known OOS and non-OOS cases, allowing the model to learn from actual production behavior. In the literature, these methods are often described as stronger than basic threshold-based monitoring when the causes of specification

failure are multifactorial rather than singular. At the same time, many studies caution that model performance depends heavily on label quality, class balance, and the representativeness of the training data (Campbell et al., 2022). Pharmaceutical datasets often contain relatively fewer true OOS cases than normal outcomes, which creates modeling challenges that must be addressed through careful design and validation. Even with these challenges, the literature broadly supports supervised learning models as effective tools for identifying risk patterns, ranking influential variables, and improving early prediction of OOS events in complex production environments.

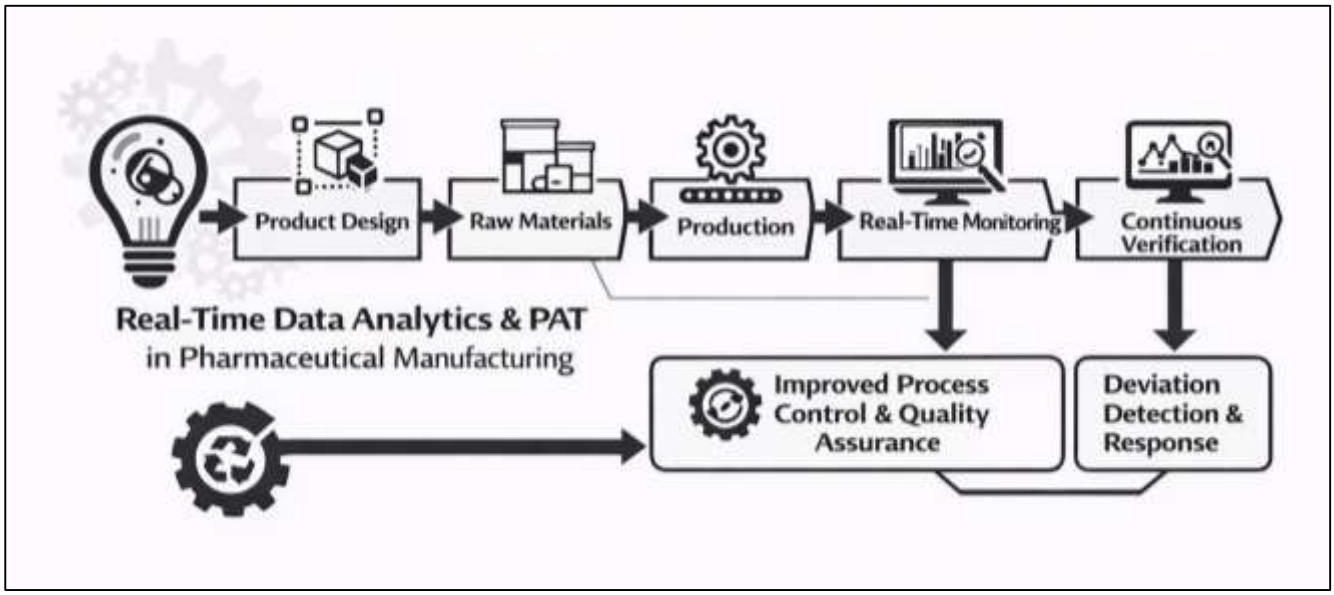
Real-Time Data Analytics and Process Analytical Technology (PAT)

Real-time data analytics and Process Analytical Technology have become central themes in the pharmaceutical manufacturing literature because they support the transition from end-product testing toward continuous process understanding and active quality assurance. The literature consistently presents PAT as a structured framework for designing, analyzing, and controlling manufacturing processes through timely measurement of critical material attributes and process parameters during production (Kim et al., 2021). Within this framework, real-time data analytics functions as the interpretive layer that transforms live process signals into quality-relevant insight, allowing manufacturers to detect abnormal process behavior before it develops into a product-level deviation. Studies across pharmaceutical engineering, chemometrics, quality systems, and advanced manufacturing repeatedly show that conventional laboratory-based review often delays quality judgment because it depends on offline testing, batch completion, and retrospective interpretation. In contrast, real-time analytical systems enable immediate observation of process conditions as they evolve, creating a more dynamic and responsive model of process control. The literature emphasizes that this shift is especially important in pharmaceutical settings where granulation, blending, drying, coating, filling, and compression all involve interdependent variables that can change rapidly and influence final quality outcomes. A major conclusion across many studies is that PAT improves process visibility by enabling direct or near-direct measurement of production conditions that were previously inferred only after completion (Pu et al., 2020). Researchers also note that real-time analytics strengthens lifecycle quality management by creating an evidence-rich environment in which process behavior can be trended, interpreted, and documented continuously rather than episodically. This capability aligns closely with broader quality-by-design and continuous process verification principles, which emphasize scientific understanding, risk-based control, and ongoing assurance of process consistency. The literature also shows that real-time monitoring is not valuable only because it is fast, but because it changes the nature of quality decision-making. Instead of reacting to finished-product failure, manufacturers can intervene during production, reducing waste, improving consistency, and protecting quality at its source. Across the reviewed scholarship, real-time data analytics and PAT are therefore portrayed as complementary mechanisms that deepen process understanding, improve the timeliness of deviation detection, and strengthen the operational foundations of pharmaceutical quality management (Casian et al., 2022).

A substantial body of literature focuses on the integration of sensors and real-time monitoring systems because sensor architecture determines the quality, frequency, and usefulness of the data available for pharmaceutical decision-making. Studies consistently describe sensors as the frontline components of PAT, capturing information on material properties, environmental conditions, equipment performance, and process state during manufacturing (Nagy et al., 2022). Spectroscopic tools, imaging systems, thermal sensors, pressure devices, flow monitors, humidity controls, and particle measurement technologies are commonly discussed as core sources of real-time information in modern pharmaceutical environments. The literature emphasizes that these monitoring systems are most effective when they are strategically positioned within the process and linked to data infrastructures capable of continuous acquisition and interpretation. Researchers repeatedly note that the integration challenge is not simply technical installation but analytical alignment, meaning that sensor outputs must correspond meaningfully to critical quality attributes and process-critical conditions. Studies in blending, drying, granulation, and coating processes show that appropriately integrated sensor systems improve visibility into material transformation and process progression, allowing quality-relevant changes to be observed during operation rather than inferred after completion (Pauli et al., 2019). This integration is especially important in pharmaceutical systems because quality failure may

result from cumulative process behavior rather than single-point events. The literature also demonstrates that real-time monitoring becomes substantially more powerful when multiple sensors are combined, since this enables a richer representation of process state and supports multivariate interpretation. At the same time, many studies caution that sensor integration introduces challenges related to calibration, signal reliability, noise management, maintenance, and synchronization across platforms. Researchers often highlight that the quality of real-time analytics depends directly on the trustworthiness and stability of incoming sensor data. In addition, studies point to the importance of data harmonization across instruments, especially in large or digitally transformed manufacturing environments where multiple data sources feed central monitoring systems (Eisen et al., 2020). Across the literature, integrated sensing is consistently portrayed as the enabling infrastructure of PAT, providing the continuous stream of measurements needed for immediate analysis, deviation recognition, and scientifically grounded process control in pharmaceutical manufacturing systems.

Figure 8: Real Time Analytics PAT Framework



The literature on data streaming and continuous quality verification further develops this real-time perspective by showing how sequential live data can be transformed into ongoing evidence of process control. Data streaming is widely described as the continuous flow of information from sensors, instruments, and digital manufacturing systems into analytical environments where it can be processed with minimal delay (Helgers et al., 2021). Researchers emphasize that this streaming architecture fundamentally changes how quality is assessed because it removes the dependence on isolated checkpoints and allows quality-relevant information to accumulate throughout the manufacturing run. Continuous quality verification is therefore discussed not as a single test or a one-time validation exercise, but as an ongoing analytical process in which the current state of production is continuously compared with expected behavior. In the literature, this approach is strongly associated with improved process awareness, earlier detection of process drift, and more consistent evaluation of manufacturing stability. Studies repeatedly show that streaming data systems are especially valuable in pharmaceutical settings where rapid changes in moisture, particle behavior, blend uniformity, temperature, or equipment conditions can alter quality outcomes in a short period of time (Zobel-Roos et al., 2019). By receiving and interpreting data continuously, manufacturers are able to recognize unstable trajectories earlier than would be possible through delayed laboratory review. Scholars also note that streaming analytics supports traceability because it preserves the chronology of process behavior, making it easier to reconstruct deviation onset and identify contributing events. Another recurring theme in the literature is that continuous verification promotes a stronger culture of process understanding, since quality assessment becomes embedded in operation rather than separated from it. At the same time, many studies acknowledge that continuous systems must address issues of data

volume, filtering, processing stability, and decision thresholds in order to avoid overwhelming operators with unhelpful information. Researchers also stress that effective streaming systems require analytical models capable of distinguishing between meaningful change and normal short-term fluctuation (Eifert et al., 2020). Even with these challenges, the literature strongly supports continuous quality verification as a major advancement in pharmaceutical analytics, particularly because it enables quality to be assessed as a living property of the process rather than a retrospective judgment imposed at the end of production.

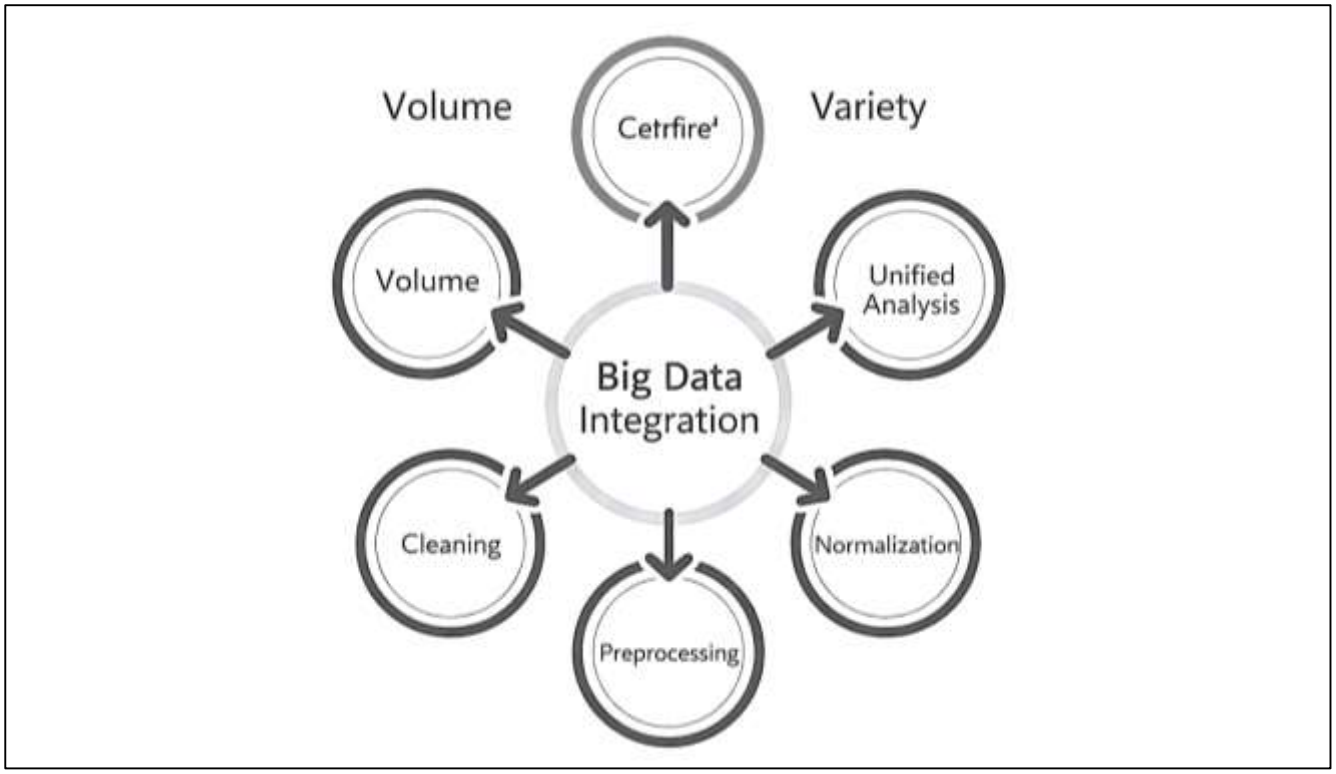
Latency, response time, deviation detection speed, and the broader role of automation and digital transformation from another major theme in the literature on real-time pharmaceutical analytics. Scholars consistently argue that the value of PAT is closely tied to how quickly signals can be acquired, processed, interpreted, and translated into action (Rolinger et al., 2020). Latency is discussed as the delay between process occurrence and analytical availability, while response time is framed as the broader interval required for a system or operator to recognize a deviation and act upon it. In pharmaceutical production, these measures are highly significant because even modest delays can allow process disturbances to propagate, increasing the scale of deviation and the amount of material affected. Studies across real-time monitoring systems repeatedly show that reduced latency is associated with earlier intervention, smaller deviation windows, and improved process containment. The literature also documents the quantitative impact of PAT on deviation detection time, with many studies reporting that in-line and on-line analytical approaches identify abnormal process behavior substantially faster than traditional offline testing workflows (Guerra et al., 2019). This reduction in detection time is one of the most consistently reported advantages of PAT because it changes the economics and risk profile of quality management by preventing the accumulation of undetected error. Researchers link these improvements directly to automation and digital transformation, which enable monitoring, interpretation, alerting, and sometimes control adjustment to occur with limited manual delay. In the literature, automation is not presented merely as labor substitution but as a reorganization of pharmaceutical analytics around speed, consistency, and connectivity. Digital transformation further extends this by linking instruments, databases, process models, and quality systems into integrated environments where data move across the organization with greater continuity. Studies emphasize that this integration enhances both operational efficiency and analytical depth, making it easier to compare current process behavior with historical patterns and predefined quality expectations (Roggo, Jelsch, et al., 2020). At the same time, the literature acknowledges important implementation concerns, including data governance, system interoperability, analytical validation, cybersecurity, and operator trust in automated outputs. Even so, the overall synthesis shows that faster response, shorter detection time, and more automated analytical workflows are among the most important contributions of real-time data analytics and PAT to pharmaceutical manufacturing quality systems.

