Acta Scientific MEDICAL SCIENCES (ISSN: 2582-0931)

Volume 10 Issue 7 July 2026

Review Article

The Impact of Artificial Intelligence and Digital Health on Modern Clinical Research: Applications, Challenges, and Ethical Aspects

Simona Napoli[1], Daniela Maria Capuano[2,3] and Roberto Verna[1-3]*

[1]Research and Training Center for Health and Well-Being, Italy

[2]Academy for Health and Clinical Research, Italy

[3]World Association of Societies of Pathology and Laboratory Medicine, Italy

***Corresponding Author:** Roberto Verna, Research and Training Center for Health and Well-Being, Italy.

Received: June 15, 2026

Published: July 03, 2026

DOI: 10.31080/ASMS.2026.10.2260

Abstract

In recent years, clinical research has been undergoing a profound transformation thanks to technological innovation and the increasing digitalization of healthcare processes. In particular, Artificial Intelligence and Digital Health are becoming the key drivers in the evolution of the design, conduct, and analysis of clinical studies. The integration of these technologies promises to make clinical research more efficient and more patient-centered, responding not only to the needs of increasingly complex healthcare systems but also, and above all, to the importance of generating concrete and reliable data in a short time. The use of Artificial Intelligence has proven to be a valuable support for the analysis of large volumes of heterogeneous data; the challenge lies in the enormous amount of data that are constantly generated from different sources and require specific processing to verify their truthfulness and accuracy. Digital Health, on the other hand, encompasses a range of technological solutions: electronic health records, wearable devices, mobile applications, and telemedicine systems that enable remote physician-patient interaction and continuously generate real-time health data. Artificial Intelligence and Digital Health are opening new perspectives for the use of Real World Data and for the real-time monitoring of clinical outcomes. In modern clinical research, these tools are applied in numerous phases of a clinical study's life cycle: experimental design, patient recruitment, study monitoring, and data analysis. Artificial Intelligence and Digital Health offer innovative solutions that reduce time and costs, improve data quality, and increase patient participation during clinical studies. However, many challenges and critical issues still remain, such as data reliability, the risk of algorithmic bias, and the difficulty of clinically validating the models used. Furthermore, integrating these technologies into regulatory processes and Good Clinical Practice (GCP) requires an adaptation of the professional skills of personnel involved in clinical studies and an evolution of the applied models, as there is a change in processes and organization, a change in mindset, and a change in culture, to which healthcare professionals must adapt; to do so, they must remain informed and properly trained. Another highly relevant aspect concerns the ethical and legal implications of Artificial Intelligence and Digital Health in clinical research. The cornerstone of clinical research lies in protecting patient integrity. The principle of patient centrality requires that every technology or procedure be implemented exclusively under the rigorous supervision of qualified personnel, ensuring that innovation never disregards safety standards and appropriate direct clinical monitoring. Supporting this, European and international regulations ensure the ethical and safe use of new technologies in modern clinical research. Through these considerations, this review aims to analyze the impact of Artificial Intelligence and Digital Health on modern clinical research, with particular attention to their main applications, operational challenges, and ethical aspects. The objective is to provide a comprehensive overview of the potential and limitations of these technologies, highlighting the central role of the researcher in their responsible and informed use, so that clinical research may also align itself with the digital era.

AI and digital health: theoretical framework and integration into clinical research

Artificial intelligence in healthcare: Definition and main technologies

History teaches us that every era is marked by significant change. Thanks to technological innovation, we are witnessing a continuous revolution that increasingly reduces human intervention while often enhancing process efficiency. Artificial Intelligence (AI) is a branch of computer science that develops systems capable of simulating human abilities such as learning, planning, analyzing, solving problems, and making decisions. Therefore, AI refers to a set of tools and technologies that enable a machine to perceive, act, think, analyze data, solve problems, and reach conclusions in a manner similar to human beings. The first definitions of AI date back to 1955, during a meeting organized by McCarthy, Minsky, Shannon, and Rochester [1], where participants discussed how every aspect of learning or any other characteristic of intelligence could, in principle, be described with such precision that a machine could be made to simulate it. During this meeting, AI was described as the science and engineering of making intelligent machines, particularly intelligent computer programs. A distinctive feature of AI is that it has experienced periods of great interest alternating with periods of limited attention; this fluctuation in interest has been referred to as the "AI winter." AI is closely associated with the concept of automation because it is capable of deducing information it already possesses, enabling it to generate and process an immense amount of data and information while performing complex repetitive operations. It is a decision-oriented system based on inputs (data), internal models, and outputs (actions or recommendations). In healthcare, AI plays a major role in risk stratification, prediction of clinical events, diagnostic image analysis, and decision support. Among the main technologies are:

- **Machine Learning (ML):** Machine learning is a subset of AI that focuses on creating systems capable of learning from data, identifying patterns, and making decisions with limited human intervention. This enables machines to learn automatically from data without being explicitly programmed for a specific task. The three main types of machine learning include supervised learning, unsupervised learning, and reinforcement learning. Supervised learning involves training the system on labeled data with the goal of learning to associate inputs with the correct outputs, making it useful for classification and prediction. Unsupervised learning involves analyzing unlabeled data to identify hidden structures or natural groupings, allowing information to be organized into meaningful subgroups or enabling dimensionality reduction. Reinforcement learning involves an algorithm learning by taking actions within an environment and receiving rewards or penalties, meaning the machine learns to perform a task through interaction with its environment based on a reward system. The integration of machine learning into clinical research enables the use of entire information systems to analyze large quantities of heterogeneous data from clinical, genomic, or digital-device sources and to identify complex patterns and relationships that are not immediately observable. Through machine learning, heterogeneous information, including textual clinical notes and diagnostic images, can be used to develop predictive models that are more comprehensive than conventional statistical approaches. In particular, machine learning is applied in the early stages of drug discovery, contributing to the identification of new therapeutic targets or the repurposing of existing molecules for other therapeutic uses. It also improves the efficiency of clinical trials by supporting patient selection and enabling more flexible and adaptive study models. Another important area is personalized medicine, where algorithms can predict individual responses to treatments, guiding more targeted therapeutic decisions.
- **Deep Learning (DL):** Deep learning is a subset of machine learning that is loosely inspired by the human brain and uses artificial neural networks composed of multiple layers to learn directly from data and classify, predict, or identify complex structures. Learning occurs through the process of error backpropagation, in which neural networks learn from their mistakes. Unlike traditional statistical models, which require the preliminary selection of relevant variables, deep neural networks can autonomously extract informative features from raw data. The use of deep learning in clinical research represents an advanced evolution of machine learning. Its application has proven particularly effective in analyzing complex and unstructured data through these deep neural networks. In clinical research, these techniques can process medical images, biological signals, and clinical

texts with a level of accuracy often comparable to, and sometimes exceeding, human performance. Deep learning has several applications in healthcare: one of the most significant is biomedical imaging, where it enables the identification of diagnostic patterns in X-rays, CT scans, and magnetic resonance imaging, supporting both research and clinical practice. It is also used in the analysis of genomic and molecular data, contributing to biomarker identification and the understanding of pathological mechanisms. Furthermore, in clinical trials, these models are useful for improving patient stratification and predicting therapeutic outcomes.

- **Natural Language Processing (NLP):** NLP is the branch of AI that enables the processing of natural language by teaching computers to understand, interpret, and generate human language. In healthcare, it refers to the technology that enables the automated processing of natural language, useful for analyzing medical records, reports, clinical notes, and scientific literature.
- In conclusion, machine learning is the brain of AI, deep learning is the nervous system, and natural language processing is the system's ability to communicate as naturally as possible. The current healthcare system is characterized by an exponential production of data. Topol [2] highlights how the increasing availability of clinical, genomic, and digital data exceeds human processing capacity, making it necessary to rely on Artificial Intelligence systems to integrate and interpret them effectively. However, it is essential to emphasize that Artificial Intelligence does not replace healthcare professionals but acts as a Decision Support System, facilitating the analysis of contexts characterized by a high degree of informational complexity. Therefore, AI has become a key tool in healthcare because it enables the identification of correlations and latent patterns within large volumes of heterogeneous clinical data.

