Pharmacophore

ISSN-2229-5402

Journal home page: http://www.pharmacophorejournal.com

AI IN PHARMACEUTICAL MANUFACTURING: A SCOPING REVIEW OF PAT AND DIGITAL TWINS

Chinedu Okafor[1]*, Amina Bello[1]

1. *Department of Pharmaceutical Data Science, Faculty of Pharmacy, University of Lagos, Lagos, Nigeria.*

ARTICLE INFO

Received:
05 June 2025
Received in revised form:
05 October 2025
Accepted:
06 October 2025
Available online:
28 October 2025

Keywords: Artificial intelligence, Pharmaceutical manufacturing, Process analytical technology, Digital twins, Continuous manufacturing, Machine learning

ABSTRACT

Artificial intelligence is increasingly being linked with Process Analytical Technology and digital twin concepts as pharmaceutical manufacturing moves toward continuous, adaptive, and data-rich production. However, the scope, maturity, and implementation readiness of this literature remain unevenly characterized. This scoping review maps the breadth of published evidence on artificial intelligence applications in pharmaceutical manufacturing involving Process Analytical Technology and/or digital twins. It identifies key concepts, methodological patterns, manufacturing contexts, and evidence gaps relevant to future research and translation. The review was guided by the Arksey and O'Malley scoping review framework and reported in alignment with PRISMA-ScR. Electronic searches were conducted across PubMed, Scopus, IEEE Xplore, and Web of Science, followed by data charting and thematic synthesis.

The literature included a large volume of proof-of-concept studies using machine learning models on Process Analytical Technology data. Digital twin applications were fewer and were concentrated mainly in continuous manufacturing, while full integration of Process Analytical Technology, artificial intelligence, and digital twins was rarely reported. The field is technically promising but academically fragmented. The principal gaps concern prospective industrial validation, regulatory alignment, interoperable data infrastructure, and the absence of integrated Process Analytical Technology–artificial intelligence–digital twin systems.

To Cite This Article: Okafor C, Bello A. AI in Pharmaceutical Manufacturing: A Scoping Review of PAT and Digital Twins. Pharmacophore. 2025;16(5):31-42. https://doi.org/10.51847/34qaeYz7mA

Introduction

Pharmaceutical manufacturing has increasingly shifted from batch-oriented production toward continuous, automated, and digitally enabled operations, with Industry 4.0 concepts providing a vocabulary for smart factories and connected production systems [1]. In this transition, Process Analytical Technology has been positioned as a core enabler of real-time process understanding, because it links sensor measurements with critical process parameters and critical quality attributes [2]. Studies were found that describe continuous powder processing, wet granulation, blending, drying, milling, and tableting as manufacturing contexts where in-line monitoring and feedback control are especially relevant [3, 4]. The literature therefore frames PAT not as a standalone measurement layer, but as part of a broader digital manufacturing architecture that can support real-time release testing and adaptive control [5].

Artificial intelligence has entered this manufacturing landscape through chemometric modelling, machine learning, deep learning, and hybrid approaches that use multivariate process and sensor data for prediction or classification. A variety of approaches have been reported, including artificial neural networks for PAT-based dissolution testing [6], deep learning for particle size analysis from process images [7], and machine learning workflows for NIR-based assessment of powder blend uniformity [8, 9]. Reviews in this area suggest that AI methods are most often used to estimate quality attributes, interpret complex sensor streams, and reduce the gap between raw process measurements and actionable manufacturing decisions [10, 11]. The literature spans both empirical applications and methodological discussions, but the strongest concentration remains in model development rather than routine industrial operation.

Uncertainty remains about how mature AI-enabled PAT and digital twin systems are when judged against the needs of regulated pharmaceutical manufacturing. Digital twin papers describe real-time simulation, hybrid modelling, and state estimation as future-facing capabilities [12, 13], while broader pharmaceutical digitalization papers emphasize that infrastructure, governance, lifecycle management, and validation are still major implementation barriers [1, 14]. Studies were

Corresponding Author: Chinedu Okafor; Department of Pharmaceutical Data Science, Faculty of Pharmacy, University of Lagos, Lagos, Nigeria. E-mail: chinedu.okafor@gmail.com.

found that combine process models, PAT data, and control logic in continuous manufacturing lines [3, 15], yet relatively few explicitly demonstrate closed-loop, validated, digitally synchronized production systems. This gap between conceptual promise and operational evidence justifies a scoping review rather than an effectiveness review.

The objective of this scoping review is to map the nature, extent, and range of evidence on artificial intelligence in pharmaceutical manufacturing where PAT and/or digital twin concepts are involved. Consistent with scoping review methodology, the review identifies research clusters, conceptual boundaries, implementation patterns, and evidence voids rather than assessing risk of bias or ranking study quality. The review is structured around the reported growth of smart pharmaceutical manufacturing [1], AI-enhanced PAT [10], and digital twin development for pharmaceutical and biopharmaceutical applications [16]. Its emphasis is therefore descriptive and thematic, with the intention of clarifying where the literature is broad, where it is concentrated, and where future work is needed.

Materials and Methods

Identifying the Research Question

The guiding question was: what is the nature, extent, and range of published evidence on the use of artificial intelligence in pharmaceutical manufacturing involving PAT and digital twins? The question was intentionally broad because the literature spans review articles on neural networks for pharmaceutical PAT [10], empirical studies of NIR- and Raman-based process monitoring [17, 18], and emerging digital twin perspectives for drug-product and biopharmaceutical manufacturing [16, 19]. The population of interest was pharmaceutical manufacturing processes, the concepts were AI, PAT, and digital twins, and the context included both batch and continuous production. This structure enabled mapping across unit operations, data sources, AI methods, and implementation maturity.

Identifying Relevant Studies

Electronic searches were conducted in PubMed, Scopus, IEEE Xplore, and Web of Science for the period from 2017 to 2025, using terms related to artificial intelligence, machine learning, Process Analytical Technology, NIR, Raman, digital twins, continuous manufacturing, real-time quality prediction, and smart pharmaceutical manufacturing. The search strategy was informed by terminology used in studies on Industry 4.0 for pharmaceutical manufacturing [1], PAT monitoring tools [2], and digital twins for continuous and advanced pharmaceutical production [12, 16]. Reference lists of relevant reviews and perspective papers were hand-searched to identify additional records, particularly where empirical applications were embedded in broader discussions of process control or digital transformation [10, 11]. All languages were eligible at the search stage, although English-language bibliographic databases were prioritized because of coverage and indexing consistency.