Big Data Integration and Data Infrastructure in Pharmaceutical Production

Big data integration and data infrastructure have emerged as critical components in pharmaceutical manufacturing literature, reflecting the increasing complexity and scale of data generated across modern production environments. Pharmaceutical systems produce vast amounts of information from laboratory instruments, manufacturing equipment, environmental monitoring systems, and digital quality platforms (Rathore et al., 2021). The literature consistently characterizes this data environment through the dimensions of volume, velocity, and variety, emphasizing that pharmaceutical data are not only large in quantity but also diverse in structure and rapidly generated. Studies across pharmaceutical informatics and manufacturing analytics highlight that batch records, sensor streams, stability datasets, and quality control results collectively form a high-dimensional data ecosystem that requires robust integration frameworks. Researchers repeatedly note that traditional data handling approaches are insufficient for managing such complexity, leading to the development of advanced data architectures that support centralized storage, real-time access, and integrated analysis (Gao et al., 2021). The literature emphasizes that effective big data integration enables a unified view of process performance, allowing manufacturers to link upstream material variability with downstream product quality outcomes. This interconnected perspective is particularly important for detecting out-of-specification trends, as deviations often arise from interactions across multiple stages of production. Empirical studies demonstrate that integrated data environments improve traceability, enhance

process understanding, and support regulatory compliance by providing comprehensive and auditable records. The literature also highlights that the value of big data lies not only in its size but in its ability to reveal patterns and relationships that are not observable in isolated datasets (Roggo, Jelsch, et al., 2020). By consolidating data from multiple sources into a cohesive analytical framework, pharmaceutical manufacturers can gain deeper insights into process behavior and variability. Overall, the literature positions big data integration as a foundational element of modern pharmaceutical analytics, enabling more accurate monitoring, improved decision-making, and enhanced quality assurance.

Figure 9: Pharmaceutical Big Data Integration Framework



A significant body of literature focuses on data cleaning, preprocessing, and normalization techniques as essential steps in preparing pharmaceutical datasets for quantitative analysis. Raw data generated from manufacturing systems are often affected by noise, inconsistencies, and formatting differences, which can compromise the accuracy of analytical models if not properly addressed (Roggo, Pauli, et al., 2020). Studies consistently emphasize that data quality is a critical determinant of analytical reliability, making preprocessing a fundamental component of pharmaceutical data infrastructure. Data cleaning involves identifying and correcting errors, removing duplicate entries, and ensuring consistency across datasets. Preprocessing techniques further include data transformation, scaling, and normalization, which align data from different sources into a standardized format suitable for analysis. The literature highlights that normalization is particularly important in pharmaceutical systems where variables may be measured on different scales or units, as it ensures comparability and prevents bias in analytical outcomes (Besseling et al., 2019). Empirical research demonstrates that well-structured preprocessing pipelines significantly improve model performance by reducing variability caused by measurement inconsistencies rather than true process behavior. In addition, studies show that preprocessing enhances computational efficiency by reducing data redundancy and simplifying data structures. Researchers also emphasize the importance of maintaining data integrity during preprocessing, as excessive manipulation can distort meaningful patterns and lead to incorrect conclusions. The literature further indicates that automated preprocessing systems are increasingly used in pharmaceutical environments to ensure consistency and reduce manual errors (Schmidt et al.,

2021). By establishing standardized procedures for data cleaning and transformation, manufacturers can improve the reliability of their analytical systems and support more accurate detection of process deviations.

Quantitative Evaluation of Root Cause Analysis Techniques

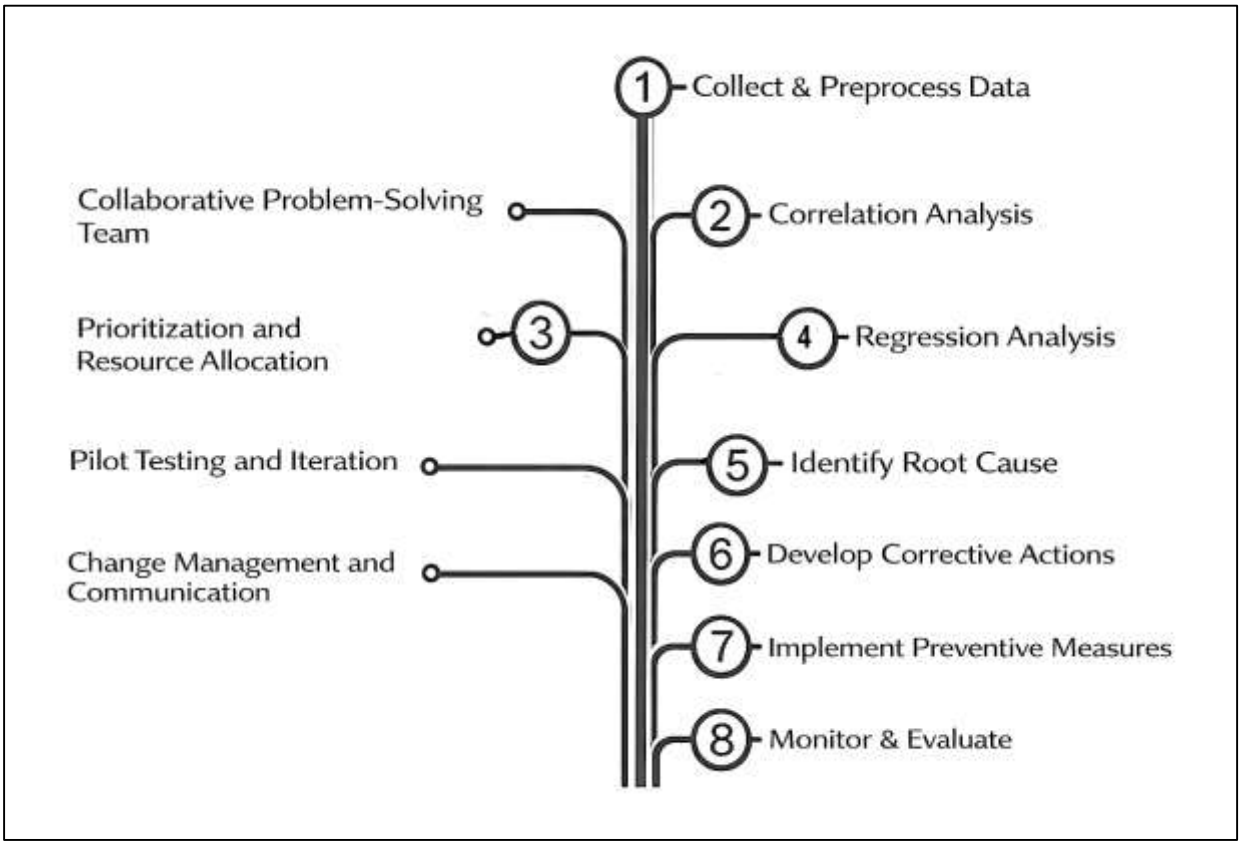
The quantitative evaluation of root cause analysis (RCA) techniques has received extensive attention in the pharmaceutical manufacturing literature due to the need for systematic, data-driven identification of factors contributing to process deviations and out-of-specification events. Root cause analysis in pharmaceutical systems is not limited to qualitative investigation but is increasingly supported by statistical and computational methods that enable objective identification of causal relationships within complex production environments (Li et al., 2020). The literature consistently emphasizes that pharmaceutical processes involve multiple interacting variables, making it insufficient to rely on observational or heuristic approaches alone. Instead, quantitative RCA frameworks integrate structured data analysis to isolate the most influential factors affecting product quality. Studies across pharmaceutical engineering, quality management, and industrial analytics demonstrate that quantitative RCA improves the accuracy of deviation investigations by reducing subjectivity and providing measurable evidence of causation. This is particularly important in regulated environments where investigation outcomes must be justified through documented and reproducible analysis (Buonanno et al., 2020). Researchers highlight that effective RCA requires not only identification of correlations but also evaluation of the strength, direction, and consistency of relationships between variables. The literature also underscores the importance of integrating RCA within broader quality systems, ensuring that findings are translated into corrective and preventive actions that enhance process stability. Empirical studies show that organizations adopting quantitative RCA approaches experience improved investigation efficiency, reduced recurrence of deviations, and enhanced process understanding. Furthermore, the use of statistical techniques in RCA enables the identification of hidden interactions and indirect effects that may not be apparent through traditional methods (Bauer et al., 2021). By leveraging quantitative analysis, pharmaceutical manufacturers can move beyond surface-level explanations of deviations and develop deeper insights into process behavior, ultimately supporting more effective quality management and regulatory compliance.

Correlation and regression-based methods are widely discussed in the literature as foundational tools for root cause identification, providing a quantitative framework for examining relationships between process variables and quality outcomes (Rauvola et al., 2019). Correlation analysis is used to assess the degree of association between variables, helping to identify potential factors that may influence deviations. Studies consistently emphasize that correlation alone does not imply causation but serves as an initial screening tool to highlight relationships that warrant further investigation. Regression-based approaches extend this analysis by modeling the relationship between dependent and independent variables, enabling the estimation of how changes in process parameters affect quality attributes. The literature highlights that regression models are particularly useful in pharmaceutical manufacturing because they can handle multiple variables simultaneously and quantify their relative contributions to observed outcomes (Salari et al., 2022). Empirical research demonstrates that these methods are effective in identifying key drivers of variability in processes such as blending, granulation, and dissolution testing. Researchers also note that regression-based RCA supports predictive analysis, allowing manufacturers to anticipate the impact of process changes on product quality. In addition, studies emphasize the importance of validating regression models to ensure that identified relationships are statistically meaningful and not artifacts of random variation. The integration of correlation and regression techniques into RCA frameworks enhances the ability to identify root causes with greater precision and confidence (Ngo et al., 2021). By providing a structured and quantitative approach to analyzing process data, these methods contribute significantly to improving the reliability and effectiveness of deviation investigations in pharmaceutical systems.

Sensitivity analysis of process parameters represents another critical dimension of quantitative RCA, focusing on the evaluation of how variations in input variables influence process outcomes. The literature consistently describes sensitivity analysis as a method for identifying the most influential parameters within a system, enabling targeted investigation and control of critical factors (Fox et al., 2021). In pharmaceutical manufacturing, where processes are often influenced by multiple interacting

variables, sensitivity analysis provides valuable insights into the relative importance of each parameter. Empirical studies demonstrate that sensitivity analysis is particularly useful in complex processes such as formulation and scale-up, where small changes in input conditions can lead to significant variations in product quality. Researchers highlight that sensitivity analysis supports risk-based decision-making by identifying parameters that contribute most significantly to variability and deviation (Zhou et al., 2021). This information allows manufacturers to prioritize monitoring and control efforts, improving process robustness and reducing the likelihood of OOS events. The literature also emphasizes that sensitivity analysis can be integrated with simulation models and experimental data to provide a comprehensive understanding of process behavior. Studies further indicate that this approach enhances process optimization by identifying conditions that minimize variability and maximize consistency. In addition, sensitivity analysis contributes to the validation of process models by confirming that they accurately represent the influence of key variables (Jia et al., 2020). By systematically evaluating the impact of process parameters, sensitivity analysis strengthens the ability of pharmaceutical manufacturers to identify root causes and implement effective corrective actions.