Digital health: Definition and tools

The WHO defines Digital Health, abbreviated as DH, as a term rooted in the concept of eHealth, which encompasses the use of technology as a means of information and communication to support health and related fields. The term Digital Health generates confusion because the same concept is expressed through different terms; hence the need to establish a common language, which materialized in 2019 with the initiative of the Digital Medicine Society (DiMe) in proposing a taxonomy [3]. Digital Health refers to the set of technologies and systems used in prevention, diagnosis, and patient monitoring. Today, digital technology makes it possible to address the current healthcare challenge: chronic diseases. We are witnessing an aging population, with increasing numbers of people living with conditions from which they cannot recover and with which they must learn to coexist through appropriate medication use, greater awareness of healthy lifestyles, and continuous, responsible self-monitoring while playing an active role in their own care. Digitalization becomes a valuable aid for patients not only in self-management but also in reducing the distance in the physician-patient relationship, as patients can be monitored remotely by healthcare professionals. Digital Health includes tools such as electronic health records, wearable devices, mobile health applications (mHealth), telemedicine, sensors, remote monitoring systems, and clinical data analytics platforms; through these digital tools, clinical research moves from the hospital center directly into the patient's home. Digital Health has enabled the digitalization of healthcare systems, but to be effective, technological components must be integrated with organizational processes and the professional competencies of healthcare personnel. Electronic health records represent the core of this system because they are the digital tool used by physicians and healthcare professionals to document and share patient clinical information in real time, generating a large volume of heterogeneous data. The emergence of mobile health (mHealth), a subset of eHealth, has introduced a paradigm of continuous monitoring based on wearable devices and mobile applications capable of collecting physiological data in real time. Steinhubl., *et al.* [4] highlight how these tools make it possible to overcome the traditional episodic collection of clinical information, promoting a dynamic and longitudinal view of a patient's health status. Telemedicine is also part of DH, representing a new way of managing patients remotely by overcoming physician-patient distance. Through telemedicine, patients generate data and information through reports, sensors, and questionnaires, which physicians must analyze in order to provide prescriptions. Physicians must be capable of interpreting data to prescribe digital therapeutics, digital rehabilitation programs, and digital support interventions. In the field of clinical research, the most significant development is represented by decentralized clinical trials, in which the use of DH enables remote monitoring of participants, reducing the need for in-person visits and improving patient access

to clinical studies. In this context, data from wearable devices and digital sensors make it possible to define new endpoints that are more continuous and sensitive than conventional episodic measurements, allowing researchers to observe patients not only in their real-world environment but also, and above all, in real time. All of this allows healthcare professionals to pay closer attention to patient health status and increase patient participation in studies: patients perceive that they are being continuously and consistently monitored. This leads to greater clinical relevance of results and enables more targeted decisions based on a broader pool of information.

Integration between artificial intelligence and digital health

In healthcare, we are witnessing a true revolution thanks to the combination of Artificial Intelligence (AI) and Digital Health. The latter generates a vast volume of heterogeneous data that AI must analyze and interpret before transforming them into predictive or decision-support information. Their convergence entails a methodological shift, moving toward a more dynamic system that enables continuous monitoring, predictive capabilities, and personalized decision-making. Artificial Intelligence, as defined by Russell and Norvig [5], can be understood as a set of systems capable of acting rationally based on available information. However, these systems require a structured and continuous digital environment in order to operate effectively. This is precisely where Digital Health intervenes, providing the infrastructure for the production, storage, and interoperability of healthcare data. This perspective emphasizes that value does not lie in a single technology but in their functional integration. DH serves as a repository of data collected through electronic health records, wearable devices, telemonitoring systems, and mobile applications, producing Real-World Data characterized by high volume and variability. AI intervenes by transforming these data into predictive models, risk stratification systems, and decision-support tools. This process does not replace the traditional scientific method but amplifies it, because their integration becomes an important tool for detecting complex patterns that escape conventional linear models. In their article, Rajkomar., *et al*. [6] emphasize that machine learning is not merely a new technical tool but a fundamental technology for processing clinical data meaningfully, going beyond the capabilities of traditional statistical techniques and improving clinical decision-making, leading to prognosis, diagnosis, and personalized treatment. This is precisely the core of AI-DH integration: the ability to take digital data and transform them into actionable insights. Without Digital Health, Artificial Intelligence would lack suitable data; without Artificial Intelligence, Digital Health would remain mere digitalization. Therefore, without digitized data, the development of AI models for clinical research would not be possible. The real innovation, then, is not technological in the strict sense but lies in their integration, which concerns the interaction between data production and computational capacity. Kelly., *et al*. [7] focus precisely on this point, emphasizing that the availability of high-quality, standardized, and interoperable data is the essential prerequisite for any medical algorithm with clinical impact. Without a robust digital infrastructure, even the most sophisticated algorithms remain theoretical. Digital Health creates the infrastructure, Artificial Intelligence enhances its value, and effective ethical and regulatory governance ensures their implementation is suitable for the continuous and ongoing protection of patients.

Relevance for clinical research

The integration of AI and DH into clinical research is not merely a technological innovation but implies a deeper transformation, affecting the methodological infrastructure and reshaping research processes and procedures. Their use is therefore becoming increasingly relevant in contemporary clinical research because it directly impacts the methodological, organizational, and operational foundations of scientific experimentation. Conventional clinical research guarantees high validity through protocol standardization and rigorous control of variables, but it presents major structural limitations related to the representativeness of real-world populations, the generalizability of results, and the ability to observe longitudinal clinical dynamics. These limitations can be overcome through the integration of AI and DH. The use of AI during the design and conduct of clinical studies enables the simulation of predictive scenarios, refinement of endpoint definitions, and real-time monitoring of outcomes. DH tools—electronic health records, wearable devices, and telemonitoring systems—generate a constant flow of heterogeneous data which, thanks to AI, can be analyzed and transformed into valuable predictive and decision-support tools. Algorithmic analysis of large quantities of data makes it possible to identify patients with homogeneous characteristics, improving participant selection because subjects who are more aligned with the study objective are enrolled. This increases the likelihood of observing clinically

significant effects and reduces uncontrolled sample heterogeneity. More targeted selection contributes to improving the efficiency of clinical studies, reducing the number of participants required to obtain statistically robust results and accelerating clinical research timelines. This approach also strengthens the ethical dimension of research by limiting the exposure of unsuitable patients to potentially ineffective or risky treatments. Furthermore, the use of digital tools increases patient participation, considerably reducing dropout rates and thereby ensuring the success of clinical research, because fewer participants leave the study and the generated data become more abundant and reliable. All of this leads to an evolution toward more targeted, adaptive, and personalized trials aligned with the principles of precision medicine. This dynamic system contrasts with the rigidity of traditional experimental models, promoting greater methodological flexibility without compromising scientific rigor. The main advantage, however, lies in the reduction of time and costs, thereby contributing to the economic sustainability of clinical research. In conclusion, the integration of Artificial Intelligence and Digital Health in clinical research does not lie exclusively in the adoption of new technologies but in the transformation of the paradigm of scientific evidence generation. Modern clinical research is progressively moving toward a data-driven model, where data become a dynamic and continuous element of clinical research and where the integration of predictive analytics with the systematic collection of Real-World Data contributes to overcoming the traditional limitations of clinical studies, with the objective of generating scientific evidence that is more representative, personalized, and timely.