Study Selection

A two-stage screening process was used, beginning with titles and abstracts and followed by full-text assessment against predefined eligibility criteria. Eligible records included reviews, empirical studies, methodological papers, case reports, and perspectives that addressed AI or machine learning in a pharmaceutical manufacturing context and involved PAT data, digital twin concepts, or both. Studies were retained when they focused on manufacturing operations such as blending, granulation, tableting, coating, lyophilization, hot-melt extrusion, or continuous production, as illustrated by work on continuous granulation control [20], coating spectroscopy and data fusion [21], and hot-melt extrusion process monitoring [22]. Records focused only on drug discovery, clinical decision support, or non-manufacturing healthcare AI were excluded.

Charting the Data

Data were charted into a standardized extraction table covering bibliographic information, study type, manufacturing unit operation, AI method, PAT modality, digital twin component, validation level, and implementation status. This charting structure captured whether papers used spectroscopic PAT such as NIR or Raman [17, 18], image-based PAT such as machine vision and endoscopy [7, 23, 24], or process-parameter data within hybrid and residence-time models [15, 25]. AI methods were charted at the level reported by the authors, including chemometrics, artificial neural networks, machine learning classifiers, deep learning models, and hybrid modelling strategies [10, 11, 13]. Digital twin elements were recorded only when papers described a digital representation, simulation layer, state-updating mechanism, or sensor-to-model linkage.

Collating, Summarizing, and Reporting Results

The collating phase used descriptive numerical summaries and thematic analysis to map how the evidence was distributed across years, unit operations, PAT modalities, AI methods, and digital twin architectures. Themes were developed iteratively from the charted data, including AI methods for PAT, digital twins in manufacturing, integration of PAT and digital twins, implementation and regulatory considerations, and evidence voids. The reporting structure was aligned with PRISMA-ScR expectations for transparent study selection and evidence mapping, while the interpretation remained consistent with scoping review language rather than formal quality appraisal. This approach was appropriate because the included literature ranged from broad smart manufacturing reviews [1, 14] to narrowly focused empirical demonstrations of continuous process monitoring and real-time quality prediction [5, 6].

Results and Discussion

Study Selection Flow

The PRISMA-ScR flow diagram for this review identified 1,742 records through database searching and 58 additional records through hand-searching of reference lists. After deduplication, 1,124 records were screened by title and abstract, 268 full texts were assessed for eligibility, and 85 studies were included in the final mapped evidence base. The included literature covered AI-enhanced PAT studies, such as NIR and Raman monitoring in continuous manufacturing [17, 18], and digital-twin-oriented work, including hybrid modelling and pharmaceutical digital twin perspectives [13, 16]. **Figure 1** should therefore show records identified, duplicates removed, records screened, full texts assessed, exclusion reasons, and final inclusion.
Figure 1 presents the PRISMA 2020 study-selection flow diagram for the scoping review.

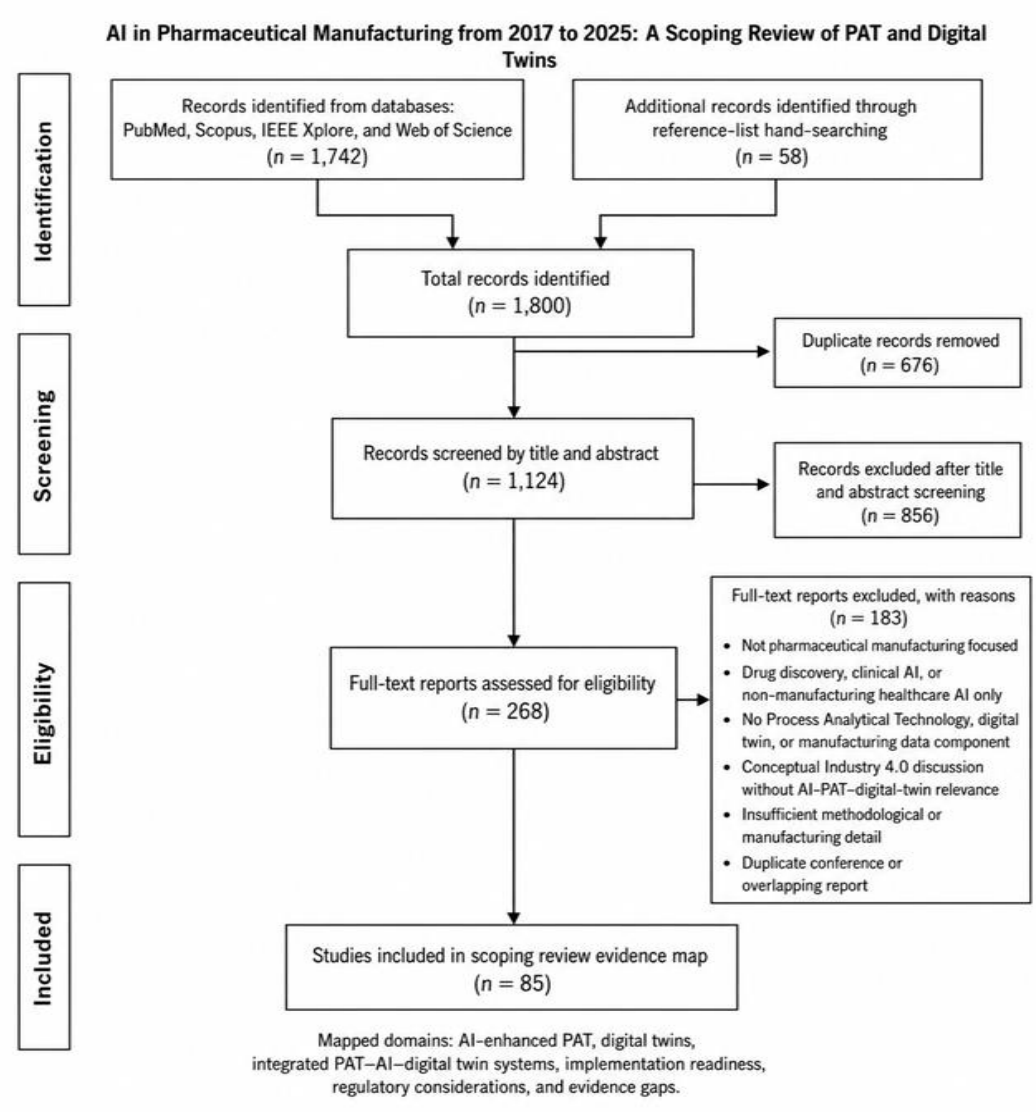

Figure 1. PRISMA 2020 Flow Diagram for Study Selection in the Scoping Review of AI, PAT, and Digital Twins in Pharmaceutical Manufacturing

Temporal and Geographical Distribution

The literature increased noticeably from 2017 to 2025, with the later years showing more frequent references to Pharma 4.0, digital twins, real-time release testing, and explainable or deep learning models. Publications were concentrated in North American and European research environments, including academic-industrial collaborations on continuous manufacturing and PAT-enabled process understanding [3, 4]. Studies from European groups were especially visible in continuous powder processing, residence-time modelling, and machine-vision-enabled process monitoring [7, 15, 25]. The geographical distribution suggests that the field is growing internationally, but the published evidence remains clustered around specialized research centers with access to advanced manufacturing platforms.