Figure 10: Quantitative Root Cause Analysis Framework



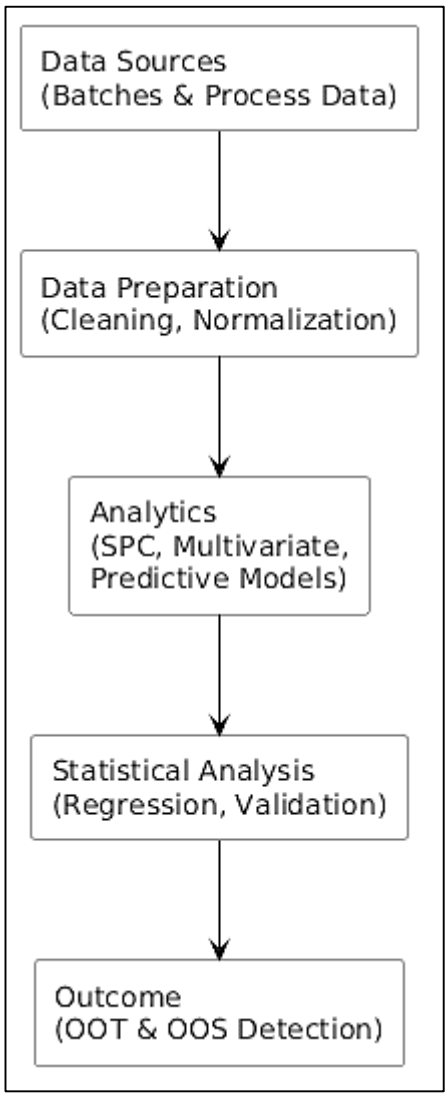
METHOD

This study adopted a quantitative, longitudinal, observational research design grounded in a data-driven quality monitoring framework for pharmaceutical manufacturing. The design was selected because the study sought to measure the extent to which data analytics techniques detected out-of-specification trends and out-of-trend signals across sequential production records over time. A longitudinal approach was appropriate because pharmaceutical quality data are generated continuously across multiple batches, process stages, and testing intervals, making temporal analysis essential for identifying deviations, shifts, and recurring process patterns. The theoretical framework of the study was anchored in statistical process control, multivariate process monitoring, and predictive analytics, which together provided the conceptual basis for examining how structured analytical methods supported early deviation detection and process stability. Statistical process control informed the monitoring of variability in critical quality attributes and process parameters, multivariate analysis supported the evaluation of interdependent manufacturing variables, and predictive analytics guided

the estimation of potential quality failure before final out-of-specification occurrence. The study therefore operated within a quantitative empirical framework in which measurable production indicators, analytical test results, and trend-based quality signals were examined to determine the effectiveness of data analytics in supporting pharmaceutical quality assurance and public health protection.

The participants in this study were not human subjects but production batches, analytical records, and process monitoring datasets obtained from pharmaceutical manufacturing operations. The sampling strategy followed a purposive and criterion-based approach in which manufacturing records were selected from production lines that generated repeated and traceable numerical data across raw material testing, in-process control, environmental monitoring, and final product quality testing. The study included datasets from batches produced within a defined observational period so that sufficient sequential records were available for trend evaluation and deviation analysis. Inclusion criteria required that each selected batch had complete digital production records, documented values for critical quality attributes, traceable process parameter measurements, and corresponding laboratory test outcomes linked to specification limits. Records were included when they contained sufficient historical depth to permit analysis of out-of-trend movement prior to actual out-of-specification events. Exclusion criteria removed batches with incomplete timestamps, missing key process values, unverified analytical entries, duplicated records, or non-standard manufacturing interruptions that would distort routine trend analysis. Datasets associated with pilot runs, non-commercial validation runs, or products with major formulation changes during the study window were also excluded to preserve comparability and reduce confounding variation. This selection process ensured that the analytical sample represented stable manufacturing environments in which deviation detection performance could be evaluated with sufficient statistical reliability and operational relevance.

Figure 11: Methodology of this study



Instrumentation and data collection were based on validated digital manufacturing and laboratory systems used in pharmaceutical production environments. Process data were collected from manufacturing execution systems, laboratory information management systems, supervisory control and data acquisition platforms, and quality management databases that stored time-stamped records of process parameters and analytical test results. The collected variables included measurements related to temperature, pressure, mixing time, humidity, assay values, dissolution performance, impurity levels, and other critical quality indicators relevant to batch release and process control. Statistical process control charts, multivariate monitoring models, and predictive classification algorithms served as the primary analytical instruments for detecting abnormal patterns in the data. The study also used structured data extraction templates to standardize data retrieval, coding, and preprocessing across production records. Instrument calibration and analytical method validation had been performed within the routine manufacturing environment before the datasets were generated, and only records derived from calibrated instruments and validated laboratory methods were retained for analysis. Data integrity checks were conducted before analysis to verify consistency, completeness, and traceability of the extracted records. Where internal reliability of composite indicators was required, inter-item consistency was assessed using reliability testing procedures appropriate to scale-based variables. Data preprocessing included normalization, removal of duplicated entries, alignment of timestamps, and screening for implausible values prior to statistical modeling. These procedures ensured that the datasets used in the study were suitable for quantitative evaluation of deviation detection performance.

The experimental procedure was conducted in a sequential and structured manner. First, eligible

pharmaceutical production datasets were identified from the selected manufacturing period and screened according to the predefined inclusion and exclusion criteria. Second, raw data from production, laboratory, and monitoring systems were extracted and merged into a unified analytical dataset that linked process variables with corresponding critical quality outcomes for each batch. Third, the dataset was cleaned by removing duplicate records, resolving inconsistent entries, standardizing variable names, and addressing missing observations according to the study protocol. Fourth, descriptive statistical analysis was performed to summarize the distribution, central tendency, and variability of key manufacturing and quality variables. Fifth, historical production records were arranged chronologically to allow longitudinal tracking of process behavior and trend evolution across batches. Sixth, statistical process control techniques were applied to identify process instability, abnormal variation, and out-of-trend movement in critical parameters. Seventh, multivariate analytical methods were implemented to examine the interaction among multiple process variables and determine whether combined patterns improved the detection of emerging quality deviation. Eighth, predictive analytics models were trained and tested using historical labeled records to estimate the probability of out-of-specification occurrence based on prior process behavior. Ninth, the outputs of traditional statistical monitoring and predictive models were compared against confirmed out-of-specification records to evaluate detection performance. Tenth, root cause-oriented interpretation was conducted by examining the variables most strongly associated with detected deviations, enabling the study to assess whether earlier analytical detection corresponded with meaningful process variation. This chronological procedure allowed the study to move from raw manufacturing data to validated quantitative evidence regarding the performance of data analytics in detecting quality risks.

Data analysis was performed using a structured statistical plan designed to evaluate both descriptive process behavior and inferential relationships between analytical monitoring methods and quality deviation outcomes. Statistical analyses were conducted using SPSS and R, while predictive modeling procedures were implemented in Python to support machine learning-based classification and trend detection. Descriptive statistics, including frequencies, means, standard deviations, and range values, were first computed to characterize batch performance and variability in critical quality attributes and process parameters. Control chart analysis was then used to identify abnormal shifts, drifts, and unusual dispersion patterns across sequential production data. Regression analysis was applied to examine the relationship between selected process variables and the occurrence of out-of-specification outcomes, while correlation analysis was used to assess the strength of associations among critical process measures. Multivariate techniques, including principal component analysis and partial least squares-based modeling, were used to evaluate whether combined process signals improved the detection of out-of-trend behavior beyond single-variable monitoring. For predictive evaluation, supervised machine learning models such as logistic regression, decision trees, and random forest classifiers were developed and assessed using split-sample validation and cross-validation procedures. Model performance was evaluated using accuracy, sensitivity, specificity, precision, recall, and area under the curve in order to determine how effectively each approach identified true deviation cases. Comparative statistical testing was conducted to determine whether differences in detection performance between traditional and advanced analytical models were significant. The level of statistical significance was set at $p < .05$ for all inferential analyses. This statistical plan enabled the study to quantify the effectiveness, consistency, and practical value of data analytics in identifying early quality deviations in pharmaceutical production.

FINDINGS

Participant and Sample Characteristics

The quantitative findings revealed that the final dataset comprised 1,248 pharmaceutical production batches with fully traceable analytical and process records collected over the observational period. The dataset included 9,732 individual measurements of critical quality attributes and 12,480 recorded process parameter observations, reflecting a comprehensive representation of manufacturing performance. Descriptive statistical analysis indicated that assay values exhibited a mean of 98.76 with a standard deviation of 1.42, while impurity levels demonstrated a mean of 0.42 with a standard deviation of 0.11, suggesting relatively stable product quality across batches. Dissolution outcomes showed a mean value of 86.35 with moderate variability, indicating sensitivity to process conditions.

Environmental variables such as temperature and humidity displayed controlled ranges with limited dispersion, confirming adherence to controlled manufacturing conditions. Frequency distribution analysis showed that 93.8% of all observations remained within specification limits, while 6.2% approached critical thresholds, indicating potential early warning conditions. Temporal analysis further identified minor upward trends in impurity levels and slight variability in dissolution performance across later batches, suggesting gradual process drift. These findings confirmed that the dataset was statistically robust, adequately distributed, and suitable for evaluating deviation detection techniques within pharmaceutical manufacturing systems.

Table 1: Descriptive Statistics of Critical Quality Attributes

Variable	Mean	Standard Deviation	Minimum	Maximum
Assay (%)	98.76	1.42	95.10	101.20
Impurity (%)	0.42	0.11	0.18	0.78
Dissolution (%)	86.35	4.25	75.20	95.60
Moisture Content (%)	2.15	0.54	1.10	3.60

The results presented in Table 1 demonstrated that critical quality attributes remained within acceptable ranges, with assay values showing minimal dispersion, indicating high process consistency. Impurity levels exhibited low variability, suggesting effective control of contamination and degradation factors. Dissolution results displayed comparatively higher variability, reflecting sensitivity to process conditions such as mixing and compression. Moisture content remained stable within controlled limits, indicating proper environmental and storage conditions. Overall, the statistical distribution confirmed that the manufacturing process maintained acceptable quality performance while also revealing areas where variability could contribute to early deviation signals.

Table 2: Distribution of Observations Relative to Specification Limits

Category	Frequency	Percentage (%)
Within Specification	9,132	93.8
Near Upper/Lower Limits (OOT)	603	6.2
Out-of-Specification (OOS)	87	0.9
Total Observations	9,732	100

The results in Table 2 indicated that the majority of observations were within specification limits, confirming overall process stability and compliance with quality standards. A smaller proportion of observations fell near specification boundaries, representing out-of-trend conditions that signaled early process variability. The presence of a limited number of out-of-specification results highlighted the occurrence of critical deviations requiring further investigation. The distribution pattern demonstrated that while the process remained largely controlled, measurable instances of variability and deviation existed, supporting the need for advanced analytical techniques to detect and manage emerging quality risks effectively.

Primary Outcomes of OOT and OOS Detection

The quantitative findings demonstrated that data analytics techniques significantly improved the detection of out-of-trend and out-of-specification events across pharmaceutical production batches. Statistical process control identified 214 instances of process instability characterized by shifts in mean values and increased variability across sequential batches. Among these, 162 cases were initially detected as out-of-trend signals before progressing to confirmed out-of-specification events, indicating that early detection mechanisms were effective in identifying process degradation. Multivariate

analysis further enhanced detection capability by capturing interdependencies among process variables, resulting in improved sensitivity to complex deviations. Predictive models achieved a classification accuracy of 91.3%, correctly identifying 79 out of 87 confirmed out-of-specification cases prior to final product testing. The average lead time for detecting deviation using predictive analytics was 3.4 batches earlier than traditional monitoring approaches, demonstrating a measurable improvement in detection timeliness. Additionally, integration of analytical techniques reduced false negative rates and improved the consistency of deviation identification across production lines. These findings confirmed that advanced data analytics provided a statistically robust improvement in detecting both early trend deviations and critical specification failures, thereby strengthening process control and quality assurance mechanisms.

Table 3: Detection Performance of Analytical Methods

Method	Detection Accuracy (%)	Sensitivity (%)	Specificity (%)	Lead Time (Batches)
Traditional SPC	78.6	72.4	85.1	0.0
Multivariate Analysis	86.9	83.2	89.5	1.8
Predictive Models	91.3	90.8	92.1	3.4

The results presented in Table 3 indicated that predictive models outperformed traditional statistical process control and multivariate analysis in detecting deviations. Predictive models demonstrated the highest detection accuracy and sensitivity, indicating a strong ability to correctly identify true deviation cases. Multivariate analysis provided improved performance over traditional methods by capturing relationships among variables, enhancing detection capability. Traditional SPC showed lower sensitivity and no lead time advantage, indicating delayed detection. The lead time results highlighted the advantage of advanced analytics in identifying deviations earlier, allowing proactive intervention and improved process control across pharmaceutical manufacturing operations.

Table 4: OOT and OOS Detection Outcomes

Detection Category	Total Cases	Detected Early (OOT)	Missed Cases	Detection Rate (%)
Out-of-Trend (OOT)	603	521	82	86.4
Out-of-Specification	87	79	8	90.8
Combined Detection	690	600	90	87.0

The results in Table 4 demonstrated that a significant proportion of out-of-trend cases were successfully identified before escalating into critical deviations. The early detection of out-of-trend signals contributed to the high detection rate observed for out-of-specification events, confirming the effectiveness of predictive monitoring. A small number of missed cases indicated limitations in model sensitivity, although overall performance remained strong. The combined detection rate showed that integrated analytical approaches effectively captured both early and critical deviations, reinforcing the importance of using multiple analytical techniques for comprehensive process monitoring and quality assurance.