Applications of artificial intelligence and digital health in clinical research

Transformation of clinical trial design

The integration of Artificial Intelligence and Digital Health is producing a profound transformation in the design of clinical studies, modifying the methodological infrastructure of traditional clinical research. The classic randomized controlled trial model is based on predefined protocols with rigid inclusion and exclusion criteria, fixed variables and endpoints established in advance, and a relatively static structure that rarely undergoes modification during the course of the study. Data collection is characterized by high internal quality and rigorous control of variables; however, it is often poorly representative of real-world clinical practice because it derives from selected populations and artificial experimental conditions. The transformation of clinical trial design is characterized by the increasing adoption of flexible models, including adaptive clinical trials, which allow pre-specified modifications during the study based on emerging data, improving efficiency and the probability of success.[8] In this context, adaptive clinical trials play a central role because they permit predefined protocol modifications based on interim data analyses without compromising the statistical validity of the study. Furthermore, the introduction of predictive analytics tools based on machine learning algorithms enables the use of models trained to analyze large datasets originating from electronic health records, digital registries, and healthcare databases in order to simulate preliminary scenarios, optimize study design before implementation, and create more efficient protocols. Through these predictive models, it is possible to estimate patients' likelihood of responding to treatment, predict side effects, evaluate outcomes, and improve sample size calculations, thereby moving away from systems that study homogeneous patient populations and consequently reduce variability. Endpoint selection also undergoes transformation. In traditional models, endpoints are rigidly established before the study begins; in newer clinical studies, the introduction of more sophisticated analytical tools has enabled longitudinal analysis and continuous monitoring, with the ability to identify clinically relevant changes earlier and act promptly to avoid compromising the study. This creates a more comprehensive and realistic picture of the patient's clinical journey, improving the generalizability of evidence and supporting a more personalized approach to research. Consequently, there is a shift from a static model defined entirely a priori to a dynamic and adaptive model with planned modification possibilities, offering greater flexibility and allowing studies to adapt to emerging scientific evidence. This has made it possible to reduce drug development timelines, decrease the number of patients required for studies, ensure that the right drug is given to the right patient at the right time, reduce operational costs and achieve overall economic savings, and reduce errors during data analysis, thereby improving efficiency. This approach helps reduce the gap between experimental efficacy ("efficacy") and real-world effectiveness ("effectiveness"), improving the validity of results and bringing research closer to actual clinical practice. This transformation does not involve abandoning the methodological rigor of clinical studies; rather, it involves reshaping them through a balanced integration of algorithmic innovation and established principles

of randomization, control, and transparency. Moreover, the Digital Health and Artificial Intelligence tools used in clinical studies must be thoroughly validated, documented, and specifically integrated into the protocol, ensuring their effectiveness and reproducibility. The impact of integrating Artificial Intelligence and Digital Health into clinical trial design is reflected in the transition from a static and rigidly predefined model centered on experimental studies to a more flexible, adaptive, and patient-oriented model. This seeks a balance between innovation and methodological rigor, ensuring that design flexibility does not compromise scientific validity or patient safety.

Dimension	Conventional clinical trial design	Clinical trial design with Artificial Intelligence and Digital Health
Study structure	Rigid, sequential, fixed protocol	Flexible, adaptive (adaptive trials), dynamic protocol
Patient selection	Based on static criteria and homogeneous populations	Supported by predictive algorithms that identify specific subgroups and target
Data sources	Data collected in controlled settings (traditional RCTs)	Integration of Real-World Data (EHRs, wearables, mobile apps)
Data collection	Discontinuous, linked to clinical visits	Continuous, real-time through digital devices
Monitoring	On-site monitoring, costly and sporadic	Remote monitoring, automated and continuous
Study adaptability	Limited, modifications difficult during the trial	High, with modifications based on interim analyses
Patient involvement	Passive, limited to visits	Active, with digital tools and greater engagement
Timelines	Long and sequential	Reduced thanks to automation and real-time analysis
Costi	High|	Potentially reduced in the long term |
Generalizability |	Limited (selected populations)	Greater, thanks to the use of real-world data

Table 1: Comparison between conventional clinical trial design and clinical trial design based on the integration of AI and Digital Health.

The table highlights the transition from a traditional experimental model to an innovative paradigm based on the integration of Artificial Intelligence and Digital Health technologies (author's own elaboration based on: U.S. Food and Drug Administration, 2019, *Adaptive Designs for Clinical Trials of Drugs and Biologics: Guidance for Industry*. Silver Spring, MD: Food and Drug Administration; U.S. Food and Drug Administration, 2023, *Framework for FDA's Real-World Evidence Program*. Silver Spring, MD: Food and Drug Administration; European Medicines Agency, 2023, *Artificial Intelligence Workplan 2023–2028*. Amsterdam: European Medicines Agency; Sherman., *et al*. (2016). *Real-World Evidence—What Is It and What Can It Tell Us?* *New England Journal of Medicine*, 375(23), 2293–2297; Dorsey., *et al*. (2020). *Accelerating Clinical Trials Through Smart Technologies*. *NPJ Digital Medicine*, 3(79), 1–7).

Evolution of patient recruitment strategies

Patient recruitment is one of the most delicate parts of a clinical trial, to the extent that it represents one of the major causes of delay or failure of the entire process. Traditional recruitment strategies are based on physician referrals, print advertising, or local recommendations; however, they are often slow, costly, and ineffective in reaching a broad pool of potential participants, especially when eligibility criteria are complex or when populations are located far from the study center. Previously, clinical trials were centered on the study itself, and patient recruitment did not take into account a number of issues: the distance patients had to travel to reach the clinical center and the resulting costs in terms of time and money; problems related to limited mobility; and disruptions to patients' work continuity. Therefore, the challenge was not only recruiting patients and selecting suitable individuals to participate in the clinical trial, but above all maintaining the

appropriate number of participants throughout the entire study, since a large percentage dropped out during the course of the trial. Low participation was not attributable exclusively to these factors but was mainly linked to limited patient understanding and awareness of clinical research. The importance of ensuring adequate information and education about the study emerged, so that participants could be genuinely informed and actively engaged. Greater awareness is associated with better protocol adherence and lower dropout rates, significantly contributing to the overall success of the trial. This led to increased care and attention toward informed consent. The integration of AI and Digital Health (DH) has revolutionized the patient recruitment phase, as digital platforms make it possible to identify eligible patient profiles within healthcare system databases according to trial eligibility requirements by rapidly analyzing large volumes of data. These tools not only accelerate the identification of suitable candidates but also improve selection accuracy, reducing the time and resources required compared with traditional manual methods. In addition, digital approaches such as social media communication campaigns and targeted advertising are now used to extend the visibility of clinical studies beyond the clinical center and increase the number of eligible participants. These technologies also promote greater inclusiveness and participant diversity by overcoming geographical and social barriers. Consequently, the integration of digital health data and algorithmic intelligence tools is increasingly emerging as an effective strategy not only to accelerate patient recruitment but also, and above all, to enhance sample representativeness and improve the overall efficiency of clinical studies. Therefore, the evolution of patient recruitment strategies made possible by the integration of Digital Health and Artificial Intelligence not only improves the operational efficiency of trials but also constitutes a structural methodological element of modern clinical research, as it expands participation to a more heterogeneous population, enhances the selection of eligible individuals, and supports the conduct of more timely and generalizable studies. In conclusion, the true revolution in patient recruitment lies in the transition from a manual, site-centered system to an algorithmic, patient-centered system.

Remote monitoring and integration of real-world data

The monitoring phase is one of the areas of clinical research that has undergone the most significant transformation due to the introduction of digital technologies. This change is represented by decentralized clinical trials, in which patients are monitored remotely through continuous collection of clinical parameters. Patients are followed by healthcare professionals using wearable devices, biometric sensors, mobile applications, and telemedicine platforms, which allow real-time recording of physiological, behavioral, and environmental variables, constantly generating large volumes of data and reducing the distance between physician and patient. In contrast, the traditional approach involves data collection primarily during scheduled visits at the study site, with observations made at specific points in time and separated by intervals, limiting the ability to observe clinical dynamics occurring between visits due to the episodic nature of measurements. Izmailova., *et al*. [9] emphasize that remote monitoring not only reduces the need for in-person visits but also improves participant safety through the early identification of adverse event signals or worsening clinical conditions. Continuous monitoring therefore enables dynamic and ongoing assessment of a patient's health status, moving beyond the paradigm of episodic evaluation. The digitalization of the monitoring phase has led to significant advantages in terms of both cost reduction and time optimization. This is made possible by several factors: first, greater patient participation, as patients can be monitored comfortably from home while feeling more supported by healthcare professionals, resulting in lower study dropout rates. Furthermore, the continuous acquisition of real-time data provides more comprehensive and timely information, helping reduce operational workload and optimize resource utilization by personnel involved in clinical research. The combination of remote monitoring and mobile health technologies has enabled the systematic generation of Real-World Data (RWD), that is, data collected outside traditional experimental settings and directly from patients' everyday lives. Through wearable devices, mobile applications, and digital platforms, it is possible to continuously and passively collect a wide range of clinical, behavioral, and physiological information. This approach overcomes the limitations of episodic measurements typical of conventional clinical studies, offering a more complete, dynamic, and realistic representation of patient health status.