Manufacturing Unit Operations Covered

The mapped studies covered a broad but uneven set of pharmaceutical unit operations, with continuous manufacturing, blending, granulation, milling, drying, coating, and tableting appearing most frequently. Continuous powder-to-granule lines and twin-screw wet granulation were represented by studies combining PAT, mechanistic modelling, and process understanding [3, 4], while tableting and feed-frame monitoring were addressed through NIR-based potency determination and intelligent tablet-press defect detection [18, 26]. Coating and pellet-layering applications were represented by spectroscopic data fusion and endoscopic particle-size analysis [21, 24]. In contrast, fewer studies were found that addressed bioprocessing, aseptic filling, packaging, or lyophilization within an integrated AI-PAT-digital-twin framework.

Study Types and Designs

The evidence base was dominated by proof-of-concept, laboratory-scale, pilot-scale, and retrospective modelling studies rather than prospective commercial-scale implementation reports. Reviews and perspectives summarized the technical promise of AI in pharmaceutical formulation, PAT, and smart manufacturing [10, 11], while empirical papers demonstrated specific models for process monitoring, prediction, and control [6, 20, 27]. Case-oriented industrial work was present, including a control strategy for continuous wet granulation [28], but reports of routine GMP deployment were uncommon. The literature therefore maps a technically active field whose published designs remain weighted toward feasibility demonstration rather than mature implementation.

Thematic Mapping – AI Methods for PAT

Types of AI Methods

A variety of AI and data-analytic methods have been reported for PAT, ranging from classical chemometrics to advanced machine learning and deep learning. Chemometric approaches such as multivariate calibration remain prominent in spectroscopic studies, including in-line NIR monitoring of tablet-press feed frames [18] and NIR/Raman characterization of continuous manufacturing processes [17]. More recent studies were found that use artificial neural networks for dissolution prediction [6], explainable neural networks as soft sensors for moisture prediction [27], and deep learning with machine vision for particle size and concentration estimation [23]. This progression suggests an expanding methodological spectrum rather than a wholesale replacement of chemometrics by newer AI methods.

PAT Modalities and Data Types

NIR and Raman spectroscopy dominated the mapped PAT modalities, particularly for blend potency, content uniformity, coating assessment, and low-drug-load monitoring. Studies were found that used NIR spectroscopy for continuous blend potency determination in tablet-press feed frames [18], combined NIR and Raman spectroscopy for in-line monitoring of low-drug-load formulations [17], and applied spectroscopic data fusion to predict active ingredient concentration and mass gain in coated tablets [21]. Other data types included machine-vision images, endoscopic process images, residence-time distributions, feeder signals, and process parameters [7, 15, 29]. The literature therefore spans high-dimensional spectral data and time-resolved process data, but multimodal integration remains less common than single-modality modelling.

Quality Attributes Predicted

The most frequently reported quality attributes were content uniformity, blend potency, dissolution, moisture content, particle size distribution, tablet defects, and coating-related properties. Studies were found that used artificial neural networks to predict dissolution in a PAT-based setting [6], machine vision and deep learning to estimate particle size and component concentration in powder blends [23], and explainable neural networks to predict moisture content in continuous granulation [27]. Tablet-focused applications included in-line feed-frame potency monitoring [18] and intelligent detection of defective tablets using machine learning and deep learning [26]. These targets indicate that AI-PAT studies have primarily focused on measurable intermediate or final quality attributes with direct relevance to real-time release testing.

Thematic Mapping – Digital Twins in Manufacturing

Definitions and Architectures of Digital Twins

Digital twins were described across a spectrum from data-driven surrogate models to hybrid systems combining mechanistic process knowledge, real-time data streams, and predictive analytics. In pharmaceutical manufacturing, digital twin concepts were reported for continuous powder blending through artificial neural networks and residence-time distribution models [12], and for broader continuous manufacturing frameworks using hybrid modelling strategies [13]. Perspective work extended the concept to drug-product and biopharmaceutical development, describing digital twins as synchronized representations that could support prediction, optimization, and decision-making across manufacturing stages [16]. The mapped literature therefore shows conceptual diversity, with limited standardization in how digital twin fidelity, synchronization, or lifecycle maturity are defined.

AI Components within Digital Twins

AI components within pharmaceutical digital twin papers were most often used for surrogate modelling, model correction, prediction of unmeasured states, data fusion, and soft sensing. Continuous manufacturing studies described soft sensors based on residence-time distributions for content prediction [15], while digital twin work on low-dosage continuous blending

combined neural networks with process-dynamic models [12]. Broader digital twin and data-driven modelling papers described AI as a means to reduce computational burden, update model states, and connect sensor outputs with quality predictions [13, 14]. The literature therefore suggests that AI is not merely an add-on to digital twins, but a key mechanism for translating complex PAT and process data into real-time model behavior.

Integration with Real-Time Data Streams

Limited studies reported actual real-time coupling between PAT data streams and digital twin models, and many digital twin papers remained simulation-based or framework-oriented. Where integration was more developed, studies connected continuous manufacturing process data, PAT outputs, and residence-time or mechanistic models to support process understanding and predictive monitoring [3, 15]. Other work demonstrated real-time or in-line measurement technologies, such as videometric mass-flow control [29], machine-vision-based process endoscopy [7], and real-time release testing in continuous powder blending [5], but these did not always explicitly label the resulting architecture as a digital twin. Overall, the literature indicates that the technical building blocks for real-time digital twins exist, while fully synchronized PAT-AI-digital-twin implementations remain scarce.

Thematic Mapping – Integration of PAT and Digital Twins

Studies Combining PAT with Digital Twins

Only a small subset of the mapped literature explicitly combined multivariate PAT data with digital twin frameworks in pharmaceutical manufacturing. Studies were found that linked residence-time distributions, process models, and continuous manufacturing data for content prediction and process-state estimation [15], while other work combined PAT and mechanistic modelling to improve process understanding in a continuous powder-to-granule line [3]. Digital twin development for low-dosage continuous blending used artificial neural networks alongside residence-time distribution models, showing how data-driven and process-dynamic layers can be joined within a digital representation [12]. However, many papers reported PAT-enabled prediction or digital modelling separately, rather than demonstrating a fully integrated PAT-AI-digital-twin system.

