Integrating Artificial Intelligence and Big Data in Pharmaceutical Development: Ethical Considerations and Legislative Frameworks

Diego Hernández Pedraza · Mario López de la Cruz · David F. Nettleton · Alexandra Poch Solà

Abstract

The interconnected spheres of pharmaceutical development (ICH Q8), quality risk management (ICH Q9), drug substance development and manufacture (ICH Q11) and continuous manufacturing of drug substances and drug products (ICH Q13) outlined by the International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH), provide a robust framework for the pharmaceutical industry, focusing on various critical aspects of drug development and manufacturing. Within this framework, the application of Artificial Intelligence (AI) and Big Data (BD) presents an unprecedented opportunity to enhance quality assurance, streamline drug development processes, increase traceability at supply chain level, and ultimately improve quality and safety of medicinal products. This paper discusses the ethical considerations and legislative imperatives for integrating AI and BD within the ICH guidelines to ensure responsible usage and sustainable business outcomes.

Keywords artificial intelligence · big data · pharmaceutical development · medicinal products lifecycle · ethical use · responsible usage of artificial intelligence

Introduction

In the rapidly evolving landscape of pharmaceutical research and development, the integration of artificial intelligence (AI) and big data (BD) has emerged as a pivotal factor in streamlining processes and enhancing outcomes, as this offers the potential to analyse large datasets and identify potential new compounds, which can accelerate the discovery and development of new drugs. Some applications of IA and BD in the pharmaceutical landscape include the reduction of resources needed in drug development, optimization of design and operation of pharmaceutical manufacturing facilities, optimization of clinical trials design, diagnosis using sophisticated algorithms, determination of epidemiology forecasts, and development of medical devices. Within the manufacturing framework, as the pharmaceutical industry embraces a Quality by Design (QbD) approach, ensuring the responsible and ethical use of these advanced technologies is paramount to the integrity and success of drug development, regulatory compliance, and patient safety.

The pharmaceutical industry is at a pivotal point with digital technologies, as these are set to become central to drug development and manufacturing processes. These technologies offer the potential to significantly enhance the application of quality principles described in ICH guidelines (such as ICH Q8 (R2) (1), Q9 (R1) (2), Q11 (3) and Q13 (4)) during medicinal products development and lifecycle. However, such advances also introduce complex ethical and legislative challenges, particularly around data privacy, algorithmic transparency, and decision-making accountability. It has become crucial to assess these challenges to leverage digital innovation for sustainable and ethical pharmaceutical outcomes.

This paper aims to provide a comprehensive background into the adoption of AI and BD within the pharmaceutical sector, highlighting the novelty of these

technologies, laying out the objectives of their application in an ethical manner and connecting with the newly established AI regulatory framework that was endorsed on March 13th, 2024 by the European Parliament (5) (6).

Following a thorough literature review, we go on to identify gaps that should be addressed in the upcoming years to assure an ethical application of digital technologies.

Literature review

The review of current publications was designed to identify published studies that included information related to the application of AI to the pharmaceutical industry landscape, in particular to drug development and manufacture. While this paper is not a systematic review, the principles described in the PRISMA statement 2020 (7) have been partially followed for the review of current publications. The search for studies was conducted up to March 2024 using a comprehensive search in the following databases:

- MEDLINE
- SCIELO
- Database of Abstracts of Reviews of Effects (DARE)
- International HTA Database

Free full texts studies with review article style in English and Spanish published between the years 2020 and 2024 related to AI that included aspects of drug development and manufacture were considered.