As a result, RWD not only enriches the quality and quantity of available evidence but also allows the observation of changes over time, improves intervention personalization, and supports clinical decision-making based on more extensive information obtained more rapidly. Once collected through digital tools and analyzed

using Artificial Intelligence, these data transform clinical practice into scientific evidence—Real-World Evidence (RWE)—that is, concrete evidence regarding treatment safety and effectiveness [10].

Artificial Intelligence plays a crucial role in this process because it enables the analysis of large volumes of heterogeneous, multimodal, and complex data. Through machine learning algorithms, longitudinal patterns, nonlinear correlations, therapeutic effectiveness, and toxicity can be identified, contributing to a more dynamic assessment of outcomes. From a regulatory perspective, interest in the use of RWD in clinical research has grown significantly. Corrigan-Curay., *et al*. [11] highlight that regulatory authorities are progressively recognizing the value of Real-World Data in supporting decisions related to new therapeutic indications, label extensions, and post-marketing evaluations. However, the integration of remote monitoring with Real-World Data also introduces important methodological challenges, including ensuring data quality and reliability, the need for standardized formats, interoperability among different information systems, and control of potential biases. In this context, it becomes essential to ensure the integrity, consistency, and accuracy of collected data in order to preserve their scientific validity. Data collected in uncontrolled environments may be subject to measurement errors, technological disruptions, or limited patient familiarity with digital tools. Therefore, the adoption of these technologies must be accompanied not only by adequate user training but also, and above all, by rigorous validation protocols and robust data governance strategies to ensure quality, reliability, and proper use.

In conclusion, remote monitoring and the integration of Real-World Data represent a shift from an episodic, site-centered model of data collection to a continuous and ongoing model that is primarily patient-centered. This transformation enables a broader, more comprehensive, and more dynamic view of clinical practice, narrowing the gap between clinical research and real-world practice and contributing to the generation of more representative and timely evidence.

Clinical data analysis and decision support

Clinical data analysis and decision support constitute one of the areas in which the integration of Artificial Intelligence and Digital Health has found the most concrete applications and generated the most profound transformations in modern clinical research. The digitalization of healthcare systems through electronic health records, administrative databases, electronic registries, and patient-reported questionnaires has produced an extremely large and complex volume of data. In addition to ensuring that data comply with ALCOA+ regulatory requirements—that is, that they are attributable, legible, contemporaneous, original, accurate, complete, consistent, enduring, and available over time—it is necessary to ensure that principles of data integrity and quality are always respected before proceeding with analysis and decision-making. Applications of Artificial Intelligence in clinical decision support systems have evolved from rigid rule-based models to advanced systems capable of learning from data through machine learning, deep learning, and natural language processing techniques. These systems are designed to provide healthcare professionals with recommendations, outcome predictions, risk identification, and personalized therapeutic suggestions based on information derived from a patient's profile. The goal of these systems is not to replace physicians but rather to facilitate decision-making processes through automated data analysis of information that would be too vast and complex to manage using traditional methods. This predictive approach enables the anticipation of complications, patient risk stratification, and optimization of personalized treatment plans.

In recent years, this has demonstrated that the integration of Artificial Intelligence into clinical data analysis and decision support systems represents one of the major developments in modern clinical research. It not only increases the efficiency and accuracy of medical decisions but also creates a close connection between large volumes of digital data and everyday clinical practice, generating advanced predictive tools and operational support in increasingly complex and personalized settings. Artificial Intelligence does not replace the scientific method or the art of medicine; rather, it enhances their interpretive and decision-making capabilities, provided that high data quality, rigorous prospective validation, and an appropriate regulatory framework are ensured.

Critical issues, methodological limitations, and operational challenges

Data quality, reliability, and integrity

Data quality, reliability, and integrity represent the indispensable foundation of clinical research, regardless of

whether traditional or innovative methods are used. These principles, ensured by Good Clinical Practice (GCP) standards and appropriate risk management, form the very basis of the success of any clinical study. Today, the healthcare environment generates large quantities of heterogeneous data from various sources, including electronic health records, administrative databases, health registries, Real-World Data sources, and remote monitoring devices. As a result, managing data and ensuring their quality and integrity has become increasingly complex. Artificial Intelligence enables the processing and analysis of large volumes of data; however, ensuring data quality and integrity is essential because the use of incomplete, inaccurate, or inconsistent data can generate misleading, biased, or erroneous predictions. AI also offers tools capable of detecting, preventing, and resolving integrity issues throughout the data lifecycle, thereby safeguarding data reliability. In recent years, growing international attention to the issue of data integrity has highlighted the importance of developing robust and reliable digital infrastructures. To this end, specific regulations are needed to strengthen compliance and guarantee high standards of data quality. According to the FDA, data integrity is the extent to which GxP data are complete, consistent, accurate, reliable, and truthful throughout their lifecycle while complying with ALCOA+ principles. The ALCOA+ principles describe the characteristics that data must possess:

- **Attributable:** All data created and collected must be attributable to a person or computer system, with collection and generation details recorded;
- **Legible:** All data must be permanent and easily accessible throughout their lifecycle;
- **Contemporaneous:** Data must be recorded at the same time the activity is performed, avoiding errors and approximations that could produce imprecise information;
- **Original:** Data should be original rather than copies or transcriptions;
- **Accurate:** Data must reflect what actually occurred, and any modifications must be signed and justified;
- **Complete:** Omission or concealment is prohibited, since what is not documented does not exist;
- **Consistent:** Data must be consistent with their intended use;
- **Enduring:** Data must be preserved appropriately over time;
- **Available:** Data must be readily accessible.

Therefore, ensuring that data comply with ALCOA+ requirements helps prevent duplicates, labeling errors, and omissions that can generate systemic biases in AI models, compromising not only predictive reliability but also appropriate decision support. Effective data governance requires interoperability standards and security protocols to ensure that clinical information remains consistent, traceable, and reproducible over time. Lack of interoperability among healthcare information systems can lead to data fragmentation and integration difficulties, hindering analytical quality and continuity of information across different sources. A well-designed data infrastructure that adopts international interoperability standards is a key element in ensuring that data retain their value and can be effectively used for clinical and research purposes. Data reliability is closely linked to reproducibility and to the ability of models to generate consistent results when applied to different datasets or multiple contexts. The application of predictive models based on Real-World Data highlights that a model performing well in a specific context is not automatically generalizable. Reliability depends on the model's ability to maintain performance, consistency, and stability when applied to different populations or clinical environments. Therefore, beyond the intrinsic quality of the data, it is necessary for the systems used to be validated through evidence demonstrating their effectiveness, thereby ensuring that the conclusions drawn are reliable. The issue of data quality and reliability is not merely a technical matter; it directly affects patient safety and ethical and regulatory compliance. Inaccurate or uncontrolled data can lead not only to incorrect clinical decisions but also to a loss of trust in digital systems. For this reason, data management processes must be accompanied by robust governance policies, methodological transparency, and auditability, ensuring that AI operates on a solid, shared, and regulated informational foundation.