Data Flow and Model Update Strategies

The reported data-flow strategies generally involved extracting features from PAT or process measurements, passing those features into a predictive or hybrid model, and using the model output to estimate product quality or update process state. For example, soft-sensor architectures used residence-time distribution information to link upstream disturbances with downstream content predictions [15], while videometric mass-flow control demonstrated how image-based measurements could support real-time feedback in powder micro-feeding [29]. Hybrid modelling papers described AI as a correction or acceleration layer within process simulations, particularly for continuous pharmaceutical manufacturing [13]. Across the literature, model update strategies were often described conceptually, with fewer studies reporting validated closed-loop updating under routine production conditions.

Thematic Mapping – Implementation and Regulatory Aspects

Scalability and Industrial Deployment

Very few studies reported commercial-scale deployment of AI-PAT-digital-twin systems, and the literature was dominated by laboratory, pilot, and platform studies. Industrially oriented evidence was found in a control strategy for continuous wet granulation [28] and in integrated continuous powder-to-granule processing studies that combined wet granulation, drying, and milling [4]. Broader smart manufacturing discussions emphasized that scalable deployment requires robust data infrastructure, sensor reliability, model lifecycle management, and alignment between digital tools and manufacturing execution systems [1, 14]. The mapped evidence therefore suggests that technical feasibility is better represented than long-term operational performance in routine GMP environments. **Table 1** highlights the deployment evidence gap by distinguishing technical feasibility signals from the operational and regulatory evidence still needed for routine GMP adoption.

Table 1. Evidence Gap Between Technical Feasibility and GMP-Scale Deployment of AI-PAT-Digital-Twin Systems

Evidence dimension	What the mapped literature shows	Remaining gap for routine GMP use
Scale of implementation	Most studies remain laboratory, pilot, or platform-based.	More commercial-scale, multi-batch, and long-duration deployment evidence is needed.
Process integration	Continuous wet granulation and powder-to-granule studies show stronger implementation relevance [4, 28].	Evidence is still limited on full integration across manufacturing execution systems, quality systems, and batch-release workflows.
Sensor and data reliability	PAT and smart manufacturing studies emphasize sensor reliability and robust data infrastructure [1, 14].	Routine evidence is needed on sensor drift, missing data, calibration transfer, and data-governance failures.
Model lifecycle management	AI and PAT reviews identify validation, traceability, interpretability, and lifecycle control as central concerns [2, 10].	Few studies show how models are monitored, updated, locked, or revalidated after deployment.

Regulatory maturity	Real-time release testing and control-strategy examples suggest possible regulatory use cases [5, 28].	Direct evidence remains limited on regulatory interaction, approval pathways, and post-approval maintenance of adaptive AI models.

Regulatory Considerations

Regulatory considerations were most often discussed in relation to process validation, real-time release testing, control strategy definition, and the reliability of models used for quality decisions. Studies were found that addressed real-time release testing of dissolution and blend uniformity using NIR spectroscopy and machine vision [5], while control-strategy work for continuous wet granulation provided a more implementation-oriented example of how PAT and process understanding may support regulated manufacturing [28]. Reviews of PAT monitoring and AI applications emphasized that validation, traceability, interpretability, and lifecycle control remain central for acceptance of AI-enabled manufacturing tools [2, 10]. However, the mapped literature provided limited empirical evidence of direct regulatory interaction, approval pathways, or post-approval maintenance of adaptive AI models.

Workforce and Cultural Factors

Workforce and organizational factors were mentioned less frequently than technical modelling issues, even though they appear important for translating AI-PAT and digital twin systems into practice. Industry 4.0 discussions noted that smart factories require not only digital technologies, but also cross-functional expertise in manufacturing science, data science, quality assurance, and regulatory strategy [1]. Data-driven pharmaceutical process modelling papers similarly implied a need for skills in data engineering, model governance, and process analytics [14]. The literature therefore contains a visible gap around workforce readiness, training models, change management, and the cultural conditions needed for sustained adoption of AI-enabled manufacturing systems.

Research Gaps and Evidence Voids

Lack of Prospective, Commercial-Scale Studies

A major evidence void is the lack of prospective studies evaluating AI-PAT-digital-twin systems during routine commercial GMP manufacturing. Studies were found that demonstrate real-time control of continuous granulation [20], process understanding in fully continuous powder-to-granule lines [3], and real-time release testing in continuous powder blending [5], but these mainly represent focused demonstrations rather than longitudinal industrial deployment. Digital twin reviews and perspectives describe substantial potential across manufacturing and development [16, 19], yet examples of sustained, validated use in commercial release decisions remain rare. This imbalance limits the ability to determine how model performance, data drift, sensor maintenance, and operator interaction behave over extended production campaigns.

Absence of Standardized Evaluation Frameworks

No clear consensus was found on standardized metrics for digital twin fidelity, AI robustness, model transferability, or economic value in pharmaceutical manufacturing. AI reviews summarized many modelling approaches for formulation and PAT applications [10, 11], but these studies varied in how they reported model validation, external testing, and process relevance. Digital twin and hybrid modelling papers described architectures and computational strategies [12, 13], yet fidelity, synchronization frequency, uncertainty propagation, and decision thresholds were not consistently defined. The absence of common evaluation frameworks makes it difficult to compare studies or judge whether a proposed model is ready for process monitoring, real-time release testing, or control.

Under-Explored Unit Operations and Product Types

The mapped literature was concentrated in solid oral dosage manufacturing, especially blending, granulation, coating, tableting, and continuous powder processing. Studies were found that addressed NIR-based blend monitoring [18], twin-screw wet granulation [30], tablet coating using NIR and Raman data fusion [21], and granule tableting-property prediction through a PAT platform [31]. By contrast, relatively few studies addressed biopharmaceutical manufacturing, aseptic processing, sterile fill-finish, packaging, or complex delivery systems within an AI-PAT-digital-twin framework. This distribution indicates that important areas of pharmaceutical production remain under-explored despite their relevance to product quality, supply resilience, and manufacturing innovation.

Table 2 maps the maturity of the evidence base across AI-enhanced PAT, digital twin development, integrated system architecture, regulatory alignment, industrial deployment, and workforce readiness.

