International Journal of Science and Research Archive
eISSN: 2582-8185
Cross Ref DOI: 10.30574/ijsra
Journal homepage: https://ijsra.net/

(Review Article)

Check for updates

Real-time clinical decision support with explainable AI: Balancing accuracy and interpretability in high-stakes environments

Muhammad Faheem [1, *] and Aqib Iqbal [2]

[1] *Department of Information Technology Management, Cumberland University, Lebanon Tennessee United States.*
[2] *Department of Project Management, University of Law, Birmingham, United Kingdom.*

International Journal of Science and Research Archive, 2025, 16(01), 1204-1220

Publication history: Received on 22 May 2025; revised on 05 July 2025; accepted on 08 July 2025

Article DOI: https://doi.org/10.30574/ijsra.2025.16.1.1992

Abstract

Artificial intelligence (AI) is rapidly changing the face of healthcare by reshaping the way healthcare decisions are made because it provides a new Era of Data-driven healthcare decision making, especially in high-reputation companies that require very quick and trustworthy decisions. There are also, however, serious obstacles to clinical adoption of many AI models, because they are opaque, and this requires the use of explainable AI (XAI) methods. In this paper, the architecture, challenges, and applications of real-time clinical decision support systems (RT-CDSS) augmented with XAI are considered. Based on the case studies of imaging analytics, dementia prediction, and pharmacovigilance, the study examines the effects of explainability on trust, safety, as well as system usability. Major concerns covered are whether there is a trade-off between model performance and interpretability, what technical and organizational obstacles to deployment there are, and what the ethical and regulatory environment suggests regarding making AI interpretable in the clinical context. Patient-centered outcomes are also assessed along with process evaluation measures, including SHAP, LIME, and time concerning trust in clinicians. Lastly, the paper explains the emerging trends such as human-in-the-loop structures, federated learning, and consortia functions, and presents a roadmap to realize the development of RT-CDSS not just accurate but also accountable, intelligible, and ethically compliant. These results highlight the need to develop AI technology that could be easily incorporated into clinical practice to promote transparent, semitransparent, and safe medical decisions.

Keywords: Explainable Artificial Intelligence (XAI); Real-Time Clinical Decision Support Systems (RT-CDSS); Medical AI Interpretability; Trust in Healthcare AI; Human-In-The-Loop Decision Support

1. Introduction

1.1. Background and Significance of Clinical Decision Support Systems (CDSS)

Clinical Decision Support Systems (CDSS) are gaining popularity as clinical tools that support physicians' actions by presenting him or her with intelligent and filtered facts and suggestions in the context of making decisions. CDSS as a phenomenon originated due to the growing complexity of healthcare information to help doctors in making diagnostic decisions, therapeutic decisions, and preventive measures. Their adoption in the electronic health records (EHR) has been associated with increased use of guidelines, medication safety, and best healthcare quality. Due to the transformation in clinical settings, the real-time processing functions of CDSS have become a matter of immense importance. Intensive care units (ICU), emergency departments, and surgical theaters are typical examples of high acuity that require immediate and accurate interventions and in which timing decisions can have major implications for patient survival and recovery (Robert et al. [28]).

* Corresponding author: Muhammad Faheem

1.2. Rise of Artificial Intelligence in Healthcare

Artificial Intelligence (AI) transforming CDSS has altered healthcare development, giving systems the capability to learn based on large amounts of information and provide predictive, prescriptive, and diagnostic information on an unsurpassed scale and accuracy. Machine learning and deep learning have proven to be better AI algorithms in fields such as radiology, pathology, and genomics. One of the most significant developments has been the emergence of real-time analytics engines that are automated to process medical imaging adequately and screen aberrations with similar radiologist precision by Allen *et al*.,[8]. Such systems can decrease diagnostic time, decrease the presence of human error, and aid the overworked healthcare workers. Nonetheless, the interpretability of AI is difficult due to the complexity of its outputs as well as the high dimensionality, which are challenges in the context where AI systems act as black boxes, with findings and inferences without explicit explanations. The reliability of the AI-based choice then becomes vitally important during high-risk clinical settings when misinterpretation may result in life-threatening outcomes (Bernal & Mazo, [4]).

1.3. The Need for Explainability in High-Stakes Environments

Explainable artificial intelligence (XAI) has emerged as a term that is increasingly central in combating low-explainability AI models with high performance. Explainability in healthcare is a technical requirement, but it is also both ethically and practically necessary. Clinicians need to be aware of the system recommendations and reasons that they are confident in incorporating them into the care of the patient. The absence of explainability is a major obstacle to the implementation of AI in clinical practice, as it was emphasized in the systematic review conducted by Antoniadi *et al*. [2]. High-stakes settings are settings of diagnostic uncertainty, patient vulnerability, and high decision-making speed, which increase the effects of erroneous or misinterpreted AI results. The use of transparent AI has the potential to do more than create better clinical decisions, but also to generate accountability and support patient-clinician dialogue, improving healthcare delivery in general. With the application of interpretable models, in the case of predicting the outcome of mild cognitive impairment to dementia, the model will provide essential knowledge about the characteristics (e.g., age, memory testing) that mediate risk estimates, which would allow timely personalized interventions (Chun *et al*. [7]).

1.4. Research Problem and Objectives

Although remarkable AI-enhanced CDSS capabilities are mentioned, an excellent balance between the accuracy and interpretability of models has not been reached yet. Top performers are likely to use complex neural networks, which provide little in terms of internal logic, whereas simple and interpretable models can lose predictive ability. This trade-off poses an existential research challenge: how do we design real-time CDSS to provide a high degree of correctness and significant transparency, especially within high-risk clinical settings? The current paper attempts to provide an answer to the following question: How is explainable AI applied to real-time CDSS? And what are the technical, organizational, and ethical problems in this field? Can we conclude about best practices based on the latest literature? It also seeks to compare the correspondence of clinician requirements and available XAI models to discuss whether those systems can indeed assist in the production of confident and safe decisions under stress.

2. Real-Time Clinical Decision Support Systems (RT-CDSS)

2.1. Definition and Functional Architecture

A more advanced subclass of CDSS known as Real-Time Clinical Decision Support Systems (RT-CDSS) works to structure the clinician with conceptually view-aware suggestions, alerts, and anticipatory investigation all at the correct time when a conclusion on care is about to be made. In contrast to the traditional CDSS, which can be based on the regularly updated data with the IT system making the processing computing, RT-CDSS are fed by continuously updated data or near-real-time data fed by electronic health records, bedside monitors, laboratory systems, and imaging modalities. Such systems combine inference, processing, data acquisition, and communication systems to provide the final recommendations with a combination of the latest patient status. Their design is usually multilayered: data ingestion, preprocessing and normalization, inference engines that use AI or rules logic, and delivery of the information using clinical interfaces. Allen *et al*. [8] opine that the biggest secret to high-performing RT-CDSS is that they can integrate high-throughput computational analytics into time-sensitive responsiveness and frequently use edge computing or distributed frameworks to reduce latency and secure clinical usefulness.