Algorithmic bias and model opacity

The introduction of Artificial Intelligence into clinical research represents an important opportunity because it enables the optimization and increased efficiency of various phases of the clinical process. However, it is equally important to consider the associated structural risks, particularly those related to algorithmic bias and the limited transparency of decision-making models. Algorithmic bias is a systematic and repeated error that generates distorted, unfair, or discriminatory outcomes. It can produce unjustifiably favorable or unfavorable results for specific groups

of individuals and amplify pre-existing inequalities in data or in AI model training processes. This occurs when training data are unbalanced and fail to accurately and comprehensively represent reality. Bias can arise from various factors, including imbalances in training data, labeling errors, non-neutral methodological choices, and implicit systemic prejudices embedded within clinical datasets. Without rigorous oversight, such models may replicate and amplify existing inequalities, worsening healthcare equity and patient safety. In healthcare settings, algorithmic bias can have clinically significant and potentially harmful consequences. Biases may be intentional, resulting from design choices, or unintentional, most often stemming from unbalanced data. Many machine learning algorithms, particularly those based on complex deep learning architectures, are characterized by limited transparency in their decision-making processes. Unlike traditional statistical models, where the contribution of individual variables can be explicitly reconstructed, advanced models operate through multiple layers of data transformation, making it difficult to understand how specific characteristics influence the final output. This situation implies that, in many cases, even the model developers may be unable to provide a clear and interpretable explanation of the logical pathway leading to a prediction. This phenomenon is known as model opacity and is often described as the "black box" problem. In the context of clinical research, this lack of interpretability represents a significant challenge. It not only limits the possibility of rigorously validating results but may also reduce clinicians' trust and hinder the adoption of these tools in practice. Furthermore, the inability to fully understand a model's internal functioning makes it more difficult to identify systematic errors or biases, with potential implications for patient safety and the reliability of clinical decisions. Such opacity undermines not only healthcare professionals' confidence in the technology but also raises important ethical and legal questions concerning clinical responsibility, regulatory compliance, and the principle of transparency in evidence-based medicine.

Moreover, issues related to bias and model opacity may lead to the propagation of social prejudices within clinical decision-making, amplifying existing racial, gender, or socioeconomic stereotypes. Unlike traditional predictive models, which primarily operate on structured data, generative AI models are trained on enormous quantities of text from heterogeneous sources, including scientific publications, clinical databases, guidelines, forums, and web content. While this broad training base is a strength in terms of generative capabilities, it also carries the risk of incorporating and reproducing stereotypes, conceptual errors, and inequalities already present in the original textual data. One of the most common issues is the generation of inaccurate clinical recommendations for specific population groups. If information indicating that certain pathological conditions have historically been more prevalent in particular demographic groups is included in the training data, the model may develop statistical associations that influence its generated responses.

This phenomenon is driven by the quality and distribution of information associated with variables such as gender, ethnicity, or patient age, creating the risk of implicitly reinforcing systemic biases already present in the literature and clinical practice. To avoid this, greater human oversight, more rigorous validation, and continuous monitoring of generated responses are necessary.

Several strategies have been proposed to mitigate bias and opacity, including the adoption of Explainable AI (XAI) techniques that provide a more transparent understanding of algorithmic decision-making processes, the use of representative and balanced datasets, and the implementation of fairness metrics to evaluate model performance across different subgroups. Advanced methodological approaches, such as predictive auditing and statistical fairness criteria, are increasingly being recognized as necessary tools for promoting more equitable and reliable systems capable of supporting responsible clinical decisions aligned with the principles of healthcare equity and justice. Therefore, to improve fairness and correctness within healthcare processes and reduce existing inequalities, rigorous controls and continuous oversight by qualified professionals must be implemented throughout clinical research. It is essential to recognize that Artificial Intelligence is a valuable tool only when supported by the active involvement of competent human experts.

Reproducibility and clinical validation

The spread of Digital Health and Artificial Intelligence introduces new methodological opportunities, but also raises critical issues regarding reproducibility and clinical validation. Reproducibility is one of the fundamental principles of clinical research and is also an essential requirement for the application of Artificial Intelligence. While in algorithmic models the problem

mainly concerns the generalizability of predictive performance, in the field of digital health the issue extends to the entire technological ecosystem: wearable devices, remote monitoring platforms, telemedicine systems, and mobile applications generate data in uncontrolled settings characterized by high environmental and behavioral variability. The lack of standardization in devices, data collection methods, and data formats represents a significant obstacle; biometric sensors from different manufacturers may measure the same physiological parameter with different levels of accuracy and sampling frequency. This heterogeneity makes it difficult to compare results across studies and compromises the possibility of replicating evidence generated in different settings. Clinical validation of digital devices is critical because many remote monitoring tools are initially developed for wellness purposes and only later adapted to clinical contexts. In these cases, technical validation—which must demonstrate measurement accuracy against a reference standard—does not necessarily coincide with clinical validation. Therefore, the mere ability to measure a parameter does not guarantee that the information is clinically relevant or can be used in decision-making processes. Validation of electronic and technological tools used in clinical research consists of generating documented evidence capable of guaranteeing, with a high degree of reliability, that such tools operate consistently, accurately, and in accordance with the function for which they were designed throughout their entire lifecycle of use. In the context of decentralized trials, for example, validation assumes an even greater importance, as remote monitoring enables continuous data collection in home environments but introduces variables that are difficult to control, such as adherence to device use, user errors, environmental conditions, and interruptions in data transmission. All these factors can affect the quality and completeness of datasets, impacting the robustness of scientific conclusions. Similarly, regarding the use of Artificial Intelligence in modern clinical research, algorithmic models may demonstrate high performance on a specific dataset, but this does not automatically guarantee that such results are reproducible in different contexts, with different populations, or in real-world clinical environments. Thus, in healthcare, the reproducibility problem frequently manifests itself in the difficulty of transferring a model developed in a specific hospital setting to other contexts, since differences in clinical protocols, data recording methods, and population demographics can significantly affect algorithm performance. Rajkomar., *et al*. [6] highlight how machine learning models applied to electronic health records can achieve promising results but require validation on external datasets to demonstrate robustness and generalizability; similarly, Wiens., *et al*. [12] emphasize the importance of cross-validation strategies, testing on independent datasets, and interdisciplinary collaboration between clinicians and data scientists to mitigate this risk. This demonstrates that validation cannot be considered a one-time event but rather a structured and continuous process consisting of several phases: the planning phase, during which data are collected, risks assessed, and testing strategies defined; followed by the test execution phase, in which planned verifications are implemented and any deviations managed; subsequently, a final report is prepared and the system released; finally, the process is completed with a post-release maintenance phase, which is essential to ensure reliable performance and compliance with required standards over time. In this dynamic context, an additional element of complexity is represented by adaptive computational models, namely systems capable of updating themselves over time through the integration of new data. While such models offer a significant advantage in terms of continuous performance improvement, they also raise important concerns regarding system stability and change traceability. For this reason, validation necessarily becomes a continuous and ongoing process. Every update must not only be carefully documented but also subjected to rigorous verification and validation procedures to ensure that introduced modifications do not compromise the reliability, safety, and consistency of performance over time. Reproducibility and clinical validation represent the crucial bridge between technological innovation and the responsible application of Artificial Intelligence and Digital Health in healthcare practice. Digital tools and computerized systems supporting GXP processes must be validated according to an approach proportionate to the level of risk they pose to patient safety, product quality, and data integrity. Clinical validation of digital tools requires a multi-level approach integrating technical validation, clinical evidence, and regulatory compliance. In particular, in the context of Digital Health and Artificial Intelligence-based systems, this process is dynamic and continuous, requiring not only the initial demonstration of safety and effectiveness but also post-market monitoring aimed at ensuring reliable performance over time and in Real-World settings. The challenge is not to demonstrate that an algorithm works in a laboratory, but that it produces concrete, safe, and generalizable benefits in the real world. Otherwise, the risk is generating models that are difficult to replicate and undermining

scientific confidence in the results produced. To generate robust, comparable, and generalizable scientific evidence, it is necessary to ensure that the rapid spread of technologies does not outpace the ability to rigorously verify their effectiveness and safety.