Table 2. Evidence-Maturity Matrix for AI, PAT, and Digital Twin Applications in Pharmaceutical Manufacturing

Evidence domain	Dominant evidence pattern in the review	Typical data or technology base	Main manufacturing applications	Current maturity signal	Principal analytical limitation	Translation implication
AI-enhanced PAT	Broadest and most technically developed evidence cluster; dominated	NIR spectroscopy, Raman spectroscopy, process images,	Blend potency, content uniformity, dissolution prediction, moisture	**Moderate maturity** because AI models extend established PAT and	External validation, cross-site transferability, instrument	AI-PAT is the most immediate pathway toward real-time quality prediction,

	by proof-of-concept and pilot-scale models	endoscopic imaging, feeder signals, residence-time data, process parameters	estimation, particle-size distribution, coating quality, tablet defect detection	chemometric traditions	variability, and long-term drift are inconsistently reported	but it still requires lifecycle validation before routine GMP use
Digital twin models	Smaller and more conceptually diverse literature; often framework-oriented or simulation-based	Hybrid process models, surrogate models, residence-time distribution models, soft sensors, state-estimation layers	Continuous blending, continuous granulation, powder-to-granule processing, process simulation, model-based monitoring	**Emergent maturity** because many studies describe architectures rather than synchronized operational twins	Digital twin fidelity, synchronization frequency, model-update logic, and decision thresholds are not standardized	Digital twins need clearer maturity definitions before they can be compared, validated, or accepted as regulated manufacturing tools
Integrated PAT–AI–digital twin systems	Rare evidence cluster; most studies report PAT-enabled prediction or digital modelling separately	Multivariate PAT streams linked with predictive models, process simulations, and control logic	Real-time quality estimation, process-state updating, continuous manufacturing monitoring, potential closed-loop control	**Low maturity** because fully integrated closed-loop demonstrations are uncommon	Limited evidence of persistent data connectivity, validated model updating, interoperability, and quality-system documentation	Integration, not algorithm development alone, is the central bottleneck for next-stage research
Regulatory and quality-system alignment	Frequently discussed but rarely evaluated empirically	Validation protocols, control strategy documentation, audit trails, interpretability, model lifecycle controls	Real-time release testing, continuous process verification, model-supported quality decisions	**Low-to-moderate maturity** because principles are recognized but implementation evidence is limited	Few reports describe regulatory interaction, approval pathways, post-approval maintenance, or adaptive-model governance	Future studies should report how AI-PAT and digital twin tools are documented, controlled, updated, and reviewed within GMP systems
Industrial deployment and scalability	Mostly laboratory, pilot, platform, or retrospective studies; limited commercial-scale longitudinal evidence	Manufacturing execution systems, sensor networks, process-control platforms, data infrastructure, operator interfaces	Continuous manufacturing campaigns, scale-up studies, platform process monitoring	**Low maturity** because routine industrial deployment remains underrepresented	Long-term model performance, sensor maintenance, operator use, economic impact, and campaign-to-campaign robustness are rarely studied	Prospective industrial validation is necessary to move the field beyond technical feasibility
Workforce and organizational readiness	Underdeveloped evidence area despite clear implementation relevance	Cross-functional teams, data science capability, process analytics expertise, QA and regulatory collaboration	Adoption of AI-enabled monitoring, model oversight, deviation management, continuous improvement	**Very low maturity** because workforce factors are mostly implied rather than studied	Training needs, role definitions, cultural readiness, accountability, and change management are rarely addressed	Human and organizational readiness should be treated as part of technical validation rather than as a secondary implementation issue

A Field Dominated by Proof-of-Concept

The most striking pattern in the mapped literature is the predominance of proof-of-concept and pilot-scale demonstrations over routine industrial implementation. A variety of approaches have been reported for continuous granulation, powder blending, spectroscopic monitoring, and image-based process analysis [5, 7, 20], but relatively few papers describe long-term deployment across multiple products, sites, or campaigns. Reviews of AI in PAT and pharmaceutical formulation show rapid methodological expansion [10, 11], yet this expansion has not been matched by equally mature evidence on lifecycle performance. The field therefore appears technically dynamic but still transitional.

Figure 2 presents an evidence-to-implementation map of AI-enabled PAT and digital twin research in pharmaceutical manufacturing.

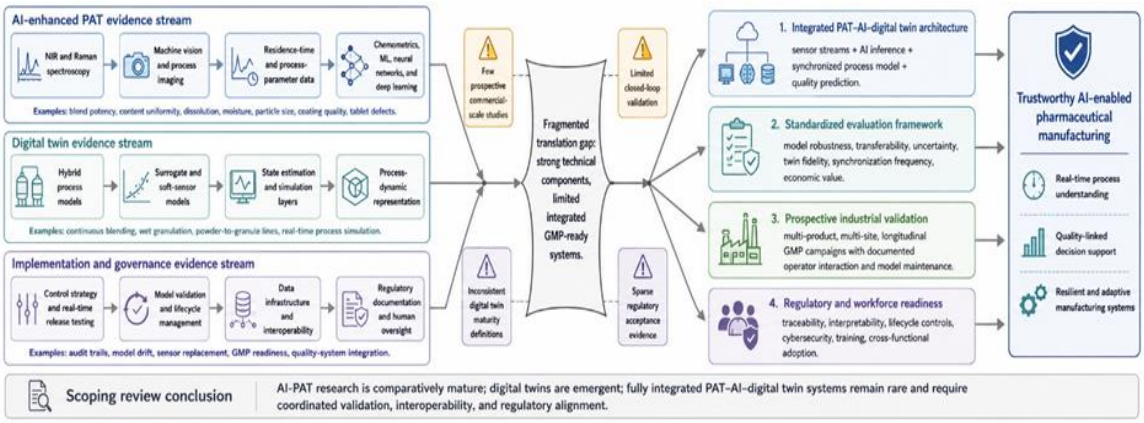

Figure 2. Evidence-to-Implementation Map of AI-Enabled PAT and Digital Twin Research in Pharmaceutical Manufacturing

AI for PAT is More Mature than Digital Twins

AI-enhanced PAT appears more mature than digital twin development because it often builds on established sensor platforms and chemometric traditions. Studies were found that use NIR and Raman spectroscopy for in-line monitoring [17, 18], machine learning-enabled NIR spectroscopy for powder blend assessment [8, 9], and neural networks for PAT-based dissolution testing [6]. Digital twin studies, in contrast, were fewer and often described frameworks, hybrid models, or future-oriented architectures rather than fully synchronized manufacturing replicas [12, 13, 16]. This difference likely reflects lower implementation barriers for AI-PAT models compared with digital twins that require persistent data connectivity, model updating, and system integration.