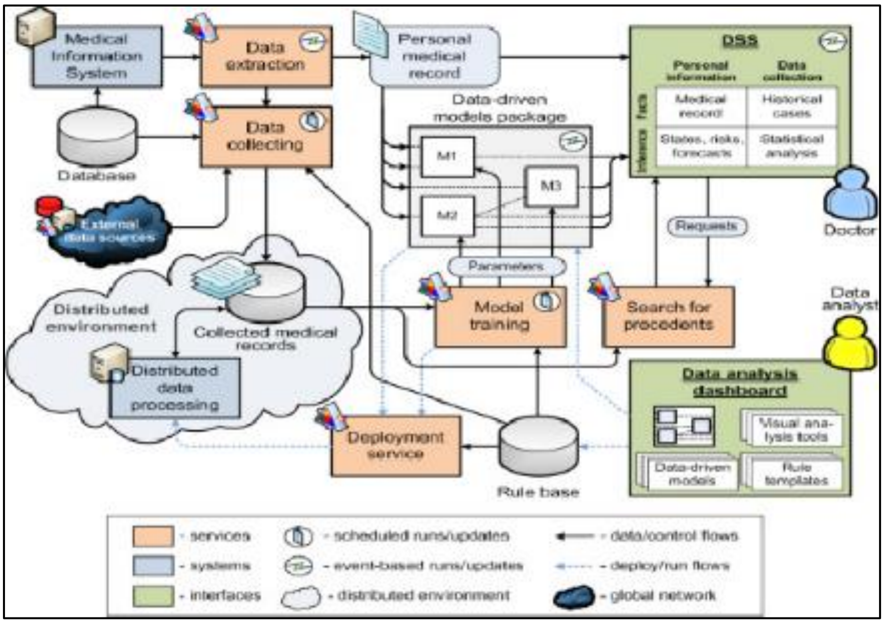

Figure 1 Functional Architecture of RT-CDSS

2.2. Integration of RT-CDSS in Clinical Workflows

The effective implementation of RT-CDSS highly relies on their incorporation into clinical practice procedures. The systems that create friction, too many notifications, or unnecessary interfaces, may be rejected by healthcare professionals regardless of their computational power. The contemporary RT-CDSSs are progressively being modeled with interoperability specifications like HL7 FHIR and DICOM to facilitate hassle-free interaction with electronic health records (EHRs), computerized physician order entry (CPOE) tools, and clinical data repositories. Effective integration implies the provision of alerts, suggestions, or visualizations in the workflow of the clinician, preferably incorporated into the EHR interface or activated by a particular patient interaction, including medication orders or diagnostic image review procedures. Real-time CDSS in clinical pharmacy practice, as presented by Robert *et al*. [28], was extremely efficient in identifying drug-related problems once the system was highly integrated into the workflow and processes of pharmacists, along with medication reconciliation. Moreover, the integration ought to consider the mental burden of clinicians transmitting only insights with high confidence that can be acted on to avoid alert exhaustion. Transparency, usability, and clinician feedback loops are key concepts of human-centered design that must be embedded into routine care without interfering with the decision authority.

2.3. Key Performance Requirements: Speed, Accuracy, and Reliability

Under high-stakes settings, the level of performance of RT-CDSS cannot be limited to the adherence of the algorithm; it should also be focused on speed, the secure and consistent work of the system, and the ability to fail. The first thing is the response time. Clinicians need to be given decision support in a clinically meaningful time window, which, depending on the nature of the problem being addressed, may be within seconds to treat a patient through resuscitation, sepsis, or triage in case of trauma. Delays can be introduced by systems unexpectedly, leading to patient harm. The paper by Allen *et al*. [8] stressed the need for real-time systems in the medical imaging analytics that involve gigabytes of complex pixel data having to be run through in sub-minute intervals, which are only achievable through GPU-accelerated and parallel computing systems. Second, precision is most important, particularly among predictive models that affect therapeutic choices or those that serve as an early-warning system. Nonetheless, reliability should not occur at the cost of high accuracy. AI models used in RT-CDSS should be interchangeable between and within patient subgroups, clinical conditions, and institutional settings, robust to missing, noisy, or biased input data. Reliability also means system uptime and fallback systems; i.e., back to rule-based alerting, where the AI engine is down. Lastly, explainability is another RT-CDSS performance demand that is unspoken because a user needs to analyze the outputs swiftly without having to technically break them apart. Antoniadi *et al*. [2] speculated that system outputs that can be explained by the user, e.g., the risk factors of a patient highlighted with a specific color, or a rationale on the internal decision of the system, are more likely to be trusted and used in time-limited cases.

2.4. Case Example of High-Performance Analytics in Imaging Systems

One of the brightest instances of RT-CDSS in practice is the automated analytics system that was created to support high-performance processing of medical images and was offered by Allen *et al*. [8]. Their system is an example of a real-time imaging decision support solution that would admit complex MRI data and quickly conduct classification tasks, e.g., to detect abnormal brain lesions. With a network-device mix of high-speed and distributed compute clusters, as well as optimized neural networks, the platform showed end-to-end performance in a clinically feasible window. The system also pioneered the ability to operate on terabytes of imaging data, but with sub-minute inference time; this was made possible by developments in GPU-based computation and parallel data pipelining. The system did not merely give diagnostic outputs but also added explainable features, including saliency maps and lesion segmentation overlay to assist radiologists' decisions. This case highlights the opportunity of real-time AI in radiology workflows where timing, accuracy, and interpretability combine to affect the process of diagnosis and patient care.

3. Explainable Artificial Intelligence (XAI) in Clinical Decision-Making

3.1. Defining XAI and Its Importance in Healthcare

Explainable Artificial Intelligence (XAI) is a package of techniques and frameworks to explain the decision-making process of machine learning (ML) models to humans. The urgency of explainability is especially pronounced in the medical field, where the decisions are life-threatening, clinics are under the state of scrutiny, and clinicians have a moral obligation to explain their actions. In contrast to traditional AI systems, which behave as black boxes, the goals of XAI systems are to make their outputs comprehensible by either highlighting associated input features, producing model-agnostic explanations, or simply using inherently interpretable models (e.g., decision trees or rule-based systems). As added by Antoniadi *et al*. [2], interpretability is the root of clinical trust ability particularly in those settings in which practitioners must combine AI findings with intricate and person-specific medical judgments. There is no explanation behind why a clinician would not want to follow the recommendation of an AI, which would be against the role of decision support.

3.2. Transparency Versus Black-Box Models

Although the convolutional neural networks and ensemble models are deep learning algorithms that have shown a remarkable record of predictive abilities, their mechanisms are not always understandable to end-users. Such a performance/transparency trade-off is the problem in clinical practice, where traceability and accountability are crucial. As stated by Bernal and Mazo [4], healthcare professionals in several nations have repeatedly voiced their concerns regarding the use of black-box AI models in clinical practice. In their results, they imply that this disclosure decreases credibility, constrains use, and complicates the insertion of AI into the multidisciplinary care workflow. Moreover, regulators and medical ethicists are increasingly supporting the so-called right to explanation in decision-making with algorithms, especially when the result of that decision process can considerably change the diagnosis, treatment plan, or prognosis of a patient. Many of the above ideas contribute to shared decision and informed consent as transparent models or models with explanatory mechanisms give clinicians additional certainty and make discussing an AI prescription easier with the patient.

3.3. Stakeholder Needs: Clinicians, Patients, and Regulators

The requirements of explainability depend on the stakeholder, each of whom has different prioritization of their expectations. The first concern is actionable insight on the part of a clinician. They need time-dependent, clinically pertinent explanations, which can be interpreted easily when one is actively making decisions. As an example, a real-time forecast of the risk of septic shock ought to tell which vital signs, lab values, or history of the patient informed the forecast, which will enable the physician to confirm the system's rationale against experience. On the contrary, patients tend to pursue understanding and assurance. They do not have to be detailed in technical information, but demand that decisions supported with AI are to agree with ethical values and human judgment. Accountability and traceability are the main priorities of regulators and hospital administrators, particularly under unfavorable circumstances. The rise in use of AI in diagnostic and prognostic systems has led to the demand that AI systems should be auditable and meet ethical standards of AI, including those contained in the European Union regime on AI (the AI Act) and the U.S. Food and Drug Administration (FDA) guidelines on Good Machine Learning Practice. XAI output alignment with stakeholder expectations is also one of the significant gaps in contemporary system design and implementation, as noted by Antoniadi *et al*. [2].

3.4. Frameworks and Techniques for Explainability in Machine Learning

Multiple approaches have been developed to improve machine learning models' interpretability in clinical decision making. These are post hoc methods like Shapley Additive ex Planation’s (SHAP), Local Interpretable Model-agnostic Explanations (LIME), and Grad-CAM on visual models, and intrinsically interpretable models such as tree-based models (decision trees, ETC) and generalized additive models. SHAP, as an example, measures the impact of individual input features on the model prediction, providing a responsible and mathematically backed rationale for the complicated decisions. LIME would be useful in understanding personal predictions by clinicians as it approximates local model behavior by modifying input features, followed by observation of output modifications. Such tools prove particularly useful when applied in real-time systems where real-time generation of feature-level rationalizations can be generated as well as understandable/visualizable. Grad-CAM methods used in clinical imaging demonstrate the areas of interest in the scans that fed the most into the classification decision given by the model, thereby effectively closing the gap between the abstraction of an algorithm and the diagnosis of a human operator Allen *et al*. [8].