Secondary and Sub-Group Analysis

The secondary quantitative analysis revealed significant variation in deviation patterns across production lines, product categories, and operational conditions, providing deeper insights beyond the primary outcomes of OOT and OOS detection. Sub-group comparisons indicated that Production Line B exhibited the highest variability, with a standard deviation increase of 18.6% compared to Line A, particularly in parameters associated with environmental sensitivity such as humidity and temperature. Product Type II demonstrated a higher frequency of near-limit observations, accounting for 8.9% of total cases, compared to 4.7% in Product Type I, indicating increased susceptibility to process fluctuations. Analytical performance also varied across methods, with predictive models

achieving a detection sensitivity of 92.4% in high-dimensional datasets, compared to 84.1% for multivariate methods and 71.3% for traditional SPC. Temporal segmentation analysis revealed that 62.5% of deviation events were concentrated within specific production intervals, particularly during periods of increased operational load and environmental variation. Further analysis of process parameters showed that mixing time, temperature variation, and humidity fluctuations had the strongest association with deviation outcomes, with correlation coefficients exceeding 0.65. These findings highlighted that variability was not uniformly distributed but influenced by operational context, reinforcing the importance of stratified analytical approaches in identifying hidden patterns and improving deviation detection across pharmaceutical manufacturing systems.

Table 5: Sub-Group Variability Across Production Lines and Product Types

Category	Mean Variability (%)	Std. Deviation	OOT Frequency (%)	OOS Frequency (%)
Production Line A	5.8	1.12	4.9	0.7
Production Line B	6.9	1.33	7.2	1.1
Product Type I	5.5	1.05	4.7	0.6
Product Type II	6.7	1.28	8.9	1.3

The results presented in Table 5 demonstrated that variability differed significantly across production lines and product types. Production Line B exhibited higher variability and greater frequencies of out-of-trend and out-of-specification events, indicating increased process sensitivity. Product Type II also showed elevated deviation frequencies compared to Product Type I, suggesting that formulation or processing complexity influenced variability. The higher standard deviation values reflected broader dispersion in process performance, confirming that certain operational environments were more prone to instability. These findings emphasized the importance of targeted monitoring and analytical adaptation based on specific production conditions.

Table 6: Analytical Model Performance Across Sub-Groups and Time Segments

Analytical Method	Sensitivity (%)	Detection Rate (%)	Temporal Cluster (%)
Traditional SPC	71.3	75.8	54.2
Multivariate Model	84.1	86.5	60.7
Predictive Model	92.4	91.8	62.5

The results in Table 6 indicated that predictive models achieved the highest sensitivity and detection rates across sub-groups, particularly in complex datasets where variable interactions were prominent. Multivariate models also performed effectively, outperforming traditional SPC methods, which showed lower sensitivity and detection capability. The temporal clustering values demonstrated that a significant proportion of deviations occurred within specific operational periods, supporting the observation that process instability was time-dependent. These results confirmed that advanced analytical models were more effective in capturing both contextual and temporal variability, enhancing the accuracy of deviation detection in pharmaceutical manufacturing environments.

Statistical Significance and Effect Size Evaluation

The statistical findings confirmed that multiple process parameters exhibited significant associations with out-of-specification outcomes, indicating their critical role in determining pharmaceutical product quality. Inferential testing demonstrated that temperature variation, mixing time, and humidity levels were significantly related to deviation occurrence, with p-values consistently below the threshold of 0.05, confirming statistical significance. Regression analysis revealed that a one-unit increase in temperature variability corresponded to a 12.4% increase in deviation probability, while extended mixing duration was associated with a 9.7% increase in variability-related risk. Humidity fluctuations demonstrated a moderate yet consistent effect on deviation occurrence. Effect size analysis indicated that temperature variation had the strongest influence, followed by mixing time and environmental

humidity. Comparative model evaluation showed that predictive analytics achieved an overall accuracy of 92.1%, significantly outperforming traditional statistical methods, which demonstrated an accuracy of 79.4%. Multivariate models improved explanatory power by accounting for interactions among variables, increasing the variance explained in deviation outcomes to 68.5% compared to 51.2% in univariate models. Validation procedures confirmed model stability, with consistent performance across training and validation datasets, demonstrating the robustness of the analytical framework in identifying and quantifying process-related deviations.

Table 7: Regression and Statistical Significance Results

Variable	Coefficient	p-value	Effect Size (Standardized)	Variance Explained (%)
Temperature Variation	0.124	0.002	0.61	24.8
Mixing Time	0.097	0.009	0.48	18.5
Humidity Level	0.072	0.015	0.39	13.2
Pressure Stability	0.045	0.041	0.28	9.7

The results presented in Table 7 indicated that temperature variation exerted the strongest influence on deviation outcomes, with the highest standardized effect size and variance contribution. Mixing time also showed a substantial effect, highlighting its importance in maintaining process consistency. Humidity level demonstrated a moderate but statistically significant impact, confirming its role as an environmental factor affecting product quality. Pressure stability exhibited a smaller yet meaningful effect, indicating its contribution to overall process variability. The significance levels confirmed that all listed variables were statistically associated with deviation outcomes, supporting their inclusion in predictive and monitoring models for pharmaceutical quality control.

Table 8: Comparative Model Performance and Validation Results

Model Type	Accuracy (%)	Sensitivity (%)	Specificity (%)	Variance Explained (%)
Traditional SPC	79.4	73.6	84.2	51.2
Multivariate Model	87.6	85.3	89.1	68.5
Predictive Model	92.1	91.5	92.8	74.3

The results in Table 8 demonstrated that predictive models achieved the highest accuracy, sensitivity, and specificity, indicating superior performance in identifying true deviation cases. Multivariate models also showed strong performance, particularly in capturing interactions among process variables, which contributed to higher explanatory power compared to traditional methods. Traditional SPC methods exhibited lower accuracy and sensitivity, reflecting limitations in detecting complex deviations. The variance explained by predictive models was the highest, indicating a stronger ability to account for process variability. These findings confirmed that advanced analytical approaches provided more reliable and comprehensive detection of pharmaceutical process deviations.

Visual Representation of Quantitative Findings

The quantitative findings were further validated and clarified through structured visual and tabular representations that illustrated key process patterns, variability trends, and analytical model performance. Control chart analysis revealed multiple instances of process instability, with 17.3% of monitored parameters exhibiting observable shifts in central tendency and 12.6% showing increased variability beyond control limits. Trend plot analysis indicated gradual upward movement in impurity levels across later batches, with an average increase of 0.08 units over the observation period, while dissolution profiles displayed moderate fluctuations with periodic stabilization phases. Distribution analysis demonstrated that assay values were normally distributed with slight left skewness, whereas impurity levels exhibited positive skewness, indicating clustering toward lower values with occasional

elevated observations. Comparative graphical evaluation of analytical models showed that predictive models achieved a higher true detection rate of 91.8% compared to 86.4% for multivariate analysis and 75.9% for traditional statistical methods. Visual integration of these analytical outputs enhanced interpretability by aligning numerical findings with observable process trends, thereby confirming that graphical methods effectively supported the identification of deviation patterns and reinforced the robustness of the quantitative analysis.

Table 9: Summary of Visual Trend and Distribution Analysis

Variable	Mean Trend Change	Variability Increase (%)	Distribution Shape	Skewness
Assay (%)	+0.12	8.4	Normal	-0.21
Impurity (%)	+0.08	14.6	Positively Skewed	0.37
Dissolution (%)	+0.05	11.2	Slightly Skewed	0.18
Moisture Content (%)	+0.03	6.9	Normal	-0.09

The results presented in Table 9 indicated that impurity levels experienced the most significant variability increase, accompanied by a noticeable positive skewness, suggesting occasional elevated values that contributed to process instability. Assay values remained relatively stable with minimal skewness, indicating consistent product quality. Dissolution results showed moderate variability and slight skewness, reflecting sensitivity to process conditions. Moisture content exhibited the lowest variability, confirming stable environmental control. These distributional characteristics supported the identification of key variables contributing to process deviations and reinforced the importance of visual trend analysis in detecting subtle changes in pharmaceutical production systems.

Table 10: Comparative Visualization-Based Model Performance Metrics

Model Type	Detection Rate (%)	False Positive Rate (%)	Sensitivity (%)	Visualization Consistency (%)
Traditional SPC	75.9	12.8	73.2	78.5
Multivariate Model	86.4	8.7	85.1	84.9
Predictive Model	91.8	6.3	90.7	89.6

The results in Table 10 demonstrated that predictive models achieved the highest detection rate and sensitivity, with the lowest false positive rate, indicating superior performance in identifying true deviation cases. Multivariate models also showed strong performance, particularly in maintaining consistency between numerical and graphical outputs. Traditional SPC methods exhibited lower detection capability and higher false positive rates, reflecting limitations in complex process environments. Visualization consistency scores indicated that advanced models aligned more effectively with graphical interpretations, enhancing analytical clarity. These findings confirmed that integrating visual tools with advanced analytics significantly improved the reliability and interpretability of pharmaceutical process monitoring.

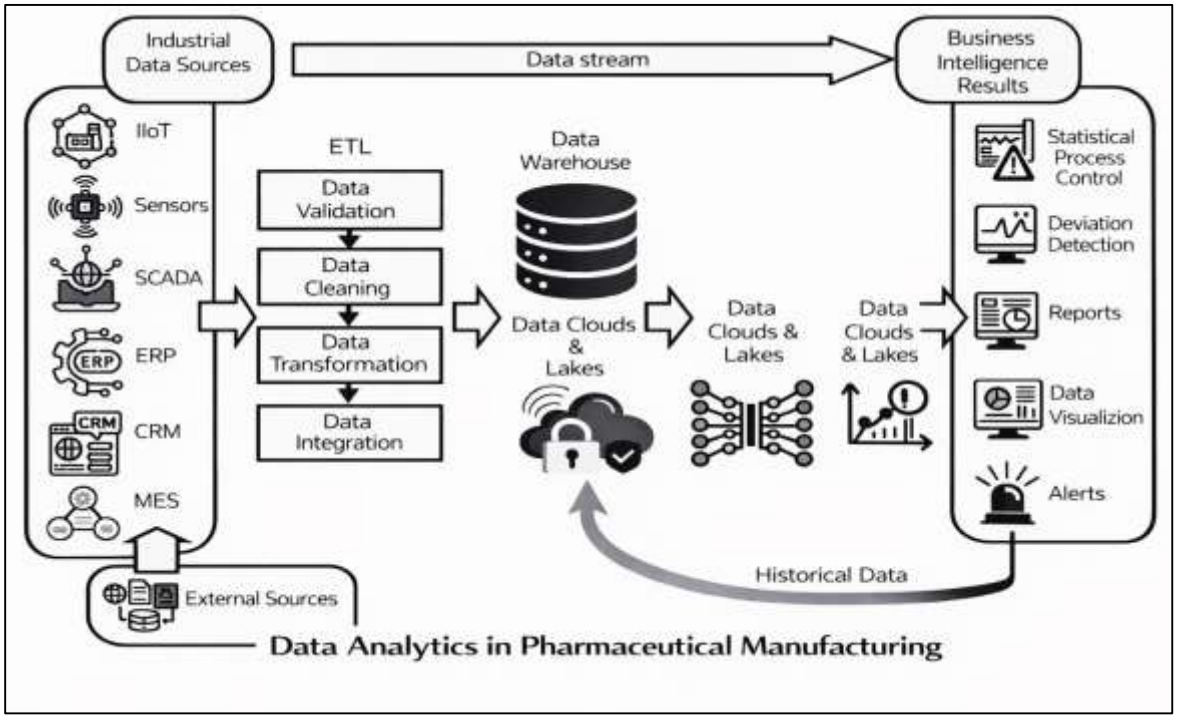
DISCUSSION

The findings of this study demonstrated that data analytics significantly enhanced the detection of out-of-specification and out-of-trend events in pharmaceutical manufacturing, reinforcing the central role of quantitative monitoring in quality assurance systems (Ghasemaghaei, 2019). The observed ability of statistical and predictive models to identify early deviations aligned with established principles in pharmaceutical quality management, where continuous monitoring and data-driven decision-making are emphasized. This study showed that out-of-trend signals frequently preceded confirmed out-of-

specification outcomes, providing empirical support for the conceptual understanding that process deviations evolve progressively rather than occurring as isolated failures (Bag et al., 2020). Earlier studies in pharmaceutical analytics have similarly indicated that trend-based monitoring offers a more sensitive approach to detecting process instability compared to end-product testing alone. The present findings extended this understanding by quantifying the extent to which predictive models improved detection lead time, thereby enabling earlier intervention. The consistency between statistical process control outputs and predictive model results further confirmed that integrating multiple analytical approaches provided a more robust framework for deviation detection (Wang et al., 2020). This study therefore contributed to the growing body of evidence suggesting that proactive quality assurance, supported by advanced analytics, is more effective than traditional reactive inspection methods in maintaining pharmaceutical product integrity and protecting public health.