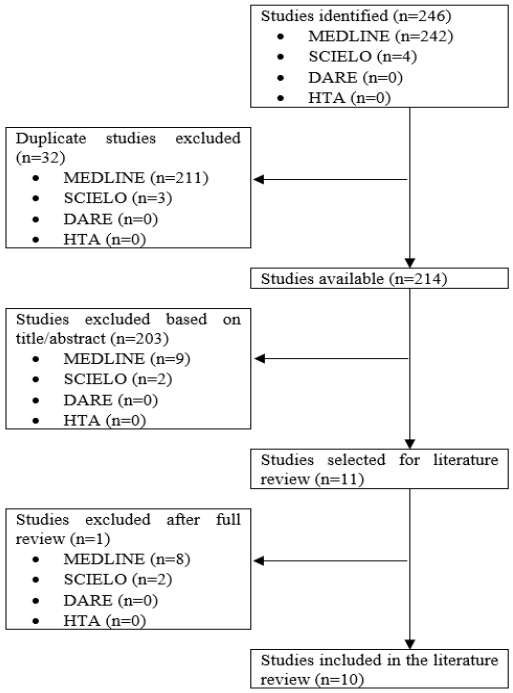
The search strategies carried out using MeSH (Medical Subject Headings) terms are outlined in Table 1.

Methods

The number of studies identified through the search strategies in various databases were pooled to eliminate possible duplicates. Out of a total of 246 identified studies, 32 were excluded due to duplication. Subsequently, titles and abstracts were screened to identify publications containing information relevant to the objective of the present study and to exclude those that did not fall within its scope. A total of 11 studies were selected with 203 excluded. Thereafter, a full-text review of the 11 selected studies was conducted. Following this review, 1 study was excluded as being out of the scope of the present study.

Ultimately, a total of 10 studies were included in the litereature review. The selection process is depicted in Figure 1. The search, identification, review, exclusion, selection, and evaluation of the studies were conducted by both authors* of the literature review (**DHP* and *MLdlC*).

Figure 1. Studies search flow and selection



Results and discussion

AI in pharmaceutical development faces significant challenges around market positioning, prediction and analysis, pharmaceutical manufacturing, quality control and assurance, and clinical trial design, primarily derived from the complexity and volume of data involved (8).

Table 1. Search Strategies

Number	Search	Publications
MEDLINE (2020 - 2024)		
1	("Artificial Intelligence"[Mesh]) AND "Drug Development"[Mesh]	172
2	("Artificial Intelligence"[Mesh]) AND "Databases, Pharmaceutical"[Mesh]	0
3	("Artificial Intelligence"[Mesh]) AND "Chemistry, Pharmaceutical"[Mesh]	5
4	("Artificial Intelligence"[Mesh]) AND "Pharmaceutical Research"[Mesh]	2
5	("Artificial Intelligence"[Mesh]) AND "Ethics, Pharmacy"[Mesh]	0
6	("Artificial Intelligence"[Mesh]) AND "Drug Industry"[Mesh]	10
7	("Artificial Intelligence"[Mesh]) AND "Drug Design"[Mesh]	48
8	("Artificial Intelligence"[Mesh]) AND "Health Care Sector"[Mesh]	5
SCIELO (2020 - 2024)		
9	("Artificial Intelligence"[Mesh]) AND "Drug Development"[Mesh]	3
10	("Artificial Intelligence"[Mesh]) AND "Drug Industry"[Mesh]	1
DARE (2020 - 2024)		
11	("Artificial Intelligence"[Mesh]) AND "Drug Development"[Mesh]	0
HTA (2020 - 2024)		
12	("Artificial Intelligence"[Mesh]) AND "Drug Development"[Mesh]	0

Among the principal challenges of the application of AI within the pharmaceutical and biomedical sectors is the need for vast amounts of reliable data. This is foundational for AI's success in pharmaceuticals yet, accessing such data across various databases can be a costly endeavour. Additionally, the opaque nature of some AI models, the so-called "Black Box" phenomenon, stirs scepticism towards AI-generated data due to the challenges in deciphering the decision-making processes of these models (9).

The ethical issues associated with the predictions made by AI, especially those that have direct implications for the health and well-being of patients, should be taken also into consideration. It is crucial to meticulously evaluate the potential effects of AI, particularly regarding any biases or discrimination that could arise, and how these might impact outcomes (10).