Table 1 Comparison of Common XAI Techniques Used in Clinical AI Systems

XAI Technique	Type	Strengths	Limitations	Use Case
SHAP	Post hoc, model-agnostic	Theoretically solid; consistent feature attribution	Computationally expensive for large datasets	Risk prediction in EHR-based models
LIME	Post hoc, model-agnostic	Simple local explanations; works with any classifier	Can produce unstable explanations depending on perturbation	Real-time clinical alerts
Grad-CAM	Post hoc, visual models	Highlights important image regions; intuitive visuals	Limited to convolutional neural networks	Diagnostic support in radiology
Decision Trees	Inherently interpretable	Transparent logic paths; easy to follow	May underperform on complex datasets	Rule-based triage systems
Attention Maps	Model-specific	Visual insight into model focus areas	Interpretation can be ambiguous	Natural language processing in diagnostics

These methods are in use of being applied in practical systems. To offer an example, Chun et al. [7] trained an interpretable ML on dementia progression and offered feature importance drawings, which allowed clinicians to know how the score on memory tests, indicators of brain volumes, and patient demographics made individual dementia progression predictions. These types of models not only increased the reliability of predictive results but also increased credibility and acceptance within the medical field.

4. Balancing Accuracy and Interpretability in High-Stakes Environments

4.1. Trade-Off Dilemma: Performance vs. Explainability

In clinical artificial intelligence (AI) systems, interpretability and predictive accuracy have always been a conflict. Such a dilemma is explained by the common fact that very precise models--usually complex architectures, such as deep neural networks or ensemble methods--tend to be opaque black box models with minimal insight into their decision-making behavior. In contrast, more transparent models like logistic regression, decision trees, or rule-based algorithms might not perform as well, when asked to learn nonlinear relationships within high-dimensional clinical data. According to Antoniadi *et al*. [2], such a trade-off cannot be simply a technical matter in clinical settings but a matter of morality and operationalism. When the price of inaccurate or unclarified recommendations may lead to patient damage or even distrust, the explanation demand turns out to be non-negotiable. But going too far on locking in accuracy (to make the model explainable) can lead to deceptive or over-simplified solutions that are unresponsive to the complexity of disease pathology or inter-patient variation. Therefore, the design of clinical decision support systems (CDSS) requires a subtle trade-off as the selected model structure must equip accountable real-time decision-making that is not only performant, but also adequately interpretable, which in turn becomes a nuance.

4.2. Risks of Misinterpretation and Cognitive Overload

Misinterpretation or cognitive overload is one of the critical issues of the clinical deployment of explainable AI. Too technical, inconsistent explanations of concepts, and poorly visualized explanations can serve to disorientate the user rather than enlighten. Clinicians are busy in many cases, as time limits and saturation with information are considered, explanations need to be concise, context-sensitive, and clinically oriented. According to a study by Bernal and Mazo [4], a large proportion of healthcare professionals in different parts of the world stated that they could not correctly interpret the AI outputs in such situations, particularly when the explanation was expressed in typical statistical language or offered a too intense visualization. This may result in over-trusting AI results without adequate knowledge, i.e., automation bias, or wholly rejecting system suggestions based on distrust. The two situations are unsafe to patients and affect overall diagnostic precision. In addition, when an explanation is not heard or taken unfairly to formulate a decision, it can even cause other kinds of errors, which can negate the whole utility of the system. To reduce this, explainable AI should not merely enable access to feature attributions or visual maps but should do so in a way that supports, not complicates, the cognitive workflow experience of the user.

4.3. Clinical Scenarios: Emergency Medicine, Oncology, Neurology

The trade-off between accuracy and interpretability is the most obvious in high-stakes fields like emergency medicine, oncology, and neurology. Triaging, sepsis alerts, and risk stratification are all undertaken in emergency departments (EDs) where time is critical and real-time prediction models are deployed. An early model of septic shock should have responses of seconds of prediction and recognize the contributory vitals, like a low blood pressure, high heart rate, or unusual lactate count, without making clinicians want to refer to other external medical records. Explainable AI is deployed in oncology to advise the treatment schedule or detect abnormalities in histopathological images. It is essential that the oncologists who need to communicate the therapeutic possibilities and prognosis to the patient and family understand how an AI system has classified a lesion as malignant through the features of the pixels or metadata Allen *et al*. [8]. Machine learning models have also gained popularity in predicting the development of cognitive impairments into dementia in neurology. Interpretable features, as demonstrated by Chun *et al*. [7], including MRI measurements, the results of neurocognitive tests, and demographic risk factors, are critical to clinical acceptance. In both cases, the effect of the AI model will be as powerful as the capability of the model to explain its rationale in a medically relevant manner.

4.4. Real-Time Constraints and the Impact on Model Selection

Other limitations are typical of real-time clinical settings that largely affect the methods of model choice and implementation. Models need to accept patient information and provide suggestions in milliseconds to a few seconds, usually in low-bandwidth, low-computational-resource, or incomplete data conditions. Such limitations prohibit the use of strongly complex models that need multiple preprocessing or long computational cycles, though they may present slim advantages in terms of accuracy. Allen et al. [8] outlined a time-efficient analysis system of brain MRI data that ensured frame-level inference by using hardware efficiently and simplifying the model by parallelism and feature selection. Hybrid solutions are also widely used in many cases: in these solutions, an equivalent of a fast and interpretable model is used initially to perform triaging or raise an alert, and only in such cases where the model necessitates confirmation is the more complex model engaged. According to Antoniadi et al. [2], a balance between performance and explanation can be achieved by ensemble strategies, in which interpretable surrogates can approximate the black-box behavior. Finally, such models, used in real-time systems, require compromises made between computational feasibility, interpretability, and clinical impact without sacrificing the clinical integrity of the decision made by the clinician.

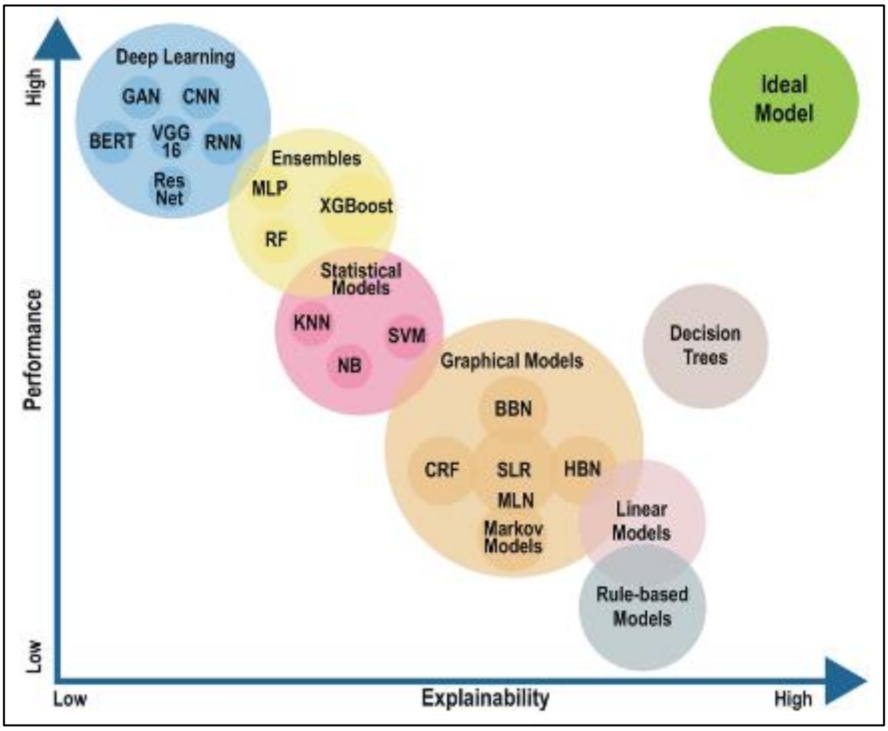

Figure 2 Trade-off Spectrum Between Interpretability and Accuracy in Clinical Models

5. Challenges of Implementing XAI in RT-CDSS

5.1. Technical Challenges: Scalability, Data Quality, and Real-Time Constraints

The technical complexity of rolling out explainable artificial intelligence (XAI) in real-time clinical decision support systems (RT-CDSS) is arguably one of the riskiest issues in achieving such a rollout in a variety of clinical settings. Real-time systems require continual data consumption, real-time low processing delay, and instant output needs. Nevertheless, a popular set of XAI frameworks are computationally expensive and involve labor-intensive feature attribution or visualization generation, which conflicts with the need for time-sensitive decision-making in clinical practice (Allen *et al*. [8]). Scalability is particularly an issue when moving pilot implementations in research hospitals to larger healthcare networks that have different IT infrastructure, electronic health record (EHR) setups, and different patient bases.