The Missing Link – Integration and Interoperability

The missing link across the evidence base is not the absence of sensors, algorithms, or process models, but the limited interoperability among them. Studies were found that provide strong individual components, including videometric mass-flow control [29], process endoscopy for particle-size measurement [7], real-time release testing with NIR and machine vision [5], and residence-time-based soft sensing [5]. However, fewer publications describe end-to-end architectures that connect PAT acquisition, AI inference, digital twin state updating, control action, and quality-system documentation. Data silos, proprietary manufacturing platforms, inconsistent metadata, and limited reporting of integration standards appear to constrain progress toward closed-loop AI-PAT-digital-twin systems.

Regulatory and Validation Hurdles

Regulatory and validation hurdles remain central because AI-PAT-digital-twin systems are often intended to influence quality decisions in highly controlled environments. PAT monitoring reviews emphasize the role of validated analytical methods, control strategies, and continuous process verification [2], while continuous wet granulation case work illustrates how control strategies may be defined around process understanding and measurement capability [28]. Explainable soft sensors and AI-based moisture prediction suggest one pathway toward more transparent models [27], but the regulatory treatment of adaptive or frequently updated models remains insufficiently described. As a result, the pathway from a successful pilot model to a regulatory-accepted real-time release or control system remains unclear. **Table 3** summarizes the main regulatory and validation barriers that must be addressed before AI-PAT-digital-twin systems can move from pilot deployment to accepted quality-control or real-time-release use.

Table 3. Regulatory and Validation Barriers for AI-PAT-Digital-Twin Implementation

Barrier	Why it matters for AI-PAT-digital-twin systems	Practical validation response
Analytical method validity	PAT signals must be linked to reliable, validated measurements before they can support quality decisions.	Define method performance criteria, measurement uncertainty, calibration range, and maintenance procedures.
Model explainability	Regulators and quality teams need to understand why the system recommends a control action or release decision.	Use interpretable soft sensors, feature-importance review, and documented mechanistic plausibility checks.
Model updating and drift	Adaptive or frequently updated models may change after validation, creating uncertainty about regulatory acceptability.	Establish change-control rules, drift monitoring, locked model versions, and revalidation triggers.
Control-strategy integration	A model is not sufficient unless it is embedded into a validated process-control strategy.	Link model outputs to predefined control actions, alarm limits, operator review steps, and batch-disposition rules.
Real-time release readiness	Moving from monitoring to real-time release requires stronger evidence that predictions can replace or support conventional testing.	Compare model predictions with reference testing, document failure modes, and define conditions where human review is required.

Emerging Opportunities

Recent methodological developments create opportunities to address some of the mapped technical gaps, particularly in explainability, edge deployment, multimodal fusion, and hybrid modelling. Explainable neural networks have already been reported for moisture prediction in continuous granulation [27], while image-based and endoscopic AI methods suggest growing capacity to interpret high-volume visual PAT data [23, 24]. Hybrid modelling strategies may also help bridge mechanistic understanding and data-driven adaptation in continuous manufacturing digital twins [13]. These developments suggest that future work can move beyond isolated prediction tasks toward transparent, maintainable, and integrated decision-support systems.

Limitations of the Scoping Review

By design, this scoping review mapped the volume, nature, and characteristics of the literature rather than assessing risk of bias or critically appraising individual studies. The included evidence was heterogeneous, spanning review articles on smart manufacturing and AI [1, 10], empirical PAT studies [17, 18], digital twin perspectives [16, 19], and unit-operation-specific demonstrations [3, 30]. This heterogeneity supports broad mapping but limits the extent to which conclusions can be drawn about comparative effectiveness, model superiority, or implementation readiness. The findings should therefore be interpreted as a structured overview of evidence patterns rather than a judgment of study quality.

Call for a Coordinated Research Agenda

The mapped evidence supports the need for a coordinated research agenda that connects AI model development, PAT implementation, digital twin architecture, regulatory expectations, and industrial validation. Studies on continuous manufacturing, real-time release testing, and hybrid modelling provide useful foundations [3, 5, 15], while broader papers on Pharma 4.0 and digital twins describe the system-level transformation required for smart manufacturing [1, 16]. Future research should prioritize interoperable data models, transparent validation protocols, longitudinal deployment studies, and pre-competitive collaboration. Such coordination would help move the field from fragmented demonstrations toward integrated and trustworthy AI-enabled manufacturing systems.

Table 4 translates the review findings into a research agenda focused on validation, interoperability, lifecycle governance, regulatory readiness, workforce adoption, and broader manufacturing coverage.

Table 4. Translational Research Agenda for Moving AI-PAT-Digital Twin Systems from Proof-of-Concept to Regulated Manufacturing Use

Translational priority	Current gap identified by the review	What future studies should report	Suggested evaluation dimensions	Expected contribution to the field
Prospective GMP-scale validation	Most studies remain laboratory-scale, pilot-scale, retrospective, or platform-based	Manufacturing scale, product type, campaign duration, number of batches, operating conditions, deviation handling, operator interaction, and quality-system documentation	Predictive accuracy over time, drift detection, batch-to-batch stability, intervention frequency, failure recovery, release-decision impact	Establishes whether AI-PAT-digital twin systems remain reliable under real manufacturing conditions
Standardized digital twin maturity framework	Digital twins are inconsistently defined across simulation, surrogate modelling, soft sensing, and synchronized process representation	Twin architecture, data-refresh frequency, synchronization mechanism, mechanistic/data-driven balance, update rules, uncertainty handling, and decision role	Fidelity, latency, state-estimation accuracy, model-update validity, uncertainty propagation, control relevance	Enables comparison between digital twin studies and prevents overuse of the term "digital twin" for simple predictive models
Interoperable data architecture	Sensors, AI models, process models, and quality systems are often reported as disconnected components	Data sources, metadata standards, integration interfaces, preprocessing pipelines, storage architecture, audit trails, cybersecurity safeguards	Data completeness, traceability, portability, interoperability, access control, reproducibility, system resilience	Addresses the main practical barrier between isolated models and deployable manufacturing intelligence
Model lifecycle and adaptive governance	Few papers explain how models are maintained after development or adapted after sensor, product, or process changes	Model versioning, retraining triggers, drift monitoring, calibration transfer, sensor replacement procedures, human approval points, documentation workflows	Stability, maintainability, validation burden, explainability, auditability, retraining safety, change-control compatibility	Aligns AI model behavior with pharmaceutical quality-system expectations
Regulatory science and real-time release readiness	Regulatory implications are discussed conceptually, but direct evidence of acceptance pathways is limited	Intended use, quality decision role, validation evidence, interpretability strategy, risk controls, real-time release linkage, regulatory engagement model	Decision criticality, traceability, method validation, uncertainty thresholds, override rules, compliance documentation	Clarifies how AI-PAT and digital twin tools can support quality decisions without weakening regulatory confidence