Health Authorities are currently implementing the use of AI in different areas. The United States Food and Drug Administration (FDA) is developing a strategic framework for incorporating AI within the medical products sector. This framework emphasizes a risk-based approach that spans the entire lifecycle of medical products. Highlighting the importance of collaboration, innovation in regulatory methods, establishment of harmonized standards, and the advancement of AI performance evaluation research, the FDA aims to ensure the responsible deployment of AI technologies in healthcare. The strategy focuses on ensuring the safety, efficacy, and regulatory adherence of these technologies, ultimately aiming to bolster public health innovation (11).

AI and BD application in the context of the ICH Guidelines has been an area of development in the last several years. AI can support the quality risk management across pharmaceutical development and manufacturing described in ICH Q9 by enabling predictive analysis and proactive identification of potential quality issues, enhancing both product safety and efficacy.

Drug substance development and manufacture principles described in ICH Q11 are enhanced with AI optimizing processes, predicting the outcomes of synthesis paths and ensuring consistency in drug substance quality. In the area of concern of ICH Q13, the application of AI means optimization of process parameters and energy usage, aiding in decision-making, prediction outcomes and risks minimization in scale-ups and technology transfers. AI can significantly amplify the pharmaceutical development principles covered by ICH Q8, advocating for a QbD approach to product and process development, supporting the Design of Experiment (DoE) activities for process scale-up by the prediction of system behaviour through digital twins, simulations and modelling, and ultimately enabling more informed decision-making and regulatory flexibility.

QbD is a systematic approach to pharmaceutical development that begins with predefined objectives and emphasizes product and process understanding and process control, based on sound science and quality risk management. Within this framework, AI and BD can play a transformative role in several key areas:

- Critical Quality Attributes (CQAs) Identification: AI algorithms can analyse large datasets from previous product development cycles to predict potential CQAs for new drugs. By learning from historical data, AI models can uncover patterns and correlations that might not be immediately apparent to human researchers, leading to a more comprehensive identification of attributes critical to product quality.
- Design Space Exploration: AI can significantly enhance the exploration of design space—the multidimensional combination and interaction of input variables (e.g., material attributes) and process parameters that have been demonstrated to provide assurance of quality. Through sophisticated modelling and simulation techniques, AI can explore vast design spaces more efficiently than traditional methods and can even enable the adjustment of equipment specifications and process parameters at the industrialisation stage. Machine learning models can simulate countless experiments to predict how changes in process parameters can affect CQAs, enabling developers to optimize their manufacturing process more rapidly and effectively.
- Risk Assessment and Mitigation Strategies: The application of AI can improve the robustness of risk assessments within the QbD framework through predictive analytics. By analysing complex datasets, AI can help identify potential risks to quality earlier in the development process. Additionally, it can help in crafting more effective mitigation strategies by predicting the impact of different risk management approaches.
- Real-time Process Monitoring and Control: By integrating AI with process analytical technology (PAT) for example based on near infrared (NIR) or Raman Spectroscopy, manufacturers can move towards a more dynamic control strategy within the established design space. AI models, trained on PAT data, can predict deviations from the desired product quality in real-time and adjust process parameters automatically to maintain quality within the specified design space. This enables not just a more consistent product quality but also enhances manufacturing efficiency by minimizing waste and reducing the need for batch rejections.

Some publications for the specific application of AI to QbD developments are available. An example is the optimization of the formulation of a poorly water-soluble drug. Following a QbD approach, the impact of critical process parameters on particle size, size distribution, and entrapment efficiency was examined. Furthermore, response surface designs were utilized for

the screening and optimization of lecithin/chitosan nanoparticles, obtaining finally an optimized formulation for the drug. (12).

Legislative and Ethical Framework for AI in QbD

Multiple AI applications on pharmaceutical environment have been discussed. However, the legislative and ethical framework for its application is still under development.

On March 13th 2024 the European Parliament ratified the Artificial Intelligence Act (13), a law designed to ensure the safe use of AI and its compliance with essential human rights while promoting innovation. The Act seeks to shield fundamental human rights, democracy, legal integrity, and environmental health from the risks posed by high-risk AI systems. At the same time, it aims to enhance innovation and position Europe as a frontrunner in AI technology. The regulation outlines specific requirements for AI, based on the level of risk and the potential impact it carries.