Integration into good clinical practice (GCP) and regulatory processes

The use of Artificial Intelligence and Digital Health is radically transforming clinical research by improving precision and accelerating processes, thereby creating the need to adapt Good Clinical Practice (GCP) principles to new technological tools. The entry into force, on July 25, 2025, of version R3 of the ICH E6 guideline (Guideline for Good Clinical Practice) represents a significant update regarding the use of computerized systems and technological tools in clinical research. The aim of GCP R3 is to promote the conscious use of digital technologies by recognizing their advantages and opportunities for clinical research while ensuring data integrity, traceability, and protection, and above all safeguarding participants. One of the most relevant aspects introduced by the new ICH E6(R3) is the inclusion of a chapter specifically dedicated to Data Governance and computerized systems: Chapter 4. This represents a significant advancement compared to the previous version, ICH E6(R2), in which validation was addressed more narrowly within a brief paragraph, section 5.5.3, concerning sponsor responsibilities. The implementation of ICH E6(R3) marks a turning point for data management through technology in clinical trials, and the inclusion of Chapter 4, entirely dedicated to computerized system validation, reflects the growing importance of technology in clinical research and strengthens the role of Data Governance as an essential element for quality and regulatory compliance. One critical issue is the plasticity of machine learning systems, as they can rapidly evolve over time through updates or the learning of new data, making the application of GCP clinical validation schemes more difficult. In the case of Artificial Intelligence algorithms, system performance may vary depending on the datasets used, the clinical context, and the characteristics of the target population. Therefore, it becomes necessary to develop more flexible clinical validation methodologies that include continuous performance monitoring, algorithm audits, and periodic updates of scientific evidence. GCP R3 excludes one-size-fits-all solutions, since the requirements of a computerized system must be calibrated according to the characteristics and complexity of the individual clinical trial. One of the most innovative elements introduced by ICH E6(R3) is the adoption of a risk-based approach, which modulates the intensity of controls and verifications according to the level of risk. In this perspective, resources are primarily concentrated on study aspects that may have the greatest impact on participant safety and data reliability. This not only makes processes more efficient by reducing unnecessary bureaucratic burdens but also strengthens the overall quality of the trial by focusing attention on truly critical elements. Competent authorities are developing specific guidelines for evaluating Artificial Intelligence and Digital Health systems used in healthcare. The concept of Software as a Medical Device (SaMD) is becoming increasingly established, identifying software and algorithms used for diagnostic or therapeutic purposes. For such tools, regulatory authorities require evidence of safety, effectiveness, and reliability comparable to that required for other medical devices, but with validation procedures adapted to the algorithmic nature of the systems. For these products to become medical devices, adequate evidence of effectiveness demonstrating their validity is essential. GCP emphasizes the importance of accurate documentation of all phases of the clinical study, ensuring traceability and transparency of the data used for training, modifications made to the model, and the adopted performance metrics. This approach is essential to enable independent verification, ensure reproducibility of results, and strengthen the confidence of healthcare professionals and patients in digital technologies. Furthermore, the integration of AI and DH into GCP and regulatory processes also implies a revision of the competencies required of researchers and oversight bodies, along with a clear definition of roles. The evaluation of algorithmic systems requires interdisciplinary expertise combining clinical, statistical, and computational knowledge. Consequently, many regulatory bodies are promoting new governance models involving experts in data science, medicine, and bioethics, with the goal of ensuring a more comprehensive assessment of Digital Health technologies and Artificial Intelligence systems. Therefore, the key points of Chapter 4 of ICH E6(R3) can be summarized as the need to define clear procedures for the use of technological systems, ensure that personnel possess adequate competencies, implement effective measures for data protection and participant privacy, ensure validation and proper management of system malfunctions, and record and track data in compliance with ALCOA+ principles. In conclusion, adapting GCP to technological innovation in clinical research encourages its use, provided that rigorous standards of data quality, participant safety, and accountability of involved parties are guaranteed.

Ethical, legal, and regulatory aspects

Privacy, data protection, and informed consent

The use of Artificial Intelligence technologies and Digital Health tools involves the processing of large amounts of data, often originating from heterogeneous sources such as electronic health records, wearable devices, mobile applications, and remote monitoring systems. The generated data fall within the category of particularly sensitive personal data, as they concern individuals' health status and may reveal extremely delicate information about their private lives. Therefore, privacy protection and personal data security represent a central dimension in the development and implementation of technological systems in healthcare. Since 2018, EU Regulation 2016/679 has been applied in all Member States of the European Union under the name General Data Protection Regulation (GDPR). However, the Regulation does not specifically govern the processing of personal data or the methods for obtaining consent for participation in clinical trials, referring instead to the applicable legislation in the relevant field (Recitals 156 and 161 of the GDPR). Clinical trial regulation in Europe is currently governed by Regulation (EU) No. 536/2014, which entered into force on January 31, 2022, repealing Directive 2001/20/EC. Among the most significant references on this topic are also the draft guidelines published by the EMA on April 7, 2022, concerning transparency of information uploaded to the Clinical Trial Information System (CTIS), entitled "Guidance document on how to approach the protection of personal data and commercially confidential information in documents uploaded and published in the Clinical Trial Information System ("Guidance")." The objective is to facilitate the free exchange of data within the European Union while strengthening personal data protection. One of the primary tools for ensuring ethical and compliant processing of health data is informed consent. It constitutes a cornerstone principle of clinical research, ensuring that patients are adequately informed about the purposes, risks, and benefits of the procedures to which they are subjected. However, in the context of digital technologies and Artificial Intelligence, the concept of informed consent assumes new complexities because patients may not be fully aware of how their data are collected, processed, and reused for research purposes or algorithm training. The GDPR establishes the principles of lawfulness, fairness, and transparency: data subjects are informed in simple, understandable, and effective language about how their data are generated and processed. To address these issues, the importance of more dynamic and transparent informed consent models is emphasized, providing greater clarity regarding data use, adequate cybersecurity measures, and data pseudonymization, with the objective of reducing the risk of patient identification and unauthorized access to health information. In particular, according to the Guidelines, clinical trial sponsors are required to ensure the coding of participants' personal data, including special-category data. In this regard, the GDPR introduced the measure of pseudonymization, defined as the processing of personal data in such a way that they can no longer be attributed to a specific individual without the use of additional information, which must be kept separately and protected through appropriate technical and organizational measures to prevent direct or indirect identification of the individual (Article 4(5), GDPR). Therefore, clinical investigators are responsible for pseudonymizing participants' medical-clinical data through the assignment of numerical codes that allow unique identification within the study without using directly identifying information such as names, addresses, or personal identification numbers. In this context, privacy protection and responsible data management are not merely regulatory obligations but essential prerequisites for the adoption of digital technologies in healthcare. Without adequate guarantees of security, transparency, and control over personal data, the dissemination of such systems risks encountering resistance from both patients and healthcare professionals. Consequently, personnel involved in clinical trial centers, particularly during preliminary interviews aimed at obtaining informed consent, must be adequately trained, including on relevant aspects of personal data protection legislation, so that they can provide clear, complete, and understandable information regarding data processing. To this end, study sponsors, when selecting centers where trials will be conducted, are required to verify the adequacy of personnel in managing such activities and provide specific training interventions if necessary. Furthermore, the European Regulation establishes the principle of purpose limitation, according to which personal data must be collected for specified, explicit, and legitimate purposes and subsequently processed in ways that are not incompatible with those purposes. Therefore, if data are used for dissemination or demonstration purposes, this must be clearly stated within the informed consent process. Data retention, on the other hand, remains regulated in a somewhat unclear and fragmented manner across Member States,

but it generally seeks to balance individual rights with the collective interest pursued through scientific research. The legislation requires specifying, consistently with the purpose of processing, how long collected data will be retained in a form that allows identification of individuals. The primary concern remains the integrity and confidentiality of data, ensuring appropriate measures to protect against risks, loss, and destruction of data, since in clinical research data quality is indispensable for demonstrating the reliability of obtained results. The growing technological development in clinical research offers the opportunity to improve attention and care in the organization and use of data while respecting the fundamental right to personal data protection of study participants. Therefore, the development of effective policies and tools for data protection is a key element in promoting the ethical, safe, and sustainable use of Artificial Intelligence and digital tools in clinical research.