Human factors and workforce implementation	Workforce readiness, accountability, and operator trust are under-explored	User roles, training requirements, operator interface design, escalation pathways, acceptance barriers, cross-functional governance model	Usability, trust calibration, workload, training effectiveness, decision accountability, adoption sustainability	Ensures that AI-enabled manufacturing systems are usable, interpretable, and governable by real manufacturing teams
Expansion beyond solid oral dosage manufacturing	Evidence is concentrated in blending, granulation, coating, tableting, and continuous powder processing	Product category, unit operation, process risk profile, PAT modality, model transferability, infrastructure requirements	Generalizability, unit-operation coverage, modality suitability, implementation complexity, quality relevance	Extends the field toward biopharmaceuticals, aseptic processing, sterile fill-finish, packaging, and complex dosage systems
Economic and operational value assessment	Most studies emphasize model performance rather than business, quality, or supply-chain impact	Cost of implementation, reduction in failed batches, release-time improvement, maintenance burden, downtime, deviation reduction, supply resilience	Return on investment, process capability, cycle-time reduction, resource utilization, quality-risk reduction	Demonstrates whether technical gains translate into meaningful manufacturing and patient-supply benefits

Strengths and Limitations of This Review

Strengths

A strength of this review is its broad scope across three intersecting domains: artificial intelligence, Process Analytical Technology, and digital twins in pharmaceutical manufacturing. The included literature captured smart manufacturing strategy [1], PAT monitoring frameworks [2], AI reviews for neural networks and formulation modelling [10, 11], and empirical studies across spectroscopy, machine vision, continuous granulation, blending, coating, and tableting [5, 18, 20, 21]. This breadth is appropriate for a scoping review because the aim was to map concepts, methods, and evidence distributions rather than estimate intervention effects. The review also used a structured charting approach to identify where the literature is mature, fragmented, or absent.

Limitations

The review is limited by possible language bias, publication bias, database coverage, and the rapidly evolving nature of AI-enabled manufacturing research. Positive proof-of-concept results may be more likely to appear in the literature than unsuccessful deployments, failed model transfers, or industrial implementation barriers. The field is also changing quickly, as reflected by recent publications on digital twins across pharmaceutical and biopharmaceutical perspectives [16], AI-based process image analysis [7], and intelligent tablet-press defect detection [26]. Consequently, this mapping may require frequent updating as more commercial, regulatory, and interoperable implementations are published.

Implications and Recommendations

For Researchers

Researchers should prioritize prospective industrial-scale pilot studies, open benchmarking datasets, and transparent validation protocols for AI-PAT and digital twin models. The current literature shows strong methodological development in NIR, Raman, image analysis, and neural-network-based soft sensing [8, 17, 23, 27], but fewer studies evaluate transferability across products, sites, instruments, or campaigns. Open-source benchmarks for PAT spectra, process images, residence-time models, and digital twin state estimation would improve comparability and reproducibility. Studies should also examine economic impact, workforce requirements, maintenance burden, and failure modes rather than focusing only on predictive accuracy.

For Industry and Regulators

Industry and regulators should consider pre-competitive mechanisms for defining validation expectations, lifecycle controls, and documentation standards for AI-enabled manufacturing systems. Control strategy studies in continuous wet granulation [28], real-time release testing work in continuous blending [5], and PAT monitoring reviews [2] indicate that the technical basis for quality-linked modelling is already developing. However, consistent expectations are needed for model updates, data drift, sensor replacement, cybersecurity, audit trails, and human oversight. Regulatory sandboxes or structured pilot programs could help clarify how AI-PAT-digital-twin systems can be evaluated before use in routine GMP decision-making.

For Funders and Policy Makers

Funders and policy makers should support translational research that bridges academic modelling, industrial manufacturing platforms, and regulatory science. The literature includes advanced demonstrations of continuous processing and hybrid modelling [3, 4, 13], but the cost and complexity of integrated AI-PAT-digital-twin testbeds may exceed the resources of single research groups. Shared infrastructure, data-sharing agreements, and public-private partnerships could accelerate progress while protecting proprietary product knowledge. Investment should also support training programs that combine pharmaceutical engineering, process analytics, data science, and quality systems.

Conclusion

AI applications in pharmaceutical manufacturing involving PAT and digital twins have expanded rapidly since 2017, but the field remains largely in the proof-of-concept stage. The evidence base shows considerable technical creativity, especially in the use of machine learning and deep learning to interpret sensor-rich manufacturing data. However, the transition from experimental demonstration to routine GMP implementation remains limited.

AI-enhanced PAT is the most mature area of the literature because it builds on established spectroscopic, imaging, and process-monitoring technologies. Digital twin development is more emergent, with many papers describing frameworks, hybrid models, or simulation-oriented architectures. The full integration of PAT, AI, and digital twins remains an important unresolved gap.

Crucial evidence voids exist in industrial-scale validation, regulatory acceptance, interoperability, and workforce readiness. The literature provides comparatively little insight into long-term model maintenance, cross-site transfer, adaptive model governance, or the practical burden of sustaining AI-enabled systems in production. These gaps are especially important because pharmaceutical manufacturing decisions are tightly connected to patient safety and product quality.

A coordinated effort across academia, industry, technology providers, and regulators is required to translate the technical promise of AI-PAT-digital-twin systems into tangible manufacturing innovation. Future work should move beyond isolated prediction models toward integrated, validated, and explainable systems that can support process understanding, real-time quality assurance, and resilient pharmaceutical supply. The next stage of the field will depend on whether these technologies can be made trustworthy, interoperable, and usable in regulated manufacturing environments.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