The AI Act thoroughly addresses how AI technologies should be incorporated within a variety of legal frameworks, including those applicable to medical devices and *in vitro* diagnostic medical devices. This comprehensive inclusion considerably affects the pharmaceutical sector, particularly companies involved in creating AI-driven medical devices or diagnostics, impacting their product development, regulatory compliance, and strategies for market entry.

Although the specifics of the risk assessment are not explicitly cited, the AI Act introduces a regulatory approach based on risk. This method sorts AI technologies by their potential risks, applying more rigorous scrutiny to applications considered high risk. Pharmaceutical firms must determine the classification of their AI tools under this framework and acquaint themselves with the necessary regulatory standards.

For pharmaceutical companies, actively navigating this changing regulatory environment is vital. Being proactive and compliant with these legal mandates can offer a strategic advantage, facilitating faster access to markets and encouraging continued innovation.

The expanded role of AI in pharmaceutical development within the new regulatory framework highlights the need for a sophisticated legislative and ethical framework that addresses:

Data Integrity and Quality: Legislation should ensure that AI systems used in pharmaceutical development are trained on high-quality, reliable data. This includes standards on data collection, storage, and processing practices.

Algorithm Validation: Given the critical role of AI in determining product quality, it is essential for legislative bodies to set standards for the validation of AI algorithms. This includes transparency in the algorithm's decision-making process, especially how it translates data into predictions or adjustments within the design space.

Continual Learning and Model Updating: AI models benefit from continual learning, incorporating new data as it becomes available. Regulations should facilitate this dynamic model improvement while ensuring that updates do not compromise product quality or patient safety.

Ethical Considerations

The core of the new European AI Act is the pledge to cultivate a secure and trustworthy ecosystem for AI technologies. This entails the utilization of AI not just for its novelty but also for its ability to enhance patient safety, protect data privacy, and ensure system reliability. Ethically, it fosters the development of AI solutions that people can trust, assuring that their rights and well-being are protected.

The legislation aims to regulate AI applications to reduce risks while simultaneously promoting innovation,

especially in critical sectors such as healthcare. It ethically motivates the pharmaceutical industry to engage in AI research and development that can yield considerable benefits for society, including better health outcomes, more efficient drug discovery processes, and customized treatments, all within the bounds of ethical practices. This approach establishes an ethical framework where innovation transcends mere progress to make a positive impact on societal health and well-being.

The integration of AI and BD into pharmaceutical practices invites scrutiny on several ethical fronts:

- Privacy and Confidentiality: With vast amounts of data being processed, ensuring patient data privacy becomes paramount. Legislation must evolve to protect individuals' data rights without stifling innovation.
- Transparency and Explicability: AI models must be transparent and their decisions, especially those affecting drug safety and efficacy, must be explainable to regulators, industry stakeholders, and the public.
- Accountability: Clear guidelines are needed to establish accountability for decisions made with the assistance of AI, especially in high-stakes contexts such as drug safety.
- Data integrity and protection: Personal data collected for other purposes can be processed within the AI regulatory scope under specific conditions, ensuring the development of innovative AI systems aligns with substantial public interest. Such data processing requires the innovative AI systems to be developed under the control and responsibility of competent authorities and based on Member State or Union law. Importantly, the data are to be processed in a functionally separate, isolated, and protected data processing environment under the control of participants, with only authorized persons having access.