Clinical responsibility and automated decision-making

The introduction of technology into clinical research raises important questions regarding professional responsibility and the management of automated decisions. Many Digital Health tools are designed to support or automate certain stages of medical decision-making, including diagnosis, rapid identification of risks, and assistance in selecting therapeutic treatments. Although these technologies can provide valuable support in various areas, they also introduce new complexities regarding the attribution of responsibility in cases of error or adverse clinical outcomes. When decisions are supported by technology, the decision-making process becomes more complex because it involves multiple actors: healthcare professionals who use the system, competent authorities that regulate it, and developers who design it. This situation makes it difficult to clearly determine who is responsible when an automated system produces an incorrect or misleading clinical recommendation [13]. Another critical issue concerns the degree of decision-making autonomy attributed to AI systems. Although such systems are designed to support clinical decision-making through the analysis of large quantities of data and the formulation of recommendations, they cannot replace the professional judgment of healthcare personnel. For this reason, their use must always occur under the supervision of competent professionals who retain ultimate responsibility for clinical decisions. Therefore, healthcare professionals must use technology as a support tool while maintaining a critical perspective and evaluating algorithmic recommendations in light of the specific context. Even as data complexity and the predictive capabilities of machine learning models increase, excessive reliance on system recommendations without appropriate critical oversight must be avoided. For this reason, the importance of maintaining a "human-in-the-loop" model is emphasized, whereby Artificial Intelligence technologies support decision-making processes but do not replace physicians. This approach combines the analytical capabilities of algorithms with the clinical judgment and experience of healthcare professionals, reducing the risk of errors arising from inappropriate use of automated systems [14]. Furthermore, it is essential to ensure an adequate level of transparency and explainability of algorithms so that clinicians can understand the logic underlying system-generated recommendations and assess their reliability. From a regulatory perspective, international institutions are progressively developing guidelines governing the use of Artificial Intelligence in clinical research. The WHO has highlighted the need to clearly define accountability mechanisms that establish roles and responsibilities among developers, technology providers, and healthcare professionals [15]. In this context, integrating AI and DH systems into existing regulatory frameworks concerning medical devices and clinical practices is essential to ensure that these tools are appropriately validated, monitored, and used safely. Therefore, the adoption of automated decision-making systems in clinical research requires a balance between technological innovation and professional responsibility. Artificial Intelligence and Digital Health technologies can provide important support for improving decision quality, but their use must occur within clear governance models that ensure transparency, human oversight, and appropriate attribution of responsibilities [16]. Only through proper integration of human expertise and technological tools will it be possible to harness the potential of technological evolution while maintaining high standards of patient safety and protection.

European regulatory framework: GDPR and AI Act

The use of Artificial Intelligence in healthcare requires a regulatory framework capable of balancing technological innovation, rights protection, and patient safety. In Europe, this balance is primarily pursued through two regulatory instruments: the General Data Protection Regulation (GDPR) and the more recent Artificial Intelligence Act (AI Act). These instruments represent the pillars of European regulation of digital technologies, aiming to ensure that the development and integration of Artificial Intelligence and Digital Health occur in accordance with ethical

principles, personal data protection, and accountability in the use of algorithms. The GDPR constitutes the primary European legal reference for personal data protection. The Regulation establishes a series of fundamental principles that must be respected in data processing, including lawfulness, fairness, transparency, data minimization, and purpose limitation [17]. In healthcare, these principles are particularly relevant because health-related data fall among sensitive personal data categories, for which the Regulation provides specific safeguards and processing conditions. In relation to Artificial Intelligence, the GDPR also imposes obligations concerning data security, data protection impact assessments, and patients' right to obtain information about how their data are used. A particularly significant aspect of the GDPR concerns automated decision-making, regulated by Article 22, which grants individuals the right not to be subject to decisions based solely on automated processing when such decisions produce legal effects or significantly affect them. In healthcare, this principle reinforces the need to maintain human oversight in decision-making processes supported by Artificial Intelligence systems, ensuring that decisions are not entrusted exclusively to algorithms. Alongside the GDPR, the European Union has developed specific legislation for the regulation of Artificial Intelligence: the AI Act, approved in 2024. It introduces a risk-based approach, classifying AI systems into different categories according to their potential impact on fundamental rights and individual safety. In particular, many AI applications in healthcare fall into the high-risk category because they can influence diagnostic, therapeutic, or organizational decisions with direct consequences for patient health. For systems classified as high-risk, the AI Act establishes stringent requirements, including the quality and governance of training data, technical documentation of systems, transparency toward users, and the implementation of appropriate human oversight mechanisms. Furthermore, the Regulation establishes obligations for continuous monitoring and evaluation of AI system performance to ensure that these technologies remain safe and effective even after market placement. In healthcare, the interaction between the GDPR and the AI Act is particularly important because while the GDPR primarily focuses on personal data protection and individual rights, the AI Act more directly regulates the technical characteristics and safety requirements of algorithmic systems. Together, they contribute to defining a comprehensive regulatory framework that promotes responsible AI development and guarantees high standards of patient protection. Therefore, the European regulatory framework represented by the GDPR and the AI Act constitutes a fundamental element in guiding the safe adoption of new technological systems, because only through an approach based on the protection of fundamental rights, transparency, and accountability of the actors involved can the potential of technology be harnessed without compromising the safety and dignity of individuals.

Integration into international regulatory frameworks (EMA/FDA)

The integration of Artificial Intelligence technologies and Digital Health tools into clinical research requires appropriate inclusion within the main international regulatory frameworks. In this context, the EMA in Europe and the FDA in the United States play a central role in the evaluation and authorization of drugs, medical devices, and digital technologies used in healthcare. They converge on a fundamental principle: it is no longer sufficient to demonstrate that Artificial Intelligence works; it is necessary to ensure that it operates reliably, traceably, and under appropriate governance. The guidelines produced do not introduce new legally binding obligations but define shared operational standards that concretely affect all phases of drug development, from clinical trials to post-marketing monitoring. This implies rethinking the use of Artificial Intelligence not merely as technological experimentation but as a structural element of regulated processes. Consequently, the evolution of digital technologies, including mobile health applications, remote monitoring systems, and AI algorithms for clinical decision support, has made it necessary to adapt traditional regulatory models to new types of software- and data-based devices. In the European context, the EMA has progressively developed strategies to promote the responsible use of Artificial Intelligence and digital technologies in research processes, drug development, and pharmacovigilance. In particular, it has highlighted the potential of AI in clinical trial design, patient recruitment, analysis of Real-World Data, and support for decision-making, making research processes shorter and more efficient. However, the adoption of these technologies requires rigorous standards of data quality, algorithm transparency, and scientific validation to ensure that the generated results are reliable and usable in regulatory processes. The EMA has therefore launched specific initiatives to explore the use of Artificial Intelligence in the pharmaceutical sector while promoting a cautious approach

based on robust scientific evidence [18]. Similarly, in the United States, the FDA has developed in recent years a specific regulatory framework for Digital Health products and Artificial Intelligence systems applied to medicine. Within this context lies the concept of Software as a Medical Device (SaMD), [19] which includes software and digital applications used for diagnostic, therapeutic, or clinical decision-support purposes. Examples include remote patient monitoring applications and machine learning-based medical image analysis algorithms. The FDA requires such technologies to demonstrate adequate levels of safety, effectiveness, and reliability before being placed on the market through validation processes similar to those required for other medical devices. Therefore, they must be validated before being used in clinical research. A particularly innovative aspect of the FDA's regulatory framework concerns the management of adaptive machine learning systems that can be updated over time based on new data. Since these systems may alter their performance after initial approval, the FDA has proposed regulatory models involving continuous monitoring, management of algorithm modifications, and performance oversight over time. This approach aims to ensure that Artificial Intelligence-based technologies maintain adequate standards of safety and effectiveness even after implementation in clinical practice. A central element in integrating Artificial Intelligence and Digital Health into international regulatory frameworks concerns the clinical validation of digital technologies. Regulatory authorities require scientific evidence demonstrating not only the technical accuracy of algorithms but also their real impact on healthcare processes and patient health outcomes. The use of real-world evidence in clinical studies involving digital tools and health data monitoring systems is assuming an increasingly important role in regulatory evaluation of technological innovations. However, the main element introduced by EMA-FDA principles is the concept of context of use, in which Artificial Intelligence is no longer evaluated in the abstract but according to the impact it produces within processes. The same technology can have profoundly different implications depending on the field and manner in which it is applied. Therefore, the focus shifts from the model itself to the context in which it operates, which must always be based on safety and reliability. The growing spread of Artificial Intelligence and Digital Health technologies necessitates greater international cooperation among regulatory authorities. Organizations such as the International Council for Harmonisation (ICH) promote the harmonization of guidelines and regulatory standards among Europe, the United States, and other international contexts. This coordination is particularly important for digital technologies, which are often developed and used globally and therefore require common standards of safety, quality, and transparency. The integration of Artificial Intelligence and Digital Health into EMA and FDA regulatory frameworks represents a crucial challenge for the governance of technological innovation in clinical research. Increasingly flexible approaches are being adopted, based on continuous monitoring of system performance, the quality of data used, and the clinical validation of digital applications. The primary objective is to ensure that the integration of AI and DH into clinical research occurs safely and transparently, thereby fostering innovation to improve the effectiveness and timelines of research. In this perspective, the goal is not to slow innovation but to make it robust, scalable, and sustainable over time.