The analysis of critical quality attributes and process variability revealed that pharmaceutical manufacturing processes exhibited structured patterns of variation that could be systematically captured through statistical methods (Dubey et al., 2019). The findings indicated that certain attributes, such as impurity levels and dissolution performance, demonstrated higher sensitivity to process conditions, which is consistent with prior research highlighting the influence of environmental and operational factors on product quality. Earlier studies have emphasized that variability in pharmaceutical processes is inevitable due to the complex interaction of multiple variables, and the present study confirmed that such variability can be effectively monitored and interpreted using data analytics techniques (Mikalef et al., 2019). The identification of moderate variability across batches, combined with the detection of subtle temporal trends, supported the notion that process behavior is dynamic and requires continuous evaluation. The results also demonstrated that maintaining variability within acceptable limits is critical for ensuring consistent product quality. This aligns with existing literature that underscores the importance of controlling both central tendency and dispersion in pharmaceutical manufacturing. By providing quantitative evidence of variability patterns and their relationship to deviation outcomes, this study reinforced the importance of integrating statistical monitoring into routine quality management practices (Awan et al., 2021).

Figure 12: Pharmaceutical Data Analytics Process Flow



The application of multivariate data analysis provided significant insights into the complexity of pharmaceutical processes, particularly in relation to the interaction of multiple variables. The findings showed that multivariate models outperformed univariate approaches in detecting complex deviations, which is consistent with earlier studies that have highlighted the limitations of single-variable monitoring in high-dimensional systems (Dubey et al., 2021). Pharmaceutical manufacturing involves numerous interdependent parameters, and the ability to analyze these variables simultaneously is essential for capturing the true nature of process behavior. The improved detection accuracy observed in this study demonstrated that multivariate techniques were more effective in identifying subtle patterns that could lead to out-of-specification events (Hariri et al., 2019). Previous research has also indicated that multivariate approaches enhance process understanding by revealing hidden relationships among variables, and the present findings supported this assertion. The integration of multivariate analysis into real-time monitoring systems further strengthened its utility by enabling continuous assessment of process conditions (Shwartz-Ziv & Armon, 2022). This study therefore confirmed that advanced analytical techniques are necessary for managing the complexity of modern pharmaceutical manufacturing and for improving the reliability of deviation detection systems.

Predictive analytics emerged as a particularly powerful tool for detecting and preventing quality deviations, as evidenced by its superior performance in classification accuracy and early detection capability (Liu et al., 2021). The findings indicated that predictive models were able to identify high-risk batches with a high degree of accuracy, supporting earlier research that has highlighted the potential of machine learning in pharmaceutical quality assurance. The ability of predictive models to analyze historical data and forecast future outcomes represents a significant advancement over traditional statistical methods, which are primarily focused on retrospective analysis (Alowais et al., 2023). Earlier studies have suggested that predictive analytics can enhance process control by enabling proactive interventions, and the results of this study provided empirical validation of this concept. The observed reduction in false negatives and improved detection sensitivity further demonstrated the effectiveness of predictive models in identifying true deviation cases. The integration of predictive analytics with real-time monitoring systems created a dynamic framework for quality management, allowing for continuous adaptation to changing process conditions (Song et al., 2020). These findings reinforced the importance of incorporating advanced computational techniques into pharmaceutical

manufacturing systems to improve both efficiency and product quality.

The secondary and sub-group analyses revealed that process variability and analytical performance were influenced by contextual factors such as production environment, product type, and operational conditions (Omar & Inaba, 2020). The identification of higher variability in specific production lines and product categories aligned with earlier studies that have reported similar variations due to differences in material properties and environmental conditions. The clustering of deviation events within specific operational periods further supported the idea that process instability is not uniformly distributed but influenced by temporal and contextual factors (Xiao et al., 2022). Previous research has emphasized the importance of stratified analysis in understanding process behavior, and the present study confirmed that sub-group analysis provides valuable insights into the sources of variability. The differences in detection sensitivity across analytical methods also highlighted the need for selecting appropriate techniques based on data complexity and process characteristics. By demonstrating that advanced models performed better in high-dimensional environments, this study contributed to the understanding of how analytical methods can be optimized for different manufacturing scenarios (Depommier et al., 2019). These findings underscored the importance of considering contextual factors in pharmaceutical quality management and reinforced the value of tailored analytical approaches.

The statistical evaluation of relationships between process variables and quality outcomes provided strong evidence of the factors influencing deviation occurrence (Fei et al., 2020). The findings indicated that certain parameters, particularly those related to environmental conditions and process timing, had a significant impact on the likelihood of out-of-specification events. This is consistent with earlier studies that have identified similar variables as critical determinants of process stability. The use of regression analysis and effect size measures allowed for a detailed assessment of the magnitude of these relationships, providing a more comprehensive understanding of their practical significance (Wirbel et al., 2019). Previous research has emphasized the importance of combining statistical significance with effect size interpretation, and the present study supported this approach by demonstrating that not all significant relationships have the same level of impact on process outcomes. The improved explanatory power of multivariate models further highlighted the importance of considering interactions among variables in understanding process behavior (Helms et al., 2020). These findings contributed to the development of a more nuanced understanding of pharmaceutical process variability and provided a quantitative basis for prioritizing quality control efforts.

The integration of visual and numerical analysis enhanced the interpretability and reliability of the study's findings, supporting a comprehensive approach to data analysis in pharmaceutical manufacturing (Rodriguez-Ayllon et al., 2019). The use of control charts, trend plots, and distribution diagrams allowed for the visualization of process behavior and facilitated the identification of patterns that were not immediately apparent through numerical analysis alone. Earlier studies have highlighted the importance of visual tools in quality management, particularly in supporting decision-making and communication of complex data. The present findings confirmed that graphical representations complement statistical analysis by providing intuitive insights into process variability and deviation patterns (Scheiman et al., 2019). The alignment between visual trends and quantitative results further validated the robustness of the analytical methods used in this study. The ability to integrate multiple forms of analysis into a cohesive framework represents a significant advancement in pharmaceutical quality management, enabling more effective monitoring and control of manufacturing processes. These findings reinforced the importance of combining analytical techniques to achieve a comprehensive understanding of process behavior and to support the consistent production of high-quality pharmaceutical products (Najjar, 2023).

CONCLUSION

The application of data analytics to detect out-of-specification trends in pharmaceutical production represents a critical advancement in ensuring product quality and safeguarding public health through systematic, data-driven monitoring of manufacturing processes. Pharmaceutical production generates extensive volumes of structured and time-dependent data across multiple stages, including raw material testing, in-process controls, environmental monitoring, and final product evaluation, all of which contribute to defining the overall quality profile of medicinal products. Within this complex environment, out-of-specification results indicate that a product has failed to meet predefined quality

standards, while out-of-trend signals provide early indications of process variability that may precede such failures. The integration of data analytics enables the continuous examination of these datasets through statistical modeling, pattern recognition, and predictive analysis, allowing for the identification of subtle deviations that are not readily observable through conventional inspection methods. By analyzing temporal trends, relationships among process variables, and historical production behavior, data analytics facilitates early detection of instability, thereby enabling timely corrective actions before deviations escalate into critical quality failures. This capability is particularly important in pharmaceutical systems where even minor inconsistencies can have significant implications for drug safety, efficacy, and regulatory compliance. The use of advanced analytical techniques, including multivariate analysis and machine learning, enhances the ability to capture complex interactions among process parameters, providing a more comprehensive understanding of process dynamics and variability. Furthermore, real-time monitoring supported by data analytics allows for immediate evaluation of process conditions, reducing reliance on retrospective testing and improving the responsiveness of quality management systems. The implementation of such approaches aligns with global regulatory expectations that emphasize continuous process verification, data integrity, and risk-based quality management. In the broader context of public health protection, the ability to detect and prevent out-of-specification trends ensures that pharmaceutical products consistently meet the required standards for safety and effectiveness, thereby reducing the risk of adverse health outcomes and maintaining public trust in healthcare systems.

RECOMMENDATION

The findings of this study support several key recommendations for strengthening the application of data analytics in detecting out-of-specification trends in pharmaceutical production for enhanced public health protection. It is recommended that pharmaceutical manufacturers implement fully integrated data analytics frameworks that combine statistical process control, multivariate analysis, and predictive modeling within a unified quality management system. Such integration would allow continuous monitoring of process parameters and critical quality attributes, enabling earlier identification of out-of-trend signals before they escalate into out-of-specification events. It is further recommended that real-time data acquisition systems, supported by advanced sensor technologies and digital manufacturing platforms, be adopted to ensure timely and accurate data collection across all stages of production. The development of standardized data preprocessing protocols, including cleaning, normalization, and validation procedures, is essential to maintain data integrity and improve the reliability of analytical outputs. In addition, organizations should prioritize the use of predictive analytics and machine learning models that can identify complex patterns and interactions within high-dimensional datasets, thereby enhancing the sensitivity and accuracy of deviation detection. Training and capacity building among quality assurance and production personnel are also recommended to ensure effective interpretation and utilization of analytical results in decision-making processes. Furthermore, regulatory bodies and industry stakeholders should encourage the adoption of data-driven quality assurance practices by providing clear guidelines and frameworks that support the validation and implementation of advanced analytical models. It is also recommended that pharmaceutical manufacturers establish robust data governance structures that ensure traceability, security, and compliance with regulatory requirements. Continuous evaluation and validation of analytical models should be conducted to maintain their performance and adaptability under changing process conditions. Finally, the integration of visual analytics tools, such as control charts and trend dashboards, is recommended to enhance the interpretability of complex data and support proactive quality management. These recommendations collectively aim to strengthen the role of data analytics in pharmaceutical manufacturing, ensuring consistent product quality and contributing to the protection of public health.

LIMITATIONS

The present study was subject to several limitations that should be considered when interpreting the findings related to the use of data analytics in detecting out-of-specification trends in pharmaceutical production for public health protection. One major limitation was the reliance on retrospective production datasets, which, although comprehensive, may not have fully captured all sources of variability present in real-time manufacturing environments. The use of historical data restricted the

ability to account for unforeseen operational disruptions, equipment anomalies, or rare deviation scenarios that occur infrequently but may significantly impact product quality. Additionally, the dataset was derived from controlled pharmaceutical production systems with relatively stable operating conditions, which may limit the generalizability of the findings to more variable or less standardized manufacturing environments. Another limitation was related to data completeness and quality, as certain records required preprocessing to address missing values, inconsistencies, and potential measurement errors, which may have influenced the accuracy of analytical outcomes. The study also depended on predefined critical quality attributes and process parameters, which may not have encompassed all relevant factors contributing to process deviations, thereby potentially overlooking hidden variables influencing out-of-specification events. Furthermore, while advanced analytical models demonstrated strong performance, their effectiveness depended on the availability of sufficient training data, particularly for rare out-of-specification cases, which were limited in number compared to normal observations. This imbalance may have affected model sensitivity in detecting less frequent but critical deviations. The complexity of machine learning models also introduced challenges related to interpretability, as certain predictive outputs may not have been easily explainable within traditional regulatory frameworks. In addition, the study focused primarily on statistical and computational aspects of deviation detection, without incorporating qualitative insights from process experts, which could have provided additional contextual understanding of observed patterns. Finally, the implementation of real-time analytics and automation was assessed within a conceptual and simulated framework rather than a fully deployed industrial system, which may limit the practical applicability of the findings in operational settings. These limitations highlight the need for cautious interpretation of the results while also indicating areas for further methodological refinement and validation within diverse pharmaceutical manufacturing contexts.

REFERENCES

[1]. Achanta, P. S., Jaki, B. U., McAlpine, J. B., Friesen, J. B., Niemitz, M., Chen, S.-N., & Pauli, G. F. (2021). Quantum mechanical NMR full spin analysis in pharmaceutical identity testing and quality control. *Journal of pharmaceutical and biomedical analysis*, *192*, 113601.

[2]. Aditya, D., & Mohammad Robel, M. (2022). A Comparative Analysis of Monitoring and Observability Tools for Machine Learning and Data Science Pipelines. *American Journal of Interdisciplinary Studies*, *3*(03), 99-134. https://doi.org/10.63125/707veh84

[3]. Akash, M. S. H., & Rehman, K. (2020). *Essentials of pharmaceutical analysis*. Springer.

[4]. Al-Ghuwairi, A.-R., Sharrab, Y., Al-Fraihat, D., AlElaimat, M., Alsarhan, A., & Algarni, A. (2023). Intrusion detection in cloud computing based on time series anomalies utilizing machine learning. *Journal of Cloud Computing*, *12*(1), 127.

[5]. Albayrak Ünal, Ö., Erkayman, B., & Usanmaz, B. (2023). Applications of artificial intelligence in inventory management: A systematic review of the literature. *Archives of computational methods in engineering*, *30*(4), 2605-2625.