References

1. Arden NS, Fisher AC, Tyner K, Yu LX, Lee SL, Kopcha M, et al. Industry 4.0 for pharmaceutical manufacturing: Preparing for the smart factories of the future. Int J Pharm. 2021;602:120554.
2. Kim EJ, Kim JH, Kim MS, Jeong SH, Choi DH. Process analytical technology tools for monitoring pharmaceutical unit operations: a control strategy for continuous process verification. Pharmaceutics. 2021;13(6):919.
3. Domokos A, Pusztai É, Madarász L, Nagy B, Gyürkés M, Farkas A, et al. Combination of PAT and mechanistic modeling tools in a fully continuous powder to granule line: Rapid and deep process understanding. Powder Technol. 2021;388:70-81.
4. Fülöp G, Domokos A, Galata D, Szabó E, Gyürkés M, Szabó B, et al. Integrated twin-screw wet granulation, continuous vibrational fluid drying and milling: A fully continuous powder to granule line. Int J Pharm. 2021;594:120126.
5. Mészáros LA, Gyürkés M, Varga E, Tacsi K, Honti B, Borbás E, et al. Real-time release testing of in vitro dissolution and blend uniformity in a continuous powder blending process by NIR spectroscopy and machine vision. Eur J Pharm Biopharm. 2024;201:114368.
6. Nagy B, Petra D, Galata DL, Démuth B, Borbás E, Marosi G, et al. Application of artificial neural networks for Process Analytical Technology-based dissolution testing. Int J Pharm. 2019;567:118464.
7. Madarász L, Mészáros LA, Köte Á, Farkas A, Nagy ZK. AI-based analysis of in-line process endoscope images for real-time particle size measurement in a continuous pharmaceutical milling process. Int J Pharm. 2023;641:123060.
8. Ali H, Muthudoss P, Ramalingam M, Kanakaraj L, Paudel A, Ramasamy G, et al. Machine learning–enabled NIR spectroscopy. Part 2: workflow for selecting a subset of samples from publicly accessible data. AAPS PharmSciTech. 2023;24(1):34.
9. Muthudoss P, Tewari I, Chi RL, Young KJ, Ann EY, Hui DN, et al. Machine learning-enabled NIR spectroscopy in assessing powder blend uniformity: clear-up disparities and biases induced by physical artefacts. AAPS PharmSciTech. 2022;23(7):277.
10. Nagy B, Galata DL, Farkas A, Nagy ZK. Application of artificial neural networks in the process analytical technology of pharmaceutical manufacturing—a review. AAPS J. 2022;24(4):74.
11. Wang S, Di J, Wang D, Dai X, Hua Y, Gao X, et al. State-of-the-art review of artificial neural networks to predict, characterize and optimize pharmaceutical formulation. Pharmaceutics. 2022;14(1):183.
12. Beke ÁK, Gyürkés M, Nagy ZK, Marosi G, Farkas A. Digital twin of low dosage continuous powder blending–Artificial neural networks and residence time distribution models. Eur J Pharm Biopharm. 2021;169:64-77.
13. Bhalode P, Chen Y, Ierapetritou M. Hybrid modelling strategies for continuous pharmaceutical manufacturing within digital twin framework. In: Computer Aided Chemical Engineering. Vol. 49. Elsevier; 2022. p. 2125-2130.
14. Dong Y, Yang T, Xing Y, Du J, Meng Q. Data-driven modeling methods and techniques for pharmaceutical processes. Processes. 2023;11(7):2096.

15. Gyürkés M, Madarász L, Záhonyi P, Köte Á, Nagy B, Pataki H, et al. Soft sensor for content prediction in an integrated continuous pharmaceutical formulation line based on the residence time distribution of unit operations. Int J Pharm. 2022;624:121950.
16. Maharjan R, Kim NA, Kim KH, Jeong SH. Transformative roles of digital twins from drug discovery to continuous manufacturing: Pharmaceutical and biopharmaceutical perspectives. Int J Pharm X. 2025;100409.
17. Harms ZD, Shi Z, Kulkarni RA, Myers DP. Characterization of near-infrared and Raman spectroscopy for in-line monitoring of a low-drug load formulation in a continuous manufacturing process. Anal Chem. 2019;91(13):8045-53.
18. De Leersnyder F, Peeters E, Djalabi H, Vanhoorne V, Van Snick B, Hong K, et al. Development and validation of an in-line NIR spectroscopic method for continuous blend potency determination in the feed frame of a tablet press. J Pharm Biomed Anal. 2018;151:274-83.
19. Davidopoulou C, Ouranidis A. Pharma 4.0-artificially intelligent digital twins for solidified nanosuspensions. Pharmaceutics. 2022;14(10):2113.
20. Dosta M, Schneider M, Geis CW, Schulte L, Kriegl JM, Gomez AM, et al. Machine learning real-time control of continuous granulation process. Int J Pharm. 2025;126244.
21. Szabó-Szőcs B, Ficzere M, Peterfi O, Galata DL. Simultaneous prediction of the API concentration and mass gain of film coated tablets using Near-Infrared and Raman spectroscopy and data fusion. Int J Pharm. 2025;668:124957.
22. Munir N, Nugent M, Whitaker D, McAfee M. Machine learning for process monitoring and control of hot-melt extrusion: Current state of the art and future directions. Pharmaceutics. 2021;13(9):1432.
23. Ficzere M, Péterfi O, Farkas A, Nagy ZK, Galata DL. Image-based simultaneous particle size distribution and concentration measurement of powder blend components with deep learning and machine vision. Eur J Pharm Sci. 2023;191:106611.
24. Péterfi O, Kállai-Szabó N, Demeter KR, Barna ÁT, Antal I, Szabó E, et al. Artificial intelligence-aided endoscopic in-line particle size analysis during the pellet layering process. J Pharm Anal. 2025;101227.
25. Gyürkés M, Madarász L, Köte Á, Domokos A, Mészáros D, Beke ÁK, et al. Process design of continuous powder blending using residence time distribution and feeding models. Pharmaceutics. 2020;12(11):1119.
26. Kim SH, Han SH. Development of an intelligent tablet press machine for the in-line detection of defective tablets using machine learning and deep learning models. Pharmaceutics. 2025;17(4):406.
27. Záhonyi P, Fekete D, Szabó E, Nagy ZK, Nagy B. Explainable artificial neural network as a soft sensor to predict the moisture content in a continuous granulation line. Eur J Pharm Sci. 2025;212:107173.
28. Chavez PF, Stauffer F, Eeckman F, Bostijn N, Didion D, Schaefer C, et al. Control strategy definition for a drug product continuous wet granulation process: Industrial case study. Int J Pharm. 2022;624:121970.
29. Madarász L, Köte Á, Gyürkés M, Farkas A, Hambalkó B, Pataki H, et al. Videometric mass flow control: A new method for real-time measurement and feedback control of powder micro-feeding based on image analysis. Int J Pharm. 2020;580:119223.
30. Zhao J, Tian G, Qu H. Pharmaceutical application of process understanding and optimization techniques: A review on the continuous twin-screw wet granulation. Biomedicines. 2023;11(7):1923.
31. Casian T, Nagy B, Lazurca C, Marcu V, Tőkés EO, Kelemen ÉK, et al. Development of a PAT platform for the prediction of granule tableting properties. Int J Pharm. 2023;648:123610.