The quality of data adds a wrench to implementation as well. Clinical data tend to be noisy, incomplete, and/or inconsistently formatted, which decreases the fidelity of the AI outputs and destabilizes the reliability of the corresponding models. As it was observed by Antoniadi *et al*. [2], machine learning models trained on well-formed research datasets often fail in practice when they are put into the messiness and noise of the real world with missing data, outliers, and data drift. Such inconsistencies also extend to explanations validity by either resulting in misinterpretation or clinicians' distrust. In addition, real-time realignment of models is a major obstacle. Keeping XAI algorithms updated based on a new clinical understanding or shifting patient demographics without corresponding sacrifices in accuracy or latency is not a trivial engineering task. Cumulatively, these issues limit the strength and functionality of XAI-powered RT-CDSS.

5.2. Organizational Barriers: Workflow Integration and Staff Training

In addition to technical hindrances, there is sufficient organizational dynamics surrounding the success or failure of XAI-CDSS implementations. One of the obstacles is a mismatch between the design of AI systems and the current clinical practice. RT-CDSS should become part of clinician practice without creating cognitive load, redundancy, and without interfering with traditional communication channels. However, most applications happen without proper consideration given to human factors, which leads to minimal use, the unwillingness of frontline personnel to use the system, or a fake interaction with the results of the AI (Bernal & Mazo,[4]). Mismatch of workflow also diminishes the chance of clinicians taking to AI suggestions, especially in instances when explanations are presented outside the context of the decision-making process that include clinical rounds or MDPs.

Training the staff is also important. To use XAI-enabled systems in practice, clinicians need not only to acquire the skills of interpreting the model output, but also to decide when and how to trust the systems. This requires systematic training on the concept of AI, system-wide capacities, and the weaknesses of inference elements through machines. Most organizations do not, though, have any formal AI literacy programs, and the training available in most cases is ad hoc and not available to everyone. Antoniadi *et al*. [2] highlighted that the lack of training can instill doubts and potentially cause misuse in the form of excessive trust in AI and over-relying on the results or outright ignoring them because of a lack of understanding. Such organizational gaps need to be filled in by interdisciplinary work, a change management approach, and system design focused on users.

5.3. Regulatory and Ethical Considerations

Complex regulatory and ethical issues are also brought using XAI in RT-CDSS. In contrast to a more rigid clinical instrumentation, AI-based systems will change over time with the introduction of new information and model adjustments. This is a challenge to regulatory agencies whose roles include the safety, effectiveness, and equity of clinical decision support technologies. Regulatory channels to adaptive AI models are also immature in most jurisdictions, which creates regulatory indecision about approval, auditability, and monitoring. As pointed out by Bernal and Mazo [4], such regulatory uncertainty sours the confidence of healthcare institutions in the adoption of XAI-CDSS at scale.

The ethical issues are every bit urgent. It has started to seem possible that AI-driven healthcare transparency goes beyond simply being preferable and becomes something of an ethical necessity, at least in cases where decisions made entail prognosis, triage, or irreversible intervention. In the absence of clear explanations, patients are put at risk of having decisions they do not know or cannot challenge made against them, which contradicts autonomy and informed consent. Besides, the possibility of algorithmic bias that is associated with the model prediction that impacts are disproportionately distributed to specific demographic or clinical subgroups should be actively eliminated by means of fairness-aware training and representation of data diversity. According to Antoniadi *et al. [*2], explainability itself may come to the rescue as the interpretable models enable stakeholders to identify and criticize unfamiliar outputs. But what exists is a fragile equation that exists between meaningful explanations, the ability to render systems vulnerable to adversarial manipulation, or suffer from information overload. To deal with them, it will be necessary to create ethical standards, open governance systems, and cross-disciplinary governance.

5.4. Case Study: Detection of Drug-Related Problems via CDSS

The example of clinical practice with XAI was conducted by Robert et al. (2023), which included the implementation of a CDSS that helped clinical pharmacists detect drug-related problems (DRPs) throughout inpatient wards. Their system had interfaced with the hospital EHR and offered real-time notifications of possible interaction, repetition, and dose mistakes. Interestingly, interpretability features were incorporated into the system so that the pharmacists could view the exact criteria that were prompted by each contraindicated medication, renal levels, or age groups. The introduction of these explanations enhanced the confidence of the user and raised the acceptance levels of interventions amongst the physicians.

But there were some challenges that were created. First, the team found that the system produced a high number of alerts with some considered low-value or unnecessary. It resulted in the phenomenon of alert fatigue, which is widely mentioned in the literature on the use of CDSS and is one that can negatively affect overall decision support performance. Second, even though the system was highly interpretable, not every pharmacist was up to par with learning the model rules or the reasoning behind the system, and this underscores the gap in AI literacy. Lastly, the project emphasized the challenge of version controlling system logic according to changes in clinical guidance or formularies that were frequently done manually and needed requalification. Such questions highlight the complexities of the scalability and sustainability of XAI-CDSS.

Table 2 Summary of Reported Challenges in XAI-CDSS Implementations

Challenge Area	Description	Implications
Scalability	Difficulty deploying systems across different settings with variable infrastructure	Limits widespread adoption and increases deployment costs
Data Quality	Incomplete, noisy, or inconsistent input data	Reduces model reliability and explanation accuracy
Real-Time Constraints	High computational demand of explainability tools	Delays in clinical response and reduced system utility
Workflow Integration	Misalignment with clinician routines and digital infrastructure	Decreases usage rates and decision impact
Staff Training	Lack of AI literacy among clinicians and pharmacists	Increases risk of misuse or rejection of system outputs
Regulatory Uncertainty	Ambiguity in certification and oversight for adaptive AI systems	Slows implementation and raises legal concerns
Ethical Issues	Risks of opacity, bias, and uninformed patient consent	Undermines trust and clinical accountability
Alert Fatigue	Overgeneration of low-value notifications	Diminishes attention to high-priority recommendations
Maintenance Burden	Complexity in updating models with new guidelines or data	Reduces system agility and scalability

6. Applications and Case Studies

6.1. Medical Imaging Analytics and Real-Time AI

Medical imaging analytics can be classified as one of the first advanced and mature applications of explainable AI in real-time clinical decision-making. Radiology is one of the fields that AI high-throughput models have benefited a great deal from because of their ability to analyze complicated visual data and provide clinical-workflow-compatible diagnostic output. Allen *et al*. [8] proposed a typical concept that combines real-time analytics with sophisticated imaging methods and, in this regard, with a particular stress on magnetic resonance imaging (MRI). South Korean models were able to work with a large dataset on is smaller time scale and provide results of inference in near real-time by utilizing parallel computing and a developed neural architecture.

One of the essential properties of the system was that it had explainability modules built in; as such, along with a classification as either binary or probabilistic, clinicians could see which areas of the image were used to inform the model to use as part of its decision. It was done through the help of saliency maps and attention heat maps formed through gradient-based procedures. The explainability demands did not affect the real-time behavior of the system, which is an important finding due to the possibility of integrating transparent AI in high-performance systems. Furthermore, clinicians deemed the explanations as useful in aligning machine-generated understanding with their radiological knowledge and, hence, made informed decisions and minimized diagnostic uncertainty. This case renders the feasibility of technical innovation in the architecture of AI to hold up the speeds and interpretability interest without a trade-off with diagnostic fidelity.

