

## Exploring AI in Healthcare Quality and Safety



By Ajay K Gupta, CISSP, MBA, CEO HSR.health and Maria Costino, MS, RN, CPHQ

**Abstract:** As Artificial Intelligence continues to revolutionize healthcare, the quality challenges that medical AI solutions present are numerous. Healthcare Quality and Safety professionals will have the opportunity to play a crucial role in ensuring its safe and effective implementation, from validating AI-driven decisions to ensuring that evidence-based standards are upheld. This white paper provides a general overview of AI's impact addressing specific concerns for quality professionals – including the challenge of monitoring, evaluating, and maintaining safety standards in a rapidly evolving AI landscape.

# Exploring AI in Healthcare Quality and Safety

## Table of Contents

<b>EXECUTIVE SUMMARY .....</b>	<b>3</b>
<b>GATEKEEPERS OF MEDICAL AI .....</b>	<b>4</b>
<b>THE POTENTIAL OF MEDICAL AI .....</b>	<b>4</b>
<b>QUALITY CHALLENGES IN AI IMPLEMENTATION .....</b>	<b>5</b>
<b>VALIDATION OF AI-DRIVEN DECISIONS.....</b>	<b>5</b>
<b>PROACTIVE RISK ASSESSMENT.....</b>	<b>5</b>
<b>INCIDENT RESPONSE PLANS .....</b>	<b>6</b>
<b>INTEGRATION WITH CLINICAL GUIDELINES .....</b>	<b>6</b>
<b>DIAGNOSTIC TOOLS .....</b>	<b>6</b>
<b>MEDICAL AI FOR DIAGNOSTIC RADIOLOGY .....</b>	<b>6</b>
<b>MAINTAINING EVIDENCE-BASED STANDARDS.....</b>	<b>7</b>
<b>ONGOING MONITORING AND COMPLIANCE .....</b>	<b>7</b>
<b>EXPLAINABLE AI AND TRANSPARENCY .....</b>	<b>7</b>
<b>EDUCATIONAL INITIATIVES .....</b>	<b>8</b>
<b>MONITORING AND EVALUATION .....</b>	<b>8</b>
<b>SAFETY STANDARDS AND RISK MANAGEMENT .....</b>	<b>9</b>
<b>THE ROLE OF QUALITY PROFESSIONALS .....</b>	<b>9</b>
<b>BENCHMARKS FOR EVALUATION .....</b>	<b>10</b>
<b>RECOMMENDATIONS FOR BEST PRACTICES .....</b>	<b>10</b>
<b>GUIDELINES FOR IMPLEMENTING AND SCALING AI SAFELY IN HEALTHCARE .....</b>	<b>10</b>
<b>ONGOING EDUCATION AND TRAINING .....</b>	<b>10</b>
<b>LEARNING THROUGH THE VALIDATION PROCESS .....</b>	<b>11</b>
<b>CONCLUSION .....</b>	<b>11</b>
<b>REFERENCES .....</b>	<b>12</b>
<b>APPENDIX A. SUGGESTIONS FOR QUALITY LEADERS IN THE PLANNING AND IMPLEMENTATION OF MEDICAL AI SOLUTIONS .....</b>	<b>13</b>
<b>APPENDIX B. ETHICAL CONSIDERATIONS OF AI .....</b>	<b>16</b>

## Executive Summary

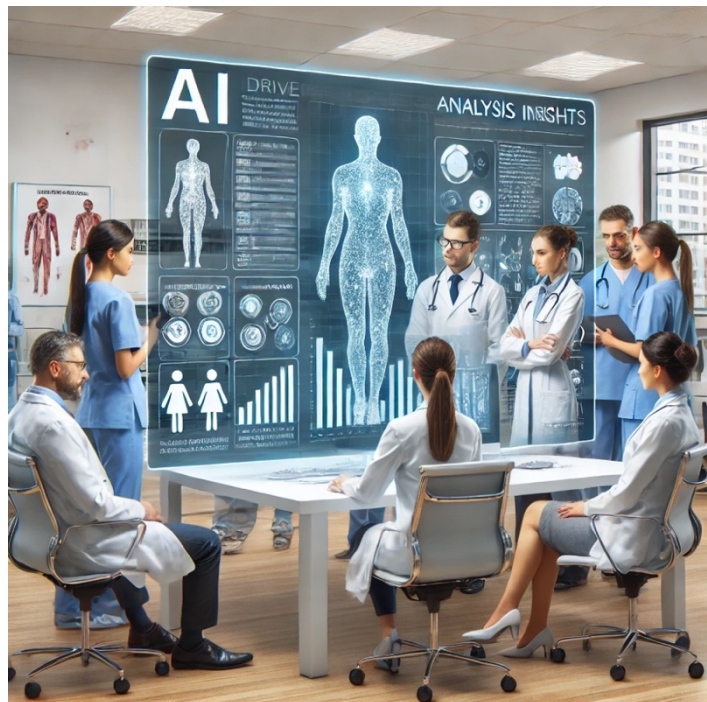
The integration of artificial intelligence (AI) into healthcare is revolutionizing patient care, diagnostics, and operational processes, offering significant potential to enhance efficiency and reduce costs. However, with this rapid adoption come challenges that health systems must address. While this is a system-wide burden, healthcare quality and safety professionals must play a key role in ensuring AI's safe and effective implementation. This white paper explores these challenges and outlines best practices for quality assurance in the deployment of medical AI solutions.