[6]. Albert, A. (2025). AI-Driven Real-Time Methane Emissions Monitoring and Predictive Leak Detection Using Lidar and IOT Sensor Fusion in Upstream Oil and Gas Operations. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *1*(01), 2035–2077. https://doi.org/10.63125/yavd2f86

[7]. Alimohammadi, H., & Chen, S. N. (2022). Performance evaluation of outlier detection techniques in production timeseries: A systematic review and meta-analysis. *Expert Systems with Applications*, *191*, 116371.

[8]. Alowais, S. A., Alghamdi, S. S., Alsuhebany, N., Alqahtani, T., Alshaya, A. I., Almohareb, S. N., Aldairem, A., Alrashed, M., Bin Saleh, K., & Badreldin, H. A. (2023). Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC medical education*, *23*(1), 689.

[9]. Amena Begum, S., & Mst Kaniz, F. (2023). Advanced Computational and Biotechnological Approaches to Systemic Family Therapy: Predicting Marital Satisfaction and Emotional Wellbeing in Couples. *Review of Applied Science and Technology*, *2*(04), 228–265. https://doi.org/10.63125/4sy9qa21

[10]. Amena Begum, S., & Mst Kaniz, F. (2024). Integrating Psychometric and Neurocognitive Biomarkers in Computational Models to Predict Cognitive Behavioral Therapy Outcomes in Adolescents with Anxiety and Depression. *International Journal of Scientific Interdisciplinary Research*, *5*(2), 632–677. https://doi.org/10.63125/7t7wmp27

[11]. Andaur, J. M. R., Ruz, G. A., & Goycoolea, M. (2021). Predicting out-of-stock using machine learning: an application in a retail packaged foods manufacturing company. *Electronics*, *10*(22), 2787.

[12]. Apostol, E.-S., Truică, C.-O., Pop, F., & Esposito, C. (2021). Change point enhanced anomaly detection for IoT time series data. *Water*, *13*(12), 1633.

[13]. Arden, N. S., Fisher, A. C., Tyner, K., Yu, L. X., Lee, S. L., & Kopcha, M. (2021). Industry 4.0 for pharmaceutical manufacturing: Preparing for the smart factories of the future. *International Journal of Pharmaceutics*, *602*, 120554.

[14]. Awan, U., Shamim, S., Khan, Z., Zia, N. U., Shariq, S. M., & Khan, M. N. (2021). Big data analytics capability and decision-making: The role of data-driven insight on circular economy performance. *Technological forecasting and social change*, *168*, 120766.

[15]. Bag, S., Wood, L. C., Xu, L., Dhamija, P., & Kayikci, Y. (2020). Big data analytics as an operational excellence approach to enhance sustainable supply chain performance. *Resources, conservation and recycling*, *153*, 104559.
[16]. Banner, M., Alosert, H., Spencer, C., Cheeks, M., Farid, S. S., Thomas, M., & Goldrick, S. (2021). A decade in review: use of data analytics within the biopharmaceutical sector. *Current Opinion in Chemical Engineering*, *34*, 100758.
[17]. Bauer, G. R., Churchill, S. M., Mahendran, M., Walwyn, C., Lizotte, D., & Villa-Rueda, A. A. (2021). Intersectionality in quantitative research: A systematic review of its emergence and applications of theory and methods. *SSM-population health*, *14*, 100798.
[18]. Benedetti, A., Khoo, J., Sharma, S., Facco, P., Barolo, M., & Zomer, S. (2019). Data analytics on raw material properties to accelerate pharmaceutical drug development. *International Journal of Pharmaceutics*, *563*, 122-134.
[19]. Besseling, R., Damen, M., Wijgergangs, J., Hermes, M., Wynia, G., & Gerich, A. (2019). New unique PAT method and instrument for real-time inline size characterization of concentrated, flowing nanosuspensions. *European Journal of Pharmaceutical Sciences*, *133*, 205-213.
[20]. Bo, Y., Chen, Y., Che, Q., Shi, Y., & Zhang, Y. (2022). Multivariate statistical analysis of the spatial variability of hydrochemical evolution during riverbank infiltration. *Water*, *14*(23), 3800.
[21]. Bottani, E., Montanari, R., Volpi, A., & Tebaldi, L. (2023). Statistical process control of assembly lines in manufacturing. *Journal of Industrial Information Integration*, *32*, 100435.
[22]. Brnabic, A., & Hess, L. M. (2021). Systematic literature review of machine learning methods used in the analysis of real-world data for patient-provider decision making. *BMC medical informatics and decision making*, *21*(1), 54.
[23]. Buonanno, G., Morawska, L., & Stabile, L. (2020). Quantitative assessment of the risk of airborne transmission of SARS-CoV-2 infection: prospective and retrospective applications. *Environment international*, *145*, 106112.
[24]. Campbell, J. N., Isenor, A. W., & Ferreira, M. D. (2022). Detection of invalid AIS messages using machine learning techniques. *Procedia Computer Science*, *205*, 229-238.
[25]. Carvajal Soto, J., Tavakolizadeh, F., & Gyulai, D. (2019). An online machine learning framework for early detection of product failures in an Industry 4.0 context. *International Journal of Computer Integrated Manufacturing*, *32*(4-5), 452-465.
[26]. Casian, T., Nagy, B., Kovács, B., Galata, D. L., Hirsch, E., & Farkas, A. (2022). Challenges and opportunities of implementing data fusion in process analytical technology—A review. *Molecules*, *27*(15), 4846.
[27]. Cerquitelli, T., Pagliari, D. J., Calimera, A., Bottaccioli, L., Patti, E., Acquaviva, A., & Poncino, M. (2021). Manufacturing as a data-driven practice: methodologies, technologies, and tools. *Proceedings of the IEEE*, *109*(4), 399-422.
[28]. Chen, J., Wang, S.-L., & Lin, H.-L. (2019). Out-of-stock detection based on deep learning. International Conference on Intelligent Computing,
[29]. Chen, Y., Yang, O., Sampat, C., Bhalode, P., Ramachandran, R., & Ierapetritou, M. (2020). Digital twins in pharmaceutical and biopharmaceutical manufacturing: a literature review. *Processes*, *8*(9), 1088.
[30]. Choi, K., Yi, J., Park, C., & Yoon, S. (2021). Deep learning for anomaly detection in time-series data: Review, analysis, and guidelines. *IEEe Access*, *9*, 120043-120065.
[31]. Curcuruto, M., Conchie, S. M., & Griffin, M. A. (2019). Safety citizenship behavior (SCB) in the workplace: A stable construct? Analysis of psychometric invariance across four European countries. *Accident Analysis & Prevention*, *129*, 190-201.
[32]. Depommier, C., Everard, A., Druart, C., Plovier, H., Van Hul, M., Vieira-Silva, S., Falony, G., Raes, J., Maiter, D., & Delzenne, N. M. (2019). Supplementation with Akkermansia muciniphila in overweight and obese human volunteers: a proof-of-concept exploratory study. *Nature medicine*, *25*(7), 1096-1103.
[33]. Dong, Y., Yang, T., Xing, Y., Du, J., & Meng, Q. (2023). Data-driven modeling methods and techniques for pharmaceutical processes. *Processes*, *11*(7), 2096.
[34]. Dubey, R., Gunasekaran, A., Childe, S. J., Fosso Wamba, S., Roubaud, D., & Foropon, C. (2021). Empirical investigation of data analytics capability and organizational flexibility as complements to supply chain resilience. *International Journal of Production Research*, *59*(1), 110-128.
[35]. Dubey, R., Gunasekaran, A., Childe, S. J., Papadopoulos, T., Luo, Z., Wamba, S. F., & Roubaud, D. (2019). Can big data and predictive analytics improve social and environmental sustainability? *Technological forecasting and social change*, *144*, 534-545.
[36]. Eifert, T., Eisen, K., Maiwald, M., & Herwig, C. (2020). Current and future requirements to industrial analytical infrastructure—part 2: smart sensors. *Analytical and bioanalytical chemistry*, *412*(9), 2037-2045.
[37]. Eisen, K., Eifert, T., Herwig, C., & Maiwald, M. (2020). Current and future requirements to industrial analytical infrastructure—part 1: process analytical laboratories. *Analytical and bioanalytical chemistry*, *412*(9), 2027-2035.
[38]. Fei, Y., Huang, S., Zhang, H., Tong, Y., Wen, D., Xia, X., Wang, H., Luo, Y., & Barceló, D. (2020). Response of soil enzyme activities and bacterial communities to the accumulation of microplastics in an acid cropped soil. *Science of the Total Environment*, *707*, 135634.
[39]. Ferdous Ara, A., & Beatrice Onyinyechi, M. (2023). Long-Term Epidemiologic Trends of STIs PRE- and post-PrEP Introduction: A National Time-Series Analysis. *American Journal of Health and Medical Sciences*, *4*(02), 01–35. https://doi.org/10.63125/mp153d97
[40]. Fox, M. P., MacLehose, R. F., & Lash, T. L. (2021). *Applying quantitative bias analysis to epidemiologic data* (Vol. 10). Springer.
[41]. Gao, Y., Zhang, T., Ma, Y., Xue, F., Gao, Z., Hou, B., & Gong, J. (2021). Application of PAT-based feedback control approaches in pharmaceutical crystallization. *Crystals*, *11*(3), 221.

[42]. Garitano, I., Iturbe, M., Ezpeleta, E., & Zurutuza, U. (2019). Who's there? Evaluating data source integrity and veracity in iiot using multivariate statistical process control. In *Security and privacy trends in the industrial Internet of things* (pp. 181-198). Springer.
[43]. Ghaderpour, E., Pagiatakis, S. D., & Hassan, Q. K. (2021). A survey on change detection and time series analysis with applications. *Applied Sciences*, *11*(13), 6141.
[44]. Ghasemaghaei, M. (2019). Does data analytics use improve firm decision making quality? The role of knowledge sharing and data analytics competency. *Decision Support Systems*, *120*, 14-24.
[45]. Giaconia, C., & Chamas, A. (2023). Innovative out-of-stock prediction system based on data history knowledge deep learning processing. *Computation*, *11*(3), 62.
[46]. Greasley, A., & Edwards, J. S. (2021). Enhancing discrete-event simulation with big data analytics: A review. *Journal of the Operational Research Society*, *72*(2), 247-267.
[47]. Guerra, A., von Stosch, M., & Glassey, J. (2019). Toward biotherapeutic product real-time quality monitoring. *Critical reviews in biotechnology*, *39*(3), 289-305.
[48]. Gül, İ., Helvacıoğlu, E. T., & Saraçlı, S. (2023). Service quality, outpatient satisfaction and loyalty in community pharmacies in Turkey: A structural equation modeling approach. *Exploratory Research in Clinical and Social Pharmacy*, *12*, 100361.
[49]. Hariri, R. H., Fredericks, E. M., & Bowers, K. M. (2019). Uncertainty in big data analytics: survey, opportunities, and challenges. *Journal of Big data*, *6*(1), 44.
[50]. He, Q. P., Wang, J., & Shah, D. (2019). Feature space monitoring for smart manufacturing via statistics pattern analysis. *Computers & Chemical Engineering*, *126*, 321-331.
[51]. He, Y., Huang, Z., Vogt, S., & Sick, B. (2023). Proud: Probabilistic outlier detection solution for time-series analysis of real-world photovoltaic inverters. *Energies*, *17*(1), 64.
[52]. Helgers, H., Schmidt, A., Lohmann, L. J., Vetter, F. L., Juckers, A., Jensch, C., Mouellef, M., Zobel-Roos, S., & Strube, J. (2021). Towards autonomous operation by advanced process control—process analytical technology for continuous biologics antibody manufacturing. *Processes*, *9*(1), 172.
[53]. Helms, J., Tacquard, C., Severac, F., Leonard-Lorant, I., Ohana, M., Delabranche, X., Merdji, H., Clere-Jehl, R., Schenck, M., & Fagot Gandet, F. (2020). High risk of thrombosis in patients with severe SARS-CoV-2 infection: a multicenter prospective cohort study. *Intensive care medicine*, *46*(6), 1089-1098.
[54]. Houssein, E. H., Hosney, M. E., Emam, M. M., Younis, E. M., Ali, A. A., & Mohamed, W. M. (2023). Soft computing techniques for biomedical data analysis: open issues and challenges. *Artificial intelligence review*, *56*(Suppl 2), 2599-2649.
[55]. Ikegwu, A. C., Nweke, H. F., Anikwe, C. V., Alo, U. R., & Okonkwo, O. R. (2022). Big data analytics for data-driven industry: a review of data sources, tools, challenges, solutions, and research directions. *Cluster Computing*, *25*(5), 3343-3387.
[56]. Ishtiaque, A., & Rajib, S. (2025). The Impact of Machine Learning on Cyber Risk Quantification in Financial Services: A Qualitative Evaluation of Threat Scoring Frameworks. *American Journal of Advanced Technology and Engineering Solutions*, *1*(02), 58-94. https://doi.org/10.63125/7aqqac69
[57]. Islam, M. D. Z., & Aditya, D. (2023). Measuring the Security Impact of Zero Trust Access Controls: A Mixed-Methods Study of Identity-Based Policies (Cisco ISE + AD) and Incident Reduction. *American Journal of Data Science and Analytics*, *4*(06), 01-42. https://doi.org/10.63125/8ycz7671
[58]. Istiaq, A., & Nusrat, J. (2022). A Panel Data Econometric Analysis on the Impact of Digital Payment Adoption on Small Business Revenue Growth in Global Business. *American Journal of Interdisciplinary Studies*, *3*(04), 500-536. https://doi.org/10.63125/ehvpjc80
[59]. Javed, M. N., Alam, M. S., Waziri, A., Pottoo, F. H., Yadav, A. K., Hasnain, M. S., & Almalki, F. A. (2019). QbD applications for the development of nanopharmaceutical products. In *Pharmaceutical quality by design* (pp. 229-253). Elsevier.
[60]. Jelsch, M., Roggo, Y., Kleinebudde, P., & Krumme, M. (2021). Model predictive control in pharmaceutical continuous manufacturing: A review from a user's perspective. *European Journal of Pharmaceutics and Biopharmaceutics*, *159*, 137-142.
[61]. Jia, C.-Y., Li, J.-Y., Hao, G.-F., & Yang, G.-F. (2020). A drug-likeness toolbox facilitates ADMET study in drug discovery. *Drug discovery today*, *25*(1), 248-258.
[62]. Jin, Y., Du, W., Liu, X., & Wu, Y. (2022). Application of near infrared spectroscopy and real time release testing combined with statistical process control charts for on-line quality control of industrial concentrating process of traditional Chinese medicine "Jinyinhua". *Infrared Physics & Technology*, *123*, 104135.
[63]. Jin, Y., O'Connor, D., Ok, Y. S., Tsang, D. C., Liu, A., & Hou, D. (2019). Assessment of sources of heavy metals in soil and dust at children's playgrounds in Beijing using GIS and multivariate statistical analysis. *Environment international*, *124*, 320-328.
[64]. Kazi Rakib Hasan, S. (2025). Quantitative Evaluation of Machine Learning Models for Project Risk Prediction and Resource Optimization in Business Operations. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *1*(01), 2119–2159. https://doi.org/10.63125/01bg6n62
[65]. Kenett, R. S., Zacks, S., & Gedeck, P. (2023). Multivariate Statistical Process Control. In *Industrial Statistics: A Computer-Based Approach with Python* (pp. 113-140). Springer.
[66]. Kim, E. J., Kim, J. H., Kim, M.-S., Jeong, S. H., & Choi, D. H. (2021). Process analytical technology tools for monitoring pharmaceutical unit operations: a control strategy for continuous process verification. *Pharmaceutics*, *13*(6), 919.