6.2. Dementia Risk Prediction with Interpretable Models

The cognitive decline and the beginning of developing dementia in subjects with mild cognitive impairment (MCI) are one of the most difficult forms of disorders to predict because of the unification of symptoms and factors involved. Chun *et al*. [7] benefited from reduced interpretability in machine learning models and used interpretable machine learning models and predict the risk of dementia conversion in patients with a diagnosis of amnestic MCI. They based their method on exploiting a structured dataset that consisted of a collection of neuropsychological tests, features of MRI images, clinical histories, and demographics. The researchers have not only focused on the black-box deep learning models, but they emphasized more on explainability in the research, which was done by ensemble methods in combination with SHAP (Shapley Additive explanations) values to offer the localized and global interpretability.

The research gave the results that delayed recall test scores, volume of the hippocampus, age, and language functionality were the most contributing factors to the risk of developing dementia. The visualization of feature contribution to single predictions, which allowed the clinicians to know the reason why specific patients were marked as high-risk, gave the latter empowerment. This openness promoted clinical confidence and allowed them to be more willing to justify intervention based on active intervention, such as increased monitoring, medicinal experimentation, or thinking therapy. Moreover, the interpretability framework promoted patient-clinician interactions as the doctors could use the predictive explanations alongside patient-specific data patterns.

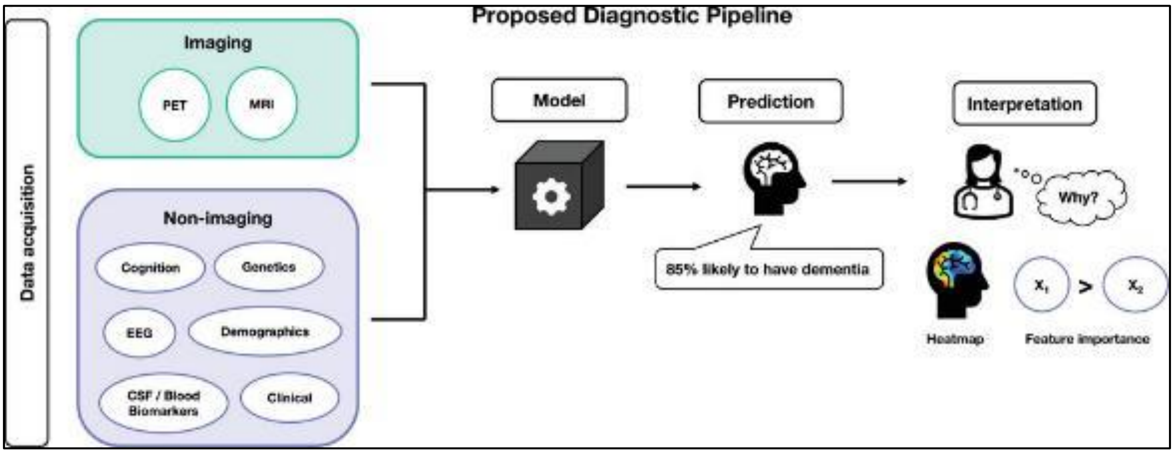

Figure 3 Interpretable ML Pipeline for Dementia Conversion Prediction

6.3. Global Insights into AI Transparency in Healthcare

Artificial intelligence transparency is both a sociotechnical issue and a global issue. Bernal and Mazo [4] performed a global investigation to examine how experts in the medical and computer sciences realm view the visibility of AI uses in medical practice. The respondent survey included more than 20 countries and revealed significant risks associated with the opacity of algorithms, their accountability, and the feasibility of explainable systems. Participants mentioned that the potential risk of errors and inefficiencies that can be improved with AI, the absence of transparent explanations, became an ongoing barrier to usage, particularly in a resource-deprived environment.

One of the most outstanding findings of the study was that there is a disconnection between the perceived transparency of AI systems and explainability in practice. Also, several respondents reported that when systems were marketed as interpretable, many of them did not provide learning that was clinically relevant and actionable. Some of the explanations were too technical or too clinically void to be used in deciding. Moreover, differences between the legal, ethical, and infrastructural frameworks across territories resulted in inconsistency in the definition and regulation of AI transparency. The researchers concluded that researchers, clinicians, and policymakers should collaborate more to develop XAI that is scalable to the globe and meets various expectations and application needs.

The case highlights the importance of aligning technological innovation to sociocultural/institutional realities. Transparency should then appear as an engineering issue, taking into consideration the other perspectives of medicine, law, ethics, and health informatics. The knowledge amassed in the international community points to the pressing necessity of more human-friendly XAI systems, which are adapted to the realities of medical personnel in various systems on each continent.

7. Evaluation Metrics for Explainable RT-CDSS

7.1. Accuracy, Precision, and Recall in Medical AI

Real-time clinical decision support systems (RT-CDSS) cannot be evaluated using the traditional performance metrics. Still, basic metrics, i.e., accuracy, precision, recall (sensitivity), and specificity, are also necessary to determine the predictive quality of the underlying artificial intelligence (AI) models. In medical use, the accuracy can be understood as the percentage of accurate predictions but could be a false indicator of accuracy with imbalanced data sets as used in clinical practice. As an example, in predicting rare diagnoses like sepsis development or malignant transformation, a high-accuracy model may not identify positive cases.

Precision and recall are thus needed. Precision quantifies the ratio of relevant predictions to all positive predictions, which is vital in ensuring that false alarms are minimized and avoiding the scenario of fatigue in this process. Recall measures the percentage of specific positive facts recognized appropriately in high-hazard situations, of which a failure to detect a conclusion has severe implications. Allen *et al*. [8] are a case in point, wherein, in their real-time imaging platform, precision-recall trade-offs in abnormal brain lesion detection were tested by making sure that the system was to prioritize clinically meaningful alerts, and that the false positives detected were kept to a minimum. These are necessary benchmarks against which to evaluate model utility before adding a requirement of interpretability.

7.2. Interpretability Metrics: SHAP, LIME, Attention Weights

The analysis of explainability has its unique collection of metrics, since conventional accuracy-measuring metrics fail to understand the explainability or interpretability of the rationale of a model. SHAP (Shapley Additive explanations) values of one of the most prevalent measures of interpretability. SHAP is a local and global explanation technique as it returns the contribution of a single input feature to the model output based on the game theory principles. It is particularly potent in the medical context, where it enables clinicians to track forecasts to the clinical characteristics, including lab measurements, demographics, or imaging outcomes (Chun *et al*. [7]).

A more common one is called LIME, or Local Interpretable Model-agnostic Explanations, and this method approximates black-box models using interpretable local surrogates. LIME can be flexible and fast, although the explanation can change as input perturbations, and meticulous validation is necessary. In practice, Grad-CAM and attention weight visualization approaches are most frequent in image-based RT-CDSS. These tools give heatmaps or areas of focus that show what regions of a radiograph or MRI had the most influence on the output of a model. Radiologists were able to detect tumor margins with the help of attention-weighted visualization, as demonstrated by Allen *et al*. [8], and this similarity in the diagnosis with humans was also affected.

Other interpretability measures are fidelity (how well explanations agree with the behavior of the actual model) and stability (how well explanations agree on similar inputs), and completeness (whether relevant information is captured by the explanation). Nevertheless, these measures are not fixed yet, and the standardization is also an unresolved question in the field.

7.3. Clinician Trust and Usability Evaluation

In addition to measures of computation, RT-CDSS evaluation is human-centered with an emphasis on clinician assessments of ranks of trust, usability, or workflow incorporation. The measurement of trust can be accomplished by validated survey measures or qualitative feedback wherein clinicians agree to the extent to which they trust, doubt, or follow the outputs of a system. As highlighted by Antoniadi *et al*. [2], perceived transparency, the clarity of the explanation, and the reliability of systems affect the level of trust extensively. As an example, explanations that are too vague or too technical dampen the eagerness to engage the system by clinicians, even in cases of high accuracy.