The U.S. healthcare system, with expenditures exceeding \$4.3 trillion annually<sup>1</sup>, faces unsustainable costs and growing demands for better outcomes. Medical AI presents a viable path toward streamlining processes and reducing expenses through enhanced data analysis, predictive modeling, and automation. Yet, the adoption of AI solution in administrative and clinical settings raises critical questions about validation, transparency, monitoring, and bias mitigation.

Medical AI must align with clinical guidelines and operation workflows designed to ensure patient safety. It covers essential aspects such as validating AI-driven decisions through prospective clinical trials, maintaining evidence-based standards, and employing comprehensive monitoring and evaluation frameworks, including adaptive learning systems.

To uphold high standards of patient care, quality professionals need practical strategies for continuous learning and involvement in AI implementation. Recommendations include leveraging concise resources, participating in pilot studies, and engaging in validation processes that build familiarity with AI oversight. Future considerations emphasize the need for adaptive monitoring systems that evolve with medical advancements to maintain transparency, equity, and trust in AI-driven healthcare solutions.

By adopting these best practices and fostering a proactive approach, healthcare institutions can maximize the benefits of medical AI while safeguarding patient outcomes and maintaining public confidence.



## Exploring AI in Healthcare Quality and Safety

### Gatekeepers of Medical AI

The integration of artificial intelligence (AI) into healthcare is progressing from theoretical applications to real-world implementation offering the potential to transform patient care, improve diagnostics, and advance operational efficiency. AI's potential to analyze vast amounts of data, identify patterns, and support clinical decision-making has made it an invaluable tool for healthcare providers. However, as this technology becomes more embedded in clinical workflows, the need for vigilant oversight and quality assurance becomes increasingly crucial.

Quality professionals, who play a key role in maintaining high standards of patient safety and care, face new challenges from medical AI. Ensuring that these systems function reliably, align with clinical guidelines, and deliver unbiased, accurate results is essential to preserving trust in AI-driven care – and in healthcare overall. With AI's potential to assist in areas ranging from diagnostic imaging to personalized treatment plans, the opportunities are substantial and so are the risks if proper quality measures are not enforced.

Maintaining robust oversight of medical AI solutions will necessitate both validating the solutions in their current state as well as tracking and adapting to ongoing advancements in new versions. Understanding these aspects are essential in navigating the complexities of AI integration and safeguarding patient outcomes.

### The Potential of Medical AI

The U.S. healthcare system faces significant challenges, with unsustainable costs and inefficiencies that strain both providers and patients. According to the Centers for Medicare & Medicaid Services (CMS), national health expenditures reached \$4.3 trillion in 2021<sup>1</sup>. The Peter G Peterson Foundation reports healthcare spending rose to \$4.8 trillion in 2023 and is projected to reach \$7.7 trillion in 2032<sup>2</sup>, outpacing economic growth and placing substantial burden on public and private sectors (see figure 1). The urgency to identify cost-saving measures and streamline healthcare delivery has never been greater, and artificial intelligence (AI) stands as a promising solution.

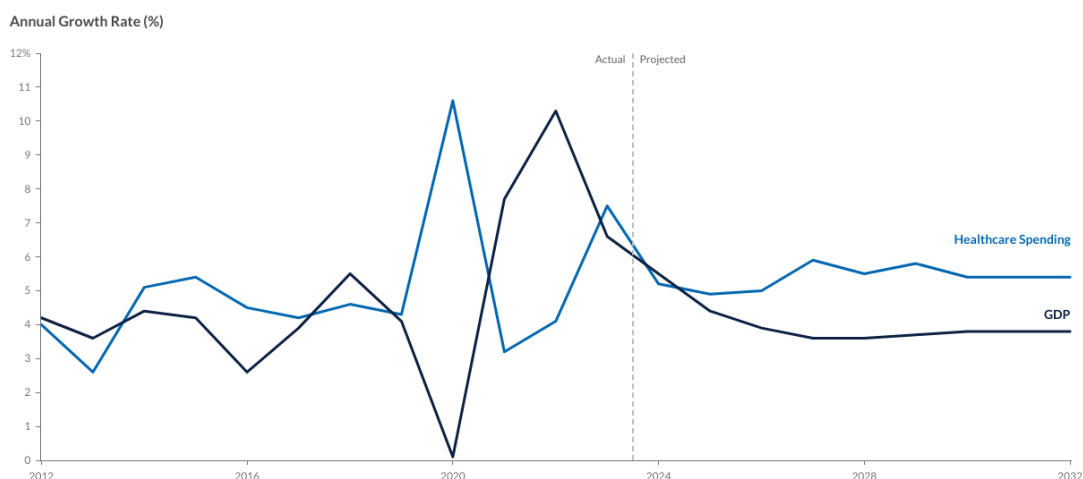


Figure 1: Healthcare spending is projected to grow faster than the economy over the next decade. Sources: Congressional Budget Office, Office of Management and Budget, and Centers for Medicare and Medicaid Services.



AI offers the potential to revolutionize healthcare by enhancing efficiency, reducing