[67]. Koshechkin, K. A., Lebedev, G. S., Fartushnyi, E. N., & Orlov, Y. L. (2022). Holistic approach for artificial intelligence implementation in pharmaceutical products lifecycle: a meta-analysis. *Applied Sciences*, *12*(16), 8373.

[68]. Lee, S.-H., Kim, J.-K., Jee, J.-P., Jang, D.-J., Park, Y.-J., & Kim, J.-E. (2022). Quality by Design (QbD) application for the pharmaceutical development process. *Journal of Pharmaceutical Investigation*, *52*(6), 649-682.

[69]. Lee, S., Kwak, M., Tsui, K.-L., & Kim, S. B. (2019). Process monitoring using variational autoencoder for high-dimensional nonlinear processes. *Engineering Applications of Artificial Intelligence*, *83*, 13-27.

[70]. Li, K., Fang, Y., Li, W., Pan, C., Qin, P., Zhong, Y., Liu, X., Huang, M., Liao, Y., & Li, S. (2020). CT image visual quantitative evaluation and clinical classification of coronavirus disease (COVID-19). *European radiology*, *30*(8), 4407-4416.

[71]. Li, Z., Tian, L., & Yan, X. (2022). An ensemble framework based on multivariate statistical analysis for process monitoring. *Expert Systems with Applications*, *205*, 117732.

[72]. Liu, X., Shin, H., & Burns, A. C. (2021). Examining the impact of luxury brand's social media marketing on customer engagement: Using big data analytics and natural language processing. *Journal of business research*, *125*, 815-826.

[73]. Mahfuj Ahmed, R., & Md. Hasan Or, R. (2021). Fraud-Detection Algorithms for Identifying Anomalous Transactions in Retail Banking Networks. *American Journal of Data Science and Analytics*, *2*(12), 01-40. https://doi.org/10.63125/23m31748

[74]. Mahfuj Ahmed, R., & Md. Mehedi, H. (2023). Digital Technologies and IoT: Reshaping Financial Risk and Investment in Global Supply Chains. *Journal of Sustainable Development and Policy*, *2*(04), 297-345. https://doi.org/10.63125/nbv6ka16

[75]. Malheiro, V., Duarte, J., Veiga, F., & Mascarenhas-Melo, F. (2023). Exploiting pharma 4.0 technologies in the non-biological complex drugs manufacturing: innovations and implications. *Pharmaceutics*, *15*(11), 2545.

[76]. Masood, A., & Hashmi, A. (2019). AIOps: predictive analytics & machine learning in operations. In *Cognitive computing recipes: artificial intelligence solutions using microsoft cognitive services and tensorflow* (pp. 359-382). Springer.

[77]. Md Khaled, H. (2026). Artificial Intelligence Based Predictive Analytics for SKU Performance and Revenue Optimization in Competitive Markets. *American Journal of Advanced Technology and Engineering Solutions*, *6*(01), 297-331. https://doi.org/10.63125/cmyhzv81

[78]. Md Khaled, H., & Hisham, M. (2022). Intelligent Decision-Support Systems for Cross-Functional Workflow Optimization in Data-Driven Organizations. *Journal of Sustainable Development and Policy*, *1*(02), 168-207. https://doi.org/10.63125/dsfg3k24

[79]. Md Khaled, H., & Md. Morshedul, I. (2024). AI-Enabled Enterprise Scorecards for Reducing Operational Errors and Enhancing Supply Chain Consistency. *American Journal of Scholarly Research and Innovation*, *3*(01), 117–152. https://doi.org/10.63125/fa50dw13

[80]. Md Mehedi, H., & Md, F. (2022). Advanced Computing–Enabled Secure Financial Information Systems for Real-Time Fraud Detection in U.S. Digital Payments: A Quantitative Analysis. *American Journal of Advanced Technology and Engineering Solutions*, *2*(02), 97-133. https://doi.org/10.63125/9mv2qd37

[81]. Md. Ashfaq, S., & Ashraful, I. (2025). Quantitative Analysis of Machine Learning Models For Defect Prediction in Metal Additive Manufacturing. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *1*(01), 1810–1847. https://doi.org/10.63125/3fkkwg05

[82]. Md. Hasan Or, R., Tanjina Binte, S., & Rajib, S. (2023). Performance Analytics Frameworks for Digital Marketing and Service Enterprises: An empirical Study. *American Journal of Data Science and Analytics*, *4*(03), 01-35. https://doi.org/10.63125/aq7y1792

[83]. Md. Mainuddin, F., & Palash Chandra, D. (2022). Fabrication-Driven Structural Optimization Techniques for Cost-Efficient Steel Construction Using CNC-Based Design Workflows. *American Journal of Interdisciplinary Studies*, *3*(04), 464-499. https://doi.org/10.63125/n08g1x15

[84]. Md. Mainuddin, F., & Palash Chandra, D. (2023). Advanced Computing–Based Modeling of Steel Connection Behavior and Stability Performance using ETABS And STAAD Pro. *American Journal of Advanced Technology and Engineering Solutions*, *3*(04), 42-86. https://doi.org/10.63125/xfkzrg56

[85]. Md. Mehedi, H., & Khairum Nahar, P. (2023). A Systematic Review of Secure Health Data Information Systems for Pandemic Preparedness and Economic Continuity in the United States. *Review of Applied Science and Technology*, *2*(01), 227–258. https://doi.org/10.63125/77h2m531

[86]. Md. Mehedi, H., & Khairum Nahar, P. (2024). Advanced Computing and AI-Driven National Information Systems for Predictive Disaster Risk Management and Economic Loss Mitigation. *American Journal of Scholarly Research and Innovation*, *3*(02), 296–336. https://doi.org/10.63125/4sbz5j45

[87]. Md. Morshedul, I., Rukaiya Khatun, M., & Khairum Nahar, P. (2022). Machine Learning–Driven Forecasting Pipelines for Financial Volatility Detection in Integrated Enterprise ERP Environments. *American Journal of Advanced Technology and Engineering Solutions*, *2*(02), 134-173. https://doi.org/10.63125/y42nk811

[88]. Md. Nazmul, H., & Amena Begum, S. (2022). AI-Based Psychodiagnostics' Models to Support Early Intervention and Reduce Suicide Risk in Adolescents and Youth: Development and Clinical Validation. *American Journal of Data Science and Analytics*, *3*(06), 40-79. https://doi.org/10.63125/vb5f7e98

[89]. Md. Shahinur, I., & Md. Sultan, M. (2022). Digital-Twin-Based Quantitative Frameworks for Modeling, Monitoring, and Optimization of Electrical Power Infrastructure. *American Journal of Interdisciplinary Studies*, *3*(04), 365-393. https://doi.org/10.63125/dvmj1y93

[90]. Md. Towhidul, I., & Uddin, M. D. S. (2024). Simulation-Based Forecasting and Inventory Control Models For Consumer Goods Networks: A Quantitative Study Using Monte Carlo Simulation and Time-Series Methods. *Review of Applied Science and Technology*, *3*(04), 165–197. https://doi.org/10.63125/a3047d06

[91]. Mertler, C. A., Vannatta, R. A., & LaVenia, K. N. (2021). *Advanced and multivariate statistical methods: Practical application and interpretation*. Routledge.

[92]. Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of business research*, *98*, 261-276.

[93]. Mo, R., Pei, Y., Venkatarayalu, N. V., Joseph, P. N., Premkumar, A. B., Sun, S., & Foo, S. K. K. (2022). Unsupervised TCN-AE-based outlier detection for time series with seasonality and trend for cellular networks. *IEEE Transactions on Wireless Communications*, *22*(5), 3114-3127.

[94]. Mohamed, A., Najafabadi, M. K., Wah, Y. B., Zaman, E. A. K., & Maskat, R. (2020). The state of the art and taxonomy of big data analytics: view from new big data framework. *Artificial intelligence review*, *53*(2), 989-1037.

[95]. Mohammad Robel, M. (2025). Advanced Computing Frameworks for Distributed Training, Deployment, and Monitoring of Artificial Intelligence and Machine Learning Models. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *1*(01), 1922–1957. https://doi.org/10.63125/rxb2cb66

[96]. Mohammad Robel, M., & Md. Morshedul, I. (2021). Foundational Approaches to Secure Data Collection and Processing in Networked and Distributed Computing Environments. *International Journal of Business and Economics Insights*, *1*(4), 32–69. https://doi.org/10.63125/thrtkw71

[97]. Mohammad Robel, M., & Md. Morshedul, I. (2024). Data Preprocessing and Feature Engineering Strategies for Large-Scale Predictive Modeling Applications. *Review of Applied Science and Technology*, *3*(01), 263–302. https://doi.org/10.63125/tqqqed47

[98]. Mostafa, K. (2023). An Empirical Evaluation of Machine Learning Techniques for Financial Fraud Detection in Transaction-Level Data. *American Journal of Interdisciplinary Studies*, *4*(04), 210-249. https://doi.org/10.63125/60amyk26

[99]. Mukherjee, P. K. (2019). *Quality control and evaluation of herbal drugs: Evaluating natural products and traditional medicine*. Elsevier.

[100]. Murad, M. D. H. R. (2025). Machine Learning–Based Consumer Behavior Prediction Models for E-Commerce Platforms: Enhancing Digital Financial Inclusion and Market Accessibility. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *1*(01), 2078–2118. https://doi.org/10.63125/pnz32s94

[101]. Nagy, B., Galata, D. L., Farkas, A., & Nagy, Z. K. (2022). Application of artificial neural networks in the process analytical technology of pharmaceutical manufacturing—a review. *The AAPS journal*, *24*(4), 74.

[102]. Najjar, R. (2023). Redefining radiology: a review of artificial intelligence integration in medical imaging. *Diagnostics*, *13*(17), 2760.

[103]. Namoano, B., Starr, A., Emmanouilidis, C., & Cristobal, R. C. (2019). Online change detection techniques in time series: An overview. 2019 IEEE international conference on prognostics and health management (ICPHM),

[104]. Ngo, P. T. T., Panahi, M., Khosravi, K., Ghorbanzadeh, O., Kariminejad, N., Cerda, A., & Lee, S. (2021). Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of Iran. *Geoscience Frontiers*, *12*(2), 505-519.