Usability is determined concerning interface responsiveness, clicking threshold to find explanations, the clarity of visualizations, and cognitive load testing. Applications of such tools as implementation evaluations are seen, for example, in the Arvizu R., Sutton E., Robert M. *et al*. [28] study, where researchers inferred whether the pharmacists and physicians considered the alerts clear, timely, and consistent with their clinical reasoning. Systems that can be incorporated into the existing EHR interface and explain the events in an understandable format (e.g., in natural language, or as graphical overlays) will rate higher in usability tests.

Besides, explanation rating or annotation as a part of the clinician feedback loop can also be applied to add XAI systems to even finer tuning. This is not only an improvement in human and machine decision alignment, but it can also be used as a training resource to improve a model iteratively. Such human-in-the-loop assessments help, over time, develop organizational trust and acceptance of a system.

7.4. Patient-Centered Outcomes

The overall assessment of explainable RT-CDSS should also cover patient-related outcomes, especially in the areas of safety, comprehension, and engagement in care. Patient satisfaction and sharing of decision-making and informed consent are becoming more valuable indicators of system value. Bernal and Mazo [2] revealed that the decision-making process, which concerns AI, is frequently predicted by the patients to be transparent and an understandable explanation available to the non-professional audience. As such, explainable models not only facilitate the use by clinicians but also allow patients to understand them.

Though no formal measurement of the degree to which a patient trusts an AI has been established, there are places where structured interviews, patient-reported experience measures (PREMs), and focus group research designs have been used. The outcomes of interest are the evaluation of patients depending on whether they enjoyed the fact that their data were used ethically, whether the decision-making process was understandable by them, and whether AI increased their communications with the clinicians or possibly decreased them.

The RT-CDSSs where interpretation of the information presented is feasible are more likely to engage the patients, which will lead to higher compliance levels (i.e., clear visualizations, summary of salient contributing factors, or analogy to clinical rules). These results are based not only on technical achievement but also on how the system can contribute to ethical, transparent, and patient-centered care.

Table 3 Evaluation Metrics Used in Reviewed Studies

Metric Category	**Specific Metrics/Methods**	**Application in Studies**
Performance Metrics	Accuracy, Precision, Recall, F1 Score	Evaluated model outputs in real-time medical imaging and EHR-based diagnostics
Interpretability Metrics	SHAP values, LIME outputs, Grad-CAM, Attention Maps, Fidelity, Stability	Used to explain predictions in dementia risk models, radiology systems, and CDSS alerts
Clinician-Cantered	Usability testing, Trust surveys, Interface feedback, Workflow alignment	Assessed in pharmacist-driven CDSS and hospital-wide AI transparency studies
Patient-Centered	Patient satisfaction, Comprehension ratings, Shared decision-making engagement	Evaluated via patient interviews and focus groups on AI-supported clinical decisions

8. Future Directions

8.1. Human-in-the-Loop and Interactive Explanations

Understanding and explaining explainable AI (XAI) in real-time clinical decision support systems (RT-CDSS) finds relevance since demand is increasing, and the future course can be directed in the direction of creating human-in-the-loop systems (HITL) that allow flexibility in explanations that are clinician-driven. HITL systems station the healthcare professional right at the decision-making loop, in which it can query, critique, and refine outputs of AI systems depending on contextual knowledge. According to Antoniadi *et al*. [2], they should not pretend that the explainability is a one-way process, and the model only provides a rationale, but rather that the explainability is a dialogic system, and they should allow the clinicians to explore alternate outcomes, query underlying features, and simulate what-if scenarios.

Interactions can promote trust and performance through AI-clinical logic alignment through explanations. Experimentally, it might take shape in real-time dashboard interfaces, showing SHAP or LIME values, manipulating the weight of features, or including counterfactual reasoning. Such tools would allow clinicians to adjust input characteristics, including lab values or symptoms, to track the way the model predictions vary. This way, users will value the internal logics of the model better and prove the results true or false. The combination of HITL approaches would result in a paradigm shift of passive interpretation into active interaction, which would lead to an increased level of accountability and decision fidelity.

8.2. Federated and Privacy-Preserving Learning in RT-CDSS

The other strategic focus is the integration of federated learning and other privacy-protecting technologies to increase the scalability and security of RT-CDSS. In the traditional machine learning systems, sensitive clinical data must be centralized, which presents a challenge to patient privacy, data governance, and institutional control, raising ethical issues and regulatory challenges. The federated learning method resolves this issue as AI models can be trained on several decentralized systems rather than on raw patient data (Bernal & Mazo,[4]). Rather, data is not shared and accumulated in favor of model updates to build a global model.

Federated architectures of RT-CDSS allow covering a wider range of institutions and still being strictly compliant with data protection regulations like HIPAA in the United States or GDPR in Europe. In addition to that, federated learning permits the development of more generalizable models that represent a wide range of patient populations and clinical procedures. In this way, the model is less biased and can be more robust, yet it does not lack individual privacy. Such systems and XAI modules can be combined to provide explanations of how a decentralized model handles features that might come from several different data sources.

The feasibility of this approach is further compounded by the application of privacy-preserving approaches, including differential privacy, multi-party computation, and homomorphic encryption. However, along with the development of the RT-CDSS ecosystem, the intersection of federated learning and interpretability-based frameworks will be crucial towards the construction of trustworthy, large-scale, and ethically aligned AI systems in healthcare.

8.3. Standardization of XAI in Clinical Practice

Although XAI in healthcare is gaining popularity amongst researchers, its application is fragmented by the absence of any unified measures of assessments, implementation procedures, and regulatory instructions. The available implementations differ drastically in design, user-friendliness, and interpretive fidelity, and a comparison of systems is hard to make, or gauge their readiness to be used in clinical practice. Robert *et al*.[28] found that across various departments of the same healthcare network, there was a considerable inconsistency in the interpretation and application of the explanations provided by CDSS, which frequently occurred because of the unique design of this interface or the level of user training.

The standardization of XAI metrics and reporting practices should be one of the priorities of future development. That incorporates alignment on the interpretation methods fitting different types of models, clinical tasks, and roles of users. As an example, SHAP could be appropriate in a feature-level clinical model, and Grad-CAM or attention-based visualization could be necessary in image-based diagnostics. The explanation of fidelity, stability, and comprehensibility should also be standardized when performing tests on clinical-grade XAI.

Moreover, regulatory bodies are starting to understand the demand for governance infrastructure in the context of AI explainability. Good Machine Learning Practices preliminary guidance has been published by the U.S. Food and Drug Administration (FDA), and the requirements of the EU Artificial Intelligence Act require transparency within high-risk systems, which includes medical device systems. The clinical and academic stakeholders should put in concerted actions that formulate the XAI certification standards, performance levels, and post-implementation monitoring protocols to support these regulatory attempts. By establishing standardization, the adoption will not only be much safer but also less legally ambiguous and institutionally risky.

8.4. Calls for Interdisciplinary Collaboration

Deliberation of the meaningful interdisciplinary collaboration among clinicians, computer scientists, ethicists, user interface designers, and policymakers would be the key to achieving success in future RT-CDSS with embedded explainability. In a lot of cases, an AI application is created technically in isolation and is not based on the realities of clinical procedures, interactions with patients, and regulatory conditions. Both sources, Antoniadi *et al. [*2] and Bernal and Mazo [4], point to the shortcoming of isolated design processes that implies that most explainable models are doomed to failure not due to the technical quality of the endeavor, but because of the poor fit to the expectations of the end-user or real-life operations.

Multidisciplinary teams offer different viewpoints in terms of designing and assessing XAI systems. Clinicians provide domain-specific expertise, signifying which characteristics are clinically important or how forecasts can be built on care pathways. The insights of ethicists and legal experts are submitted regarding transparency, accountability, and fairness. The researchers on human-computer interaction (HCI) make sure that information is presented in a mentally available form. In the meantime, computer scientists bring the algorithmic knowledge and experience to tradeoff between interpretability and performance.