Legislative Frameworks

The new European AI Legislation plays a crucial role in navigating the ethical challenges presented by AI and BD, ensuring responsible use by:

- Data Protection Regulations: Strengthen and harmonize data protection laws globally, with specific provisions for high-sensitivity sectors like healthcare and pharmaceuticals.
- Standards for Algorithmic Transparency: Develop standards mandating the transparency of AI algorithms, including documentation of data sources, modelling processes, and decision pathways.
- Framework for Accountability: Establish a clear legal framework that outlines the accountability mechanisms for AI-driven decisions in pharmaceutical development and manufacture.
- Implication for Product Development and Market Access: the integration of AI systems within various legislative frameworks, including those concerning medical devices and *in vitro* diagnostic medical devices. This integration directly impacts the pharmaceutical industry, especially companies developing AI-powered medical devices or diagnostics, as it may affect product development timelines, regulatory approval processes, and market access strategies.
- Innovation Support Measures: This aspect is crucial for pharmaceutical companies investing in AI for drug discovery, personalized medicine, and patient care optimization. Companies should stay informed about specific programs or opportunities that might arise from this regulatory framework to support innovation in the sector.

Example Case Studies: ETERNAL project

The ETERNAL project is a four-year HORIZON Research and Innovation Action running from September 2022 until August 2026. It brings together sixteen partners

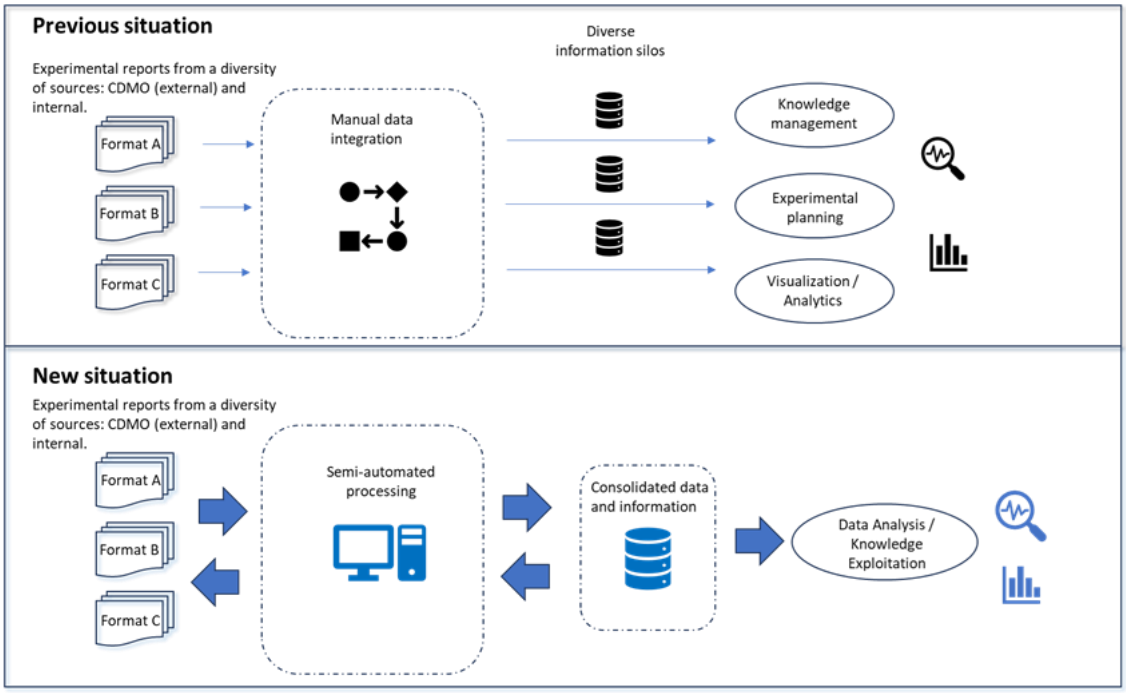
from across Europe united by the motivation to contribute to sustainable development of pharmaceutical products manufacture, use and disposal.

Within this project, several key tasks related to the digitalization of data flow and experimental process optimization are being developed.

In a first task, an exemplar complex data flow of a pharma research environment is being digitalized in order to improve the efficiency and homogeneity of the processing of documents related to empirical testing of pharmaceutical products. Large amounts of data are processed and integrated from different sources (internal and external to the company), thus involving (with reference to the list in the previous section) issues such as data integrity, data privacy, data storage, and data access. Furthermore, exhaustive validation is required to demonstrate the correct functionality of the data processing with a diversity of different data formats and volumes. The data processing system of the pharma company is already compliant to the required standards and legislation, and the regulations will be applied to the required modifications in this task of digitalization of the data flow.