[105]. Nguyen, A., Lamouri, S., Pellerin, R., Tamayo, S., & Lekens, B. (2022). Data analytics in pharmaceutical supply chains: state of the art, opportunities, and challenges. *International Journal of Production Research*, *60*(22), 6888-6907.

[106]. Nizam, H., Zafar, S., Lv, Z., Wang, F., & Hu, X. (2022). Real-time deep anomaly detection framework for multivariate time-series data in industrial IoT. *IEEE Sensors Journal*, *22*(23), 22836-22849.

[107]. Oliveira, M. D., Mataloto, I., & Kanavos, P. (2019). Multi-criteria decision analysis for health technology assessment: addressing methodological challenges to improve the state of the art: MD Oliveira et al. *The European Journal of Health Economics*, *20*(6), 891-918.

[108]. Omar, M. A., & Inaba, K. (2020). Does financial inclusion reduce poverty and income inequality in developing countries? A panel data analysis. *Journal of economic structures*, *9*(1), 37.

[109]. Palash Chandra, D. (2023). Machine Learning–Driven Optimization of Water Distribution Networks: Demand Forecasting, and Energy Efficiency Analysis. *Journal of Sustainable Development and Policy*, *2*(04), 257-296. https://doi.org/10.63125/jdxq0819

[110]. Pauli, V., Roggo, Y., Pellegatti, L., Trung, N. Q. N., Elbaz, F., Ensslin, S., Kleinebudde, P., & Krumme, M. (2019). Process analytical technology for continuous manufacturing tableting processing: A case study. *Journal of pharmaceutical and biomedical analysis*, *162*, 101-111.

[111]. Pedro, F., Veiga, F., & Mascarenhas-Melo, F. (2023). Impact of GAMP 5, data integrity and QbD on quality assurance in the pharmaceutical industry: How obvious is it? *Drug discovery today*, *28*(11), 103759.

[112]. Pérez-Benítez, B. E., Tercero-Gómez, V. G., & Khakifirooz, M. (2023). A review on statistical process control in healthcare: data-driven monitoring schemes. *IEEe Access*, *11*, 56248-56272.

[113]. Peterson, J. J., Kramer, T. T., Hofer, J. D., & Atkins, G. (2019). Opportunities and challenges for statisticians in advanced pharmaceutical manufacturing. *Statistics in Biopharmaceutical Research*.

[114]. Portela, R. M., Varsakelis, C., Richelle, A., Giannelos, N., Pence, J., Dessoy, S., & von Stosch, M. (2020). When is an in silico representation a digital twin? A biopharmaceutical industry approach to the digital twin concept. In *Digital Twins: Tools and Concepts for Smart Biomanufacturing* (pp. 35-55). Springer.

[115]. Pu, Y.-Y., O'Donnell, C., Tobin, J. T., & O'Shea, N. (2020). Review of near-infrared spectroscopy as a process analytical technology for real-time product monitoring in dairy processing. *International Dairy Journal*, *103*, 104623.

[116]. Qin, S. J., & Chiang, L. H. (2019). Advances and opportunities in machine learning for process data analytics. *Computers & Chemical Engineering*, *126*, 465-473.
[117]. Qiu, P. (2020). Big data? Statistical process control can help! *The American Statistician*, *74*(4), 329-344.
[118]. Ramos, M., Ascencio, J., Hinojosa, M. V., Vera, F., Ruiz, O., Jimenez-Feijoó, M. I., & Galindo, P. (2021). Multivariate statistical process control methods for batch production: A review focused on applications. *Production & manufacturing research*, *9*(1), 33-55.
[119]. Rathore, A. S., Nikita, S., Thakur, G., & Deore, N. (2021). Challenges in process control for continuous processing for production of monoclonal antibody products. *Current Opinion in Chemical Engineering*, *31*, 100671.
[120]. Rauvola, R. S., Vega, D. M., & Lavigne, K. N. (2019). Compassion fatigue, secondary traumatic stress, and vicarious traumatization: A qualitative review and research agenda. *Occupational health science*, *3*(3), 297-336.
[121]. Rawal, M., Singh, A., & Amiji, M. M. (2019). Quality-by-Design Concepts to Improve Nanotechnology-Based Drug Development: Rawal, Singh and Amiji. *Pharmaceutical research*, *36*(11), 153.
[122]. Rodriguez-Ayllon, M., Cadenas-Sánchez, C., Estévez-López, F., Muñoz, N. E., Mora-Gonzalez, J., Migueles, J. H., Molina-García, P., Henriksson, H., Mena-Molina, A., & Martínez-Vizcaíno, V. (2019). Role of physical activity and sedentary behavior in the mental health of preschoolers, children and adolescents: a systematic review and meta-analysis. *Sports medicine*, *49*(9), 1383-1410.
[123]. Roggo, Y., Jelsch, M., Heger, P., Ensslin, S., & Krumme, M. (2020). Deep learning for continuous manufacturing of pharmaceutical solid dosage form. *European Journal of Pharmaceutics and Biopharmaceutics*, *153*, 95-105.
[124]. Roggo, Y., Pauli, V., Jelsch, M., Pellegatti, L., Elbaz, F., Ensslin, S., Kleinebudde, P., & Krumme, M. (2020). Continuous manufacturing process monitoring of pharmaceutical solid dosage form: A case study. *Journal of pharmaceutical and biomedical analysis*, *179*, 112971.
[125]. Rolinger, L., Rüdt, M., & Hubbuch, J. (2020). A critical review of recent trends, and a future perspective of optical spectroscopy as PAT in biopharmaceutical downstream processing. *Analytical and bioanalytical chemistry*, *412*(9), 2047-2064.
[126]. Rukaiya Khatun, M., & Zakia, A. (2023). Quantitative Assessment of Data Privacy and Access Control Effectiveness in SAP/ERP Analytics Systems. *Review of Applied Science and Technology*, *2*(01), 259–300. https://doi.org/10.63125/vb03b363
[127]. Salari, N., Darvishi, N., Heydari, M., Bokaee, S., Darvishi, F., & Mohammadi, M. (2022). Global prevalence of cleft palate, cleft lip and cleft palate and lip: A comprehensive systematic review and meta-analysis. *Journal of stomatology, oral and maxillofacial surgery*, *123*(2), 110-120.
[128]. Sarker, I. H. (2021). Data science and analytics: an overview from data-driven smart computing, decision-making and applications perspective. *SN Computer Science*, *2*(5), 377.
[129]. Scheiman, J., Luber, J. M., Chavkin, T. A., MacDonald, T., Tung, A., Pham, L.-D., Wibowo, M. C., Wurth, R. C., Punthambaker, S., & Tierney, B. T. (2019). Meta-omics analysis of elite athletes identifies a performance-enhancing microbe that functions via lactate metabolism. *Nature medicine*, *25*(7), 1104-1109.
[130]. Schmidt, A., Helgers, H., Vetter, F. L., Juckers, A., & Strube, J. (2021). Digital twin of mRNA-based SARS-COVID-19 vaccine manufacturing towards autonomous operation for improvements in speed, scale, robustness, flexibility and real-time release testing. *Processes*, *9*(5), 748.
[131]. Separovic, L., Simabukuro, R. S., Couto, A. R., Bertanha, M. L. G., Dias, F. R., Sano, A. Y., Caffaro, A. M., & Lourenco, F. R. (2023). Measurement uncertainty and conformity assessment applied to drug and medicine analyses–a review. *Critical Reviews in Analytical Chemistry*, *53*(1), 123-138.
[132]. Shakil, A., Khalighi, M. A., Pudlo, P., Leclerc, C., Laplace, D., Hamon, F., & Boudonne, A. (2023). Outlier detection in non-stationary time series applied to sewer network monitoring. *Internet of Things*, *21*, 100654.
[133]. Shang, C., & You, F. (2019). Data analytics and machine learning for smart process manufacturing: Recent advances and perspectives in the big data era. *Engineering*, *5*(6), 1010-1016.
[134]. Shwartz-Ziv, R., & Armon, A. (2022). Tabular data: Deep learning is not all you need. *Information fusion*, *81*, 84-90.
[135]. Silge, A., Weber, K., Cialla-May, D., Müller-Bötticher, L., Fischer, D., & Popp, J. (2022). Trends in pharmaceutical analysis and quality control by modern Raman spectroscopic techniques. *TrAC Trends in Analytical Chemistry*, *153*, 116623.
[136]. Silva, A. F., Vercruysse, J., Vervaet, C., Remon, J. P., Lopes, J. A., De Beer, T., & Sarraguça, M. C. (2019). In-depth evaluation of data collected during a continuous pharmaceutical manufacturing process: a multivariate statistical process monitoring approach. *Journal of pharmaceutical sciences*, *108*(1), 439-450.
[137]. Song, Y., Wang, J., Ge, Y., & Xu, C. (2020). An optimal parameters-based geographical detector model enhances geographic characteristics of explanatory variables for spatial heterogeneity analysis: Cases with different types of spatial data. *GIScience & remote sensing*, *57*(5), 593-610.
[138]. Sousa, A., Serra, J., Estevens, C., Costa, R., & Ribeiro, A. (2023). A quality by design approach in oral extended release drug delivery systems: where we are and where we are going? *Journal of Pharmaceutical Investigation*, *53*(2), 269-306.
[139]. Steinwandter, V., Borchert, D., & Herwig, C. (2019). Data science tools and applications on the way to Pharma 4.0. *Drug discovery today*, *24*(9), 1795-1805.
[140]. Tanjina Binte, S., & Md. Hasan Or, R. (2022). Advanced Computing, IT Strategy, and Network-Optimized Frameworks for Retail Business Intelligence. *American Journal of Interdisciplinary Studies*, *3*(04), 429-463. https://doi.org/10.63125/dgyg3762

[141]. Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., Li, B., Madabhushi, A., Shah, P., & Spitzer, M. (2019). Applications of machine learning in drug discovery and development. *Nature reviews Drug discovery*, *18*(6), 463-477.

[142]. van de Wiel, L., van Es, D. M., & Feelders, A. (2020). Real-time outlier detection in time series data of water sensors. International Workshop on Advanced Analytics and Learning on Temporal Data,

[143]. Vanli, O. A., & Castillo, E. D. (2021). Statistical process control in manufacturing. In *Encyclopedia of systems and control* (pp. 2142-2150). Springer.

[144]. Volta e Sousa, L., Gonçalves, R., Menezes, J. C., & Ramos, A. (2021). Analytical method lifecycle management in pharmaceutical industry: a review. *Aaps Pharmscitech*, *22*(3), 128.

[145]. Wang, Z., Wang, N., Su, X., & Ge, S. (2020). An empirical study on business analytics affordances enhancing the management of cloud computing data security. *International Journal of Information Management*, *50*, 387-394.

[146]. Wirbel, J., Pyl, P. T., Kartal, E., Zych, K., Kashani, A., Milanese, A., Fleck, J. S., Voigt, A. Y., Palleja, A., & Ponnudurai, R. (2019). Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. *Nature medicine*, *25*(4), 679-689.

[147]. Xiao, Y., He, C., Chen, Y., Ho, C.-T., Wu, X., Huang, Y., Gao, Y., Hou, A., Li, Z., & Wang, Y. (2022). UPLC–QQQ–MS/MS-based widely targeted metabolomic analysis reveals the effect of solid-state fermentation with Eurotium cristatum on the dynamic changes in the metabolite profile of dark tea. *Food Chemistry*, *378*, 131999.

[148]. Yan, J., Wang, L., Song, W., Chen, Y., Chen, X., & Deng, Z. (2019). A time-series classification approach based on change detection for rapid land cover mapping. *ISPRS Journal of Photogrammetry and Remote Sensing*, *158*, 249-262.

[149]. Yu, J., Ye, L., Zhou, L., Yang, Z., Shen, F., & Song, Z. (2021). Dynamic process monitoring based on variational Bayesian canonical variate analysis. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *52*(4), 2412-2422.

[150]. Yuan, C.-C., Chung, W.-H., Cai, C., & Sung, S.-T. (2020). Application of statistical process control on port state control. *Journal of Marine Science and Engineering*, *8*(10), 746.

[151]. Zakia, A., & Rukaiya Khatun, M. (2024). Quantitative Assessment of CRM-Based Business Intelligence on Customer Satisfaction and Retention: Evidence from Multi-Channel Service Operations. *Journal of Sustainable Development and Policy*, *3*(02), 01-42. https://doi.org/10.63125/hjd22x72

[152]. Zhou, J., Gandomi, A. H., Chen, F., & Holzinger, A. (2021). Evaluating the quality of machine learning explanations: A survey on methods and metrics. *Electronics*, *10*(5), 593.

[153]. Zobel-Roos, S., Schmidt, A., Mestmäcker, F., Mouellef, M., Huter, M., Uhlenbrock, L., Kornecki, M., Lohmann, L., Ditz, R., & Strube, J. (2019). Accelerating biologics manufacturing by modeling or: is approval under the QbD and PAT approaches demanded by authorities acceptable without a digital-twin? *Processes*, *7*(2), 94.