This collaborative approach must be carried into the development of education and policy. Long-term cultural integration would be provided by training programs that can introduce medical professionals to the basics of AI and vice versa. In the same manner, interdisciplinary advisory boards will direct the ethical use and regulation of the AI systems in hospitals and health networks. Technology is just one factor in the future of explainable RT-CDSS, and thus, bringing multiple disciplines together to make such systems usable, trustworthy, and aligned with human values is the real future of explainable RT-CDSS. Deliberation of the meaningful interdisciplinary collaboration among clinicians, computer scientists, ethicists, user interface designers, and policymakers would be the key to achieving success in future RT-CDSS with embedded explainability. In a lot of cases, an AI application is created technically in isolation and is not based on the realities of clinical procedures, interactions with patients, and regulatory conditions. Both sources, Antoniadi *et al*. [2] and Bernal and Mazo [4], point to the shortcoming of isolated design processes that implies that most

explainable models are doomed to failure not due to the technical quality of the endeavor, but because of the poor fit to the expectations of the end-user or real-life operations.

Multidisciplinary teams offer different viewpoints in terms of designing and assessing XAI systems. Clinicians provide domain-specific expertise, signifying which characteristics are clinically important or how forecasts can be built on care pathways. The insights of ethicists and legal experts are submitted regarding transparency, accountability, and fairness. The researchers on human-computer interaction (HCI) make sure that information is presented in a mentally available form. In the meantime, computer scientists bring the algorithmic knowledge and experience to tradeoff between interpretability and performance.

This collaborative approach must be carried into the development of education and policy. Long-term cultural integration would be provided by training programs that can introduce medical professionals to the basics of AI and vice versa. In the same manner, interdisciplinary advisory boards will direct the ethical use and regulation of the AI systems in hospitals and health networks. Technology is just one factor in the future of explainable RT-CDSS, and thus, bringing multiple disciplines together to make such systems usable, trustworthy, and aligned with human values is the real future of explainable RT-CDSS.

9. Conclusion

This study evaluated how explainable artificial intelligence can be integrated into real-time clinical decision support systems (RT-CDSS) to enhance decision-making in high-stakes clinical environments. The analysis demonstrated that while high predictive accuracy is critical, it must be accompanied by interpretability to ensure clinician trust, safety, and actionable insight. Key findings include the technical, organizational, and ethical challenges of implementing XAI in real-time settings, as well as the importance of choosing explanation methods that are both computationally feasible and clinically understandable. The study emphasizes that successful adoption depends on human-centered design, interdisciplinary collaboration, infrastructure investment, and clinician training to bridge the gap between complex algorithms and medical reasoning. This research contributes to advancing trustworthy AI in healthcare by supporting transparent, accountable, and efficient clinical decisions—ultimately benefiting society through safer, more personalized, and ethically sound medical care.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Ahmed, Z. U., Sun, K., Shelly, M., & Mu, L. (2021). Explainable artificial intelligence (XAI) for exploring spatial variability of lung and bronchus cancer (LBC) mortality rates in the contiguous USA. Scientific reports, 11(1), 24090. https://doi.org/10.1038/s41598-021-03198-8

[2] Antoniadi, A. M., Du, Y., Guendouz, Y., Wei, L., Mazo, C., Becker, B. A., & Mooney, C. (2021). Current Challenges and Future Opportunities for XAI in Machine Learning-Based Clinical Decision Support Systems: A Systematic Review. Applied Sciences, 11(11), 5088. https://doi.org/10.3390/app11115088

[3] Banegas-Luna, A. J., Peña-García, J., Iftene, A., Guadagni, F., Ferroni, P., Scarpato, N., Zanzotto, F. M., Bueno-Crespo, A., & Pérez-Sánchez, H. (2021). Towards the Interpretability of Machine Learning Predictions for Medical Applications Targeting Personalised Therapies: A Cancer Case Survey. International journal of molecular sciences, 22(9), 4394. https://doi.org/10.3390/ijms22094394

[4] Bernal, J., & Mazo, C. (2022). Transparency of Artificial Intelligence in Healthcare: Insights from Professionals in Computing and Healthcare Worldwide. Applied Sciences, 12(20), 10228. https://doi.org/10.3390/app122010228

[5] Chatterjee, S., Saad, F., Sarasaen, C., Ghosh, S., Krug, V., Khatun, R., Mishra, R., Desai, N., Radeva, P., Rose, G., Stober, S., Speck, O., & Nürnberger, A. (2024). Exploration of Interpretability Techniques for Deep COVID-19 Classification Using Chest X-ray Images. Journal of Imaging, 10(2), 45. https://doi.org/10.3390/jimaging10020045

[6] Choi, A., Lee, K., Hyun, H., Kim, K. J., Ahn, B., Lee, K. H., Hahn, S., Choi, S. Y., & Kim, J. H. (2024). A novel deep learning algorithm for real-time prediction of clinical deterioration in the emergency department for a multimodal clinical decision support system. Scientific reports, 14(1), 30116. https://doi.org/10.1038/s41598-024-80268-7

[7] Chun MY, Park CJ, Kim J, Jeong JH, Jang H, Kim K, and Seo SW (2022). Prediction of conversion to dementia using interpretable machine learning in patients with amnestic mild cognitive impairment. Front. Aging Neurosci. 14:898940. https://doi.org/10.3389/fnagi.2022.898940

[8] Dean, N. C., Vines, C. G., Carr, J. R., Rubin, J. G., Webb, B. J., Jacobs, J. R., Butler, A. M., Lee, J., Jephson, A. R., Jenson, N., Walker, M., Brown, S. M., Irvin, J. A., Lungren, M. P., & Allen, T. L. (2022). A Pragmatic, Stepped-Wedge, Cluster-controlled Clinical Trial of Real-Time Pneumonia Clinical Decision Support. American journal of respiratory and critical care medicine, 205(11), 1330–1336. https://doi.org/10.1164/rccm.202109-2092OC

[9] E Ihongbe, I., Fouad, S., F Mahmoud, T., Rajasekaran, A., & Bhatia, B. (2024). Evaluating Explainable Artificial Intelligence (XAI) techniques in chest radiology imaging through a human-centered Lens. PloS one, 19(10), e0308758. https://doi.org/10.1371/journal.pone.0308758

[10] Fong, S., Zhang, Y., Fiaidhi, J., Mohammed, O., & Mohammed, S. (2013). Evaluation of stream mining classifiers for real-time clinical decision support system: a case study of blood glucose prediction in diabetes therapy. BioMed research international, 2013, 274193. https://doi.org/10.1155/2013/274193

[11] Ghnemat, R., Alodibat, S., & Abu Al-Haija, Q. (2023). Explainable Artificial Intelligence (XAI) for Deep Learning Based Medical Imaging Classification. Journal of Imaging, 9(9), 177. https://doi.org/10.3390/jimaging9090177

[12] Gimeno, M., San José-Enériz, E., Villar, S., Agirre, X., Prosper, F., Rubio, A., & Carazo, F. (2022). Explainable artificial intelligence for precision medicine in acute myeloid leukemia. Frontiers in immunology, 13, 977358. https://doi.org/10.3389/fimmu.2022.977358

[13] Gong, E. J., Bang, C. S., Lee, J. J., Jeong, H. M., Baik, G. H., Jeong, J. H., Dick, S., & Lee, G. H. (2023). Clinical Decision Support System for All Stages of Gastric Carcinogenesis in Real-Time Endoscopy: Model Establishment and Validation Study. Journal of medical Internet research, 25, e50448. https://doi.org/10.2196/50448

[14] Hatherley, J., Sparrow, R., & Howard, M. (2022). The Virtues of Interpretable Medical Artificial Intelligence. Cambridge quarterly of healthcare ethics : CQ : the international journal of healthcare ethics committees, 1–10. Advance online publication. https://doi.org/10.1017/S0963180122000305

[15] Islam, M. S., Hussain, I., Rahman, M. M., Park, S. J., & Hossain, M. A. (2022). Explainable Artificial Intelligence Model for Stroke Prediction Using EEG Signal. Sensors (Basel, Switzerland), 22(24), 9859. https://doi.org/10.3390/s22249859