Figure 2 illustrates the schema of the dataflow and how it will be digitised, passing from a semi-manual system to an automated one, and rationalization of processes. One current difficulty resides in the inefficiencies arising from manually processing different formats from diverse data sources which need to be integrated. Note that this is only one example workflow, a large organisation will have multiple workflows.

Figure 2. Digitalization of data flow and processes



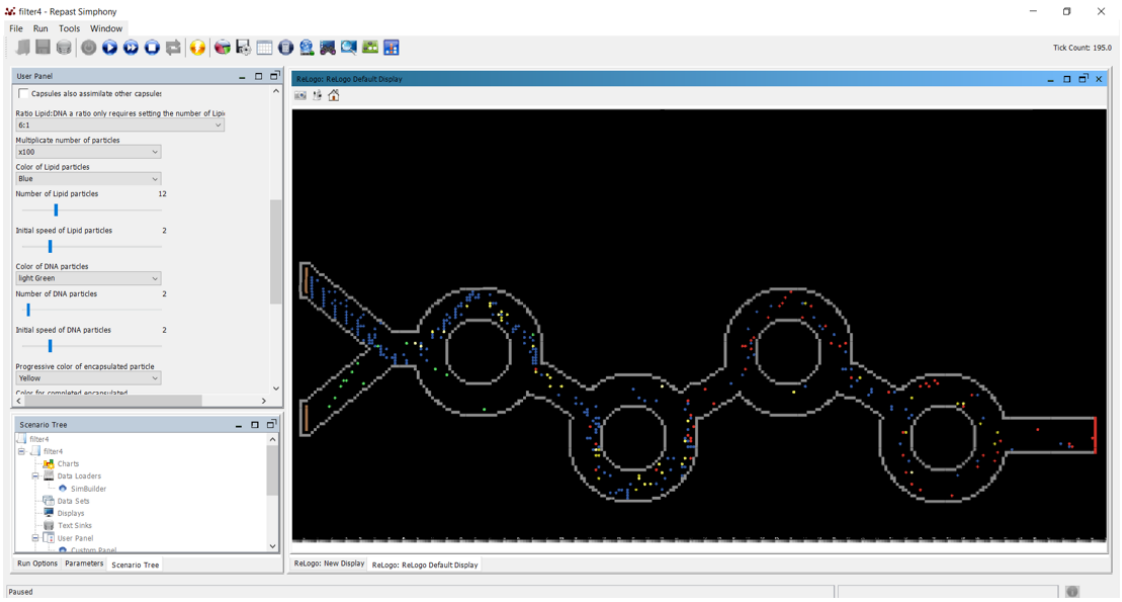
Furthermore, in terms of accountability, the human is kept "in the loop" in order to verify data formats before they are committed to be processed.

In another task, "digital twins" are being built to simulate core processes of three different pharma companies, in order to accelerate experiments and optimize key KPIs such as increasing productivity on the one hand, and reducing environmental impact, on the other. Some of the information required to digitalize the processes is IP/confidential belonging to the pharma company, together with data required to calibrate and train the models. Also, validation is a key aspect in order to demonstrate reliability, robustness and

accuracy of the models (digital twins) under different realistic conditions.

Figure 3 shows the user screen of one of the digital twins developed in the ETERNAL project, with the user controls on the left, and the process simulation visualisation on the right (micromixer for encapsulation of DNA and lipids).

Figure 3. Digital Twin (micromixer encapsulation)



In terms of accountability, as with the data flow digitalization, the human is kept “in the loop” in order to verify the experimental results of the digital twin simulations, before committing to proceeding with physical experimentation.

The design of experiment and laboratory testing of the pharma companies is already compliant to the required standards and legislation, and the regulations will be applied to the incorporation of the digital twins as an additional support tool in the laboratory setting.