Future perspectives and ethical recommendations

The development of Artificial Intelligence and Digital Health technologies has revolutionized modern clinical research by facilitating the identification of suitable patient populations, enabling remote patient monitoring, and generating large amounts of heterogeneous data that can be processed rapidly to support final decisions. However, the growing adoption of these technologies also raises important ethical, regulatory, and organizational challenges that require careful reflection on how such tools will be integrated into clinical practice and future healthcare systems. There is a clear need to develop strategies that ensure not only technically effective but also ethically responsible and socially sustainable use. One of the main future objectives is strengthening the transparency and explainability of algorithms used in Artificial Intelligence systems. Many machine learning models are characterized by a high level of complexity that makes it difficult, even for their designers, to explain the algorithm's decision-making process. For this reason, the importance of developing Explainable AI models is emphasized, capable of providing understandable explanations to healthcare professionals and patients regarding recommendations generated by algorithmic systems [20]. Greater transparency contributes to improving trust in technological evolution and encourages more informed use of AI and DH tools in modern clinical research. Another issue requiring attention in the use of Artificial Intelligence in healthcare concerns the promotion of equity and the prevention of discrimination. Machine learning algorithms may reflect or amplify biases present in the datasets used for training. This risk

is particularly relevant in healthcare, where algorithmic decisions based on non-representative data could generate disparities in study access or in the quality of generated data. Therefore, it becomes essential to ensure that datasets used for developing Artificial Intelligence systems are as complete, diverse, and representative as possible of different patient populations. Another relevant aspect concerns the need to develop governance models that accompany the entire lifecycle of Artificial Intelligence technologies and Digital Health tools, from design through implementation in clinical practice and throughout their operational lifespan. It is crucial to ensure continuous monitoring and control systems to verify that algorithmic systems maintain adequate levels of safety, reliability, and data quality over time. Competent authorities promote continuous evaluation procedures, algorithm audits, and periodic reviews of digital tools while maintaining high standards of safety, accountability, and transparency. Looking to the future, it will be important to promote greater integration among technological, clinical, and ethical competencies. The development and dissemination of effective Artificial Intelligence and Digital Health systems require collaboration among healthcare professionals, engineers, data scientists, legal experts, and ethicists. Only through an interdisciplinary approach will it be possible to design technologies that genuinely respond to patient needs and can be integrated safely and effectively into healthcare processes. The future integration of Artificial Intelligence and Digital Health into clinical research will depend on the ability to balance technological innovation with ethical responsibility. This must be supported by the development of appropriate regulations, the promotion of shared ethical standards, and strengthened collaboration among different actors within the healthcare system. This represents the key element for ensuring that these technologies truly contribute to improving clinical research.

Conclusion

The analysis developed in this work has highlighted how Artificial Intelligence and Digital Health technologies are progressively transforming the world of modern clinical research across its various stages. The adoption of tools based on machine learning algorithms, remote monitoring systems, and digital platforms for data management offers significant opportunities to improve the efficiency of clinical studies and the quality of final decisions. At the same time, these innovations introduce new challenges related to data quality, algorithm transparency, professional accountability, and the protection of patients' rights. It was highlighted that the development and integration of Artificial Intelligence and Digital Health in clinical research require a multidimensional approach capable of integrating technological, clinical, ethical, and legal expertise. A central role in this process is played by the regulatory and legal framework, which represents a fundamental tool for ensuring that technological innovation takes place in accordance with the principles of safety, transparency, and accountability. This work therefore emphasized that the value of Artificial Intelligence in clinical research does not lie exclusively in its technical capabilities, but in its integration within a governed, transparent, and quality-oriented system; without adequate governance, in fact, there is a risk of generating results that are difficult to interpret or not fully reliable, thereby compromising the decision-making process. In the future, the success of Artificial Intelligence and Digital Health technologies will depend on the ability to develop technological innovation models that place patient safety, data quality, and accountability in their use at the center. The main challenge will be to develop increasingly advanced models in clinical research based on a balance between innovation and control, between flexibility and standardization. Therefore, the introduction of these innovative technological tools must be able to support and enhance the decision-making capabilities of healthcare professionals, facilitating clinical research and accelerating the time required to bring drugs and medical devices to market. In conclusion, the integration of Artificial Intelligence and digital technologies into clinical research represents one of the most significant transformations, but their use must be promoted in a manner that is scientifically sound, ethically responsible, and regulatorily appropriate. Only through a balance between technological innovation, the protection of patients' rights, and professional accountability will it be possible to fully harness the potential of Artificial Intelligence and digital tools for the future of research.

Bibliography

1. McCarthy J., *et al*. "A proposal for the Dartmouth summer research project on artificial intelligence". *AI Magazine* 27.4 (2006): 12-14.

2. Topol EJ. "High-performance medicine: The convergence of human and artificial intelligence". *Nature Medicine* 25.1 (2019): 44-56.

3. Goldsack JC., *et al*. "Digital health, digital medicine, digital therapeutics (DTx): What's the difference?" *Digital Medicine Society (DiMe)* (2019).

4. Steinhubl S R., *et al*. "The emerging field of mobile health". *Science Translational Medicine* 7 (283 (2015): 283rv3.

5. Russell SJ and Norvig P. "Artificial intelligence: A modern approach". (4th ed.,) (2021): 15-25.

6. Rajkomar A., *et al*. "Machine learning in medicine". *The New England Journal of Medicine* 380.14 (2019): 1347-1358.

7. Kelly C J., *et al*. "Key challenges for delivering clinical impact with artificial intelligence". *BMC Medicine* 17.1 (2019): 195.

8. U.S. Food and Drug Administration (FDA). "Adaptive Designs for Clinical Trials of Drugs and Biologics: Guidance for Industry". Silver Spring, MD: U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER Center for Biologics Evaluation and Research (CBER) (2019): 7-27.

9. Izmailova ES., *et al*. "Wearable devices in clinical trials: Hype and hypothesis". *Clinical Pharmacology and Therapeutics* 104.1 (2018): 42-52.

10. Sherman R E., *et al*. "Real-world evidence—What is it and what can it tell us?" *The New England Journal of Medicine* 375.23 (2016): 2293-2297.

11. Corrigan-Curay J., *et al*. "Real-world evidence and real-world data for evaluating drug safety and effectiveness". *JAMA* 320.9 (2018): 867-868.

12. Wiens J., *et al*. "Do no harm: A roadmap for responsible machine learning for health care". *Nature Medicine* 25.9 (2019): 1337-1340.

13. Mittelstadt B D., *et al*. "The ethics of algorithms: Mapping the debate". *Big Data and Society* 3.2 (2016): 1-21.

14. Floridi L., *et al*. "AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations". *Minds and Machines* 28.4 (2018): 689-707.

15. World Health Organization. "Ethics and governance of artificial intelligence for health: WHO guidance". World Health Organization (2021): 35-37.

16. Wager K A., *et al*. "Health care information systems: A practical approach for health care management". (4th ed.,) (2017): 624.

17. European Parliament and Council of the European Union. "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)". *Official Journal of the European Union* L119 (2016): 35.

18. European Medicines Agency (EMA) & Heads of Medicines Agencies (HMA Reflection paper on the use of Artificial Intelligence (AI) in the medicinal product lifecycle, EMA/CHMP/CVMP/83833/2023, Amsterdam, EMA, (2024): 4-12.

19. U.S. Food and Drug Administration. "Proposed regulatory framework for modifications to artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD): Discussion paper and request for feedback". (2019): 1-20.

20. Ghassemi M., *et al*. "The false hope of current approaches to explainable artificial intelligence in health care". *The Lancet Digital Health* 3.11 (2021): e745-e750.