[16] Kempt, H., Heilinger, J. C., & Nagel, S. K. (2022). Relative explainability and double standards in medical decision-making: Should medical AI be subjected to higher standards in medical decision-making than doctors? Ethics and Information Technology, 24(2). https://doi.org/10.1007/s10676-022-09646-x

[17] Kerstan, S., Bienefeld, N., & Grote, G. (2024). Choosing human over AI doctors? How comparative trust associations and knowledge relate to risk and benefit perceptions of AI in healthcare. Risk Analysis, 44(4), 939–957. https://doi.org/10.1111/risa.14216

[18] Kirkendall, E. S., Ni, Y., Lingren, T., Leonard, M., Hall, E. S., & Melton, K. (2019). Data Challenges With Real-Time Safety Event Detection And Clinical Decision Support. Journal of medical Internet research, 21(5), e13047. https://doi.org/10.2196/13047

[19] Lin, HC., Wu, HC., Chang, CH. et al. Development of a real-time clinical decision support system upon the web mvc-based architecture for prostate cancer treatment. BMC Med Inform Decis Mak 11, 16 (2011). https://doi.org/10.1186/1472-6947-11-16

[20] Madhavilatha, M., & Kumari, M. G. T. P. (2024). Interpretable Artificial Intelligence in Cardiovascular Health: An In-depth Analysis of Heart Disease Data. INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT, 08(02), 1–13. https://doi.org/10.55041/ijsrem28549

[21] Naik, H., Goradia, P., Desai, V., Desai, Y., & Iyyanki, M. (2021). Explainable Artificial Intelligence (XAI) for Population Health Management – An Appraisal. European Journal of Electrical Engineering and Computer Science, 5(6), 64–76. https://doi.org/10.24018/ejece.2021.5.6.368

[22] Novakovsky, G., Dexter, N., Libbrecht, M. W., Wasserman, W. W., & Mostafavi, S. (2023). Obtaining genetics insights from deep learning via explainable artificial intelligence. Nature reviews. Genetics, 24(2), 125–137. https://doi.org/10.1038/s41576-022-00532-2

[23] Ostropolets, A., Zhang, L., & Hripcsak, G. (2020). A scoping review of clinical decision support tools that generate new knowledge to support decision making in real time. Journal of the American Medical Informatics Association : JAMIA, 27(12), 1968–1976. https://doi.org/10.1093/jamia/ocaa200

[24] Percival, J., McGregor, C., Percival, N. et al. Enabling the integration of clinical event and physiological data for real-time and retrospective analysis. Inf Syst E-Bus Manage 13, 693–711 (2015). https://doi.org/10.1007/s10257-014-0232-9

[25] Potier, A., Dufay, E., Dony, A., Divoux, E., Arnoux, L. A., Boschetti, E., Piney, D., Dupont, C., Berquand, I., Calvo, J. C., Jay, N., & Demoré, B. (2022). Pharmaceutical algorithms set in a real time clinical decision support targeting high-alert medications applied to pharmaceutical analysis. International journal of medical informatics, 160, 104708. https://doi.org/10.1016/j.ijmedinf.2022.104708

[26] Rana , M. S. ., & Shuford , J. . (2024). AI in Healthcare: Transforming Patient Care through Predictive Analytics and Decision Support Systems. Journal of Artificial Intelligence General Science (JAIGS) ISSN:3006-4023, 1(1). https://doi.org/10.60087/jaigs.v1i1.3

[27] Rebelo, A. D., Verboom, D. E., dos Santos, N. R., & de Graaf, J. W. (2023). The impact of artificial intelligence on the tasks of mental healthcare workers: A scoping review. Computers in Human Behavior: Artificial Humans, 1(2), 100008. https://doi.org/10.1016/j.chbah.2023.100008

[28] Robert, L., Cuvelier, E., Rousselière, C., Gautier, S., Odou, P., Beuscart, J.-B., & Décaudin, B. (2023). Detection of Drug-Related Problems through a Clinical Decision Support System Used by a Clinical Pharmacy Team. Healthcare, 11(6), 827. https://doi.org/10.3390/healthcare11060827

[29] Roessner, V., Uhlmann, A., Ehrlich, S., & Waltereit, R. (2024). Taming the chaos ?! - part 2: limitations of using eXplainable Artificial Intelligence (XAI) to tackle the complexity in psychotherapy of children and adolescents. European child & adolescent psychiatry, 33(6), 1617–1619. https://doi.org/10.1007/s00787-024-02487-7

[30] Sadeghi, Z., Alizadehsani, R., CIFCI, M. A., Kausar, S., Rehman, R., Mahanta, P., … Pardalos, P. M. (2024). A review of Explainable Artificial Intelligence in healthcare. Computers and Electrical Engineering, 118. https://doi.org/10.1016/j.compeleceng.2024.109370

[31] Sariyar, M., & Holm, J. (2022). Medical Informatics in a Tension Between Black-Box AI and Trust. Studies in health technology and informatics, 289, 41–44. https://doi.org/10.3233/SHTI210854

[32] Schinle, M., Erler, C., Hess, M., & Stork, W. (2022). Explainable Artificial Intelligence in Ambulatory Digital Dementia Screenings. Studies in health technology and informatics, 294, 123–124. https://doi.org/10.3233/SHTI220411

[33] Shreve, J. T., Khanani, S. A., & Haddad, T. C. (2022). Artificial Intelligence in Oncology: Current Capabilities, Future Opportunities, and Ethical Considerations. American Society of Clinical Oncology Educational Book, (42), 842–851. https://doi.org/10.1200/edbk_350652

[34] Tjoa, E., & Guan, C. (2021). A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI. IEEE transactions on neural networks and learning systems, 32(11), 4793–4813. https://doi.org/10.1109/TNNLS.2020.3027314

[35] Tosun, A. B., Pullara, F., Becich, M. J., Taylor, D. L., Fine, J. L., & Chennubhotla, S. C. (2020). Explainable AI (xAI) for Anatomic Pathology. Advances in anatomic pathology, 27(4), 241–250. https://doi.org/10.1097/PAP.0000000000000264

[36] van der Velden, B. H. M., Kuijf, H. J., Gilhuijs, K. G. A., & Viergever, M. A. (2022). Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. Medical image analysis, 79, 102470. https://doi.org/10.1016/j.media.2022.102470

[37] W. J. Allen, R. E. Gabr, G. B. Tefera, A. S. Pednekar, M. W. Vaughn, and P. A. Narayana, (2018). Platform for Automated Real-Time High Performance Analytics on Medical Image Data, in IEEE Journal of Biomedical and Health Informatics, vol. 22, no. 2, pp. https://doi.org/10.1109/JBHI.2017.2771299

[38] Ying, Z., Fan, Y., Chen, C., Liu, Y., Tang, Q., Chen, Z., Yang, Q., Yan, H., Wu, L., Lu, J., Liu, Z., Liu, J., Li, X., & Chen, Y. (2025). Real-Time AI-Assisted Insulin Titration System for Glucose Control in Patients With Type 2 Diabetes: A Randomized Clinical Trial. JAMA network open, 8(5), e258910. https://doi.org/10.1001/jamanetworkopen.2025.8910

[39] Zheng, Y., Rowell, B., Chen, Q., Kim, J. Y., Kontar, R. A., Yang, X. J., & Lester, C. A. (2023). Designing Human-Centered AI to Prevent Medication Dispensing Errors: Focus Group Study With Pharmacists. JMIR formative research, 7, e51921. https://doi.org/10.2196/51921

[40] Zhou, S. K., Greenspan, H., Davatzikos, C., Duncan, J. S., Van Ginneken, B., Madabhushi, A., … Summers, R. M. (2021). A Review of Deep Learning in Medical Imaging: Imaging Traits, Technology Trends, Case Studies with Progress Highlights, and Future Promises. Proceedings of the IEEE, 109(5), 820–838. https://doi.org/10.1109/JPROC.2021.3054390