Conclusions

The advent of sophisticated AI and BD technologies heralds a new era of potential for the pharmaceutical industry, from drug development to manufacturing processes. However, the journey towards fully leveraging these benefits demands a collaborative and responsible approach, intertwining efforts from regulatory authorities, industry stakeholders, and technology innovators. Legislation needs to adapt continually to meet the distinct challenges that AI and BD technologies introduce, guiding their ethical and sustainable application in pharmaceutical practices in alignment with the ICH guidelines Q8, Q9, Q11 and Q13. Striking a delicate balance between fostering innovation and upholding ethical responsibilities emerges as paramount for unlocking the vast opportunities digital technologies present for enhancing patient outcomes and ensuring the pharmaceutical industry's sustainable development.

The introduction of the new European AI Act cultivates a principled and balanced framework for the utilization of AI across the pharmaceutical sector and broader fields. This progressive legislation encourages the development of AI technologies that are ethical, secure, and socially advantageous, ensuring that technological progression in AI proceeds in tandem with a deep-seated respect for human rights, accountability, and the collective good of society.

This paper endeavours to bridge the gap between the ICH guidelines and the ethical, conscientious deployment of AI and BD within pharmaceutical development. It underscores the urgency for adopting forward-thinking, interactive strategies in decision-making processes within new, robust, enhanced European legislative frameworks to steer these technological integrations. Delving into specific examples or case studies could further enrich the discussion, offering more tangible insights into the application and implications of these technologies in the pharmaceutical context.

Bibliography

1. ICH Q8(R2). Pharmaceutical Development Step 4: International Council for Harmonisation (ICH); [Available from: https://database.ich.org/sites/default/files/Q8%28R2%29%20Guideline.pdf.
2. ICH Q9(R1). Quality Risk Management Step 4: International Council for Harmonization (ICH); 2023 [Available from: https://database.ich.org/sites/default/files/ICH_Q9%28R1%29_Guideline_Step4_2023_0126_0.pdf.
3. ICH Q11. Development and Manufacture of Drug Subtances (Chemical Entities and Biotechnological/Biological Entities) Step 4: International Council for Harmisation (ICH); [Available from: https://database.ich.org/sites/default/files/Q11%20Guideline.pdf.
4. ICH Q13. Continuous Manufacturing of Drug Substances and Drug Products Step 4: International Council for Harmisation (ICH); [Available from: https://database.ich.org/sites/default/files/ICH_Q13_Step4_Guideline_2022_1116.pdf.
5. Parliament E. European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council establishing a common framework for media services in the internal market (European Media Freedom Act) and amending Directive 2010/13/EU (COM(2022)0457 – C9-0309/2022 – 2022/0277(COD)) [Available from: https://www.europarl.europa.eu/doceo/document/TA-9-2024-0137_EN.pdf.
6. Parliament E. European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)).
7. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.
8. Paul D, Sanap G, Shenoy S, Kalyane D, Kalia K, Tekade RK. Artificial intelligence in drug discovery and development. Drug Discov Today. 2021;26(1):80-93.
9. Selvaraj C, Chandra I, Singh SK. Artificial intelligence and machine learning approaches for drug design: challenges and opportunities for the pharmaceutical industries. Mol Divers. 2022;26(3):1893-913.
10. Saldivar-González FI. Artificial intelligence in drug design: towards augmented intelligence. Educación Química. 2023.
11. FDA. Artificial Intelligence& Medical Products: How CBER, CDER, CDRH, and OCP are Working Together. 2024.
12. Dawoud MHS, Mannaa IS, Abdel-Daim A, Sweed NM. Integrating Artificial Intelligence with Quality by Design in the Formulation of Lecithin/Chitosan Nanoparticles of a Poorly Water-Soluble Drug. AAPS PharmSciTech. 2023;24(6):169.
13. European Media Freedom Act. European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council establishing a common framework for media services in the internal market (European Media Freedom Act)

and amending Directive 2010/13/EU (COM(2022)0457 – C9-0309/2022 – 2022/0277(COD)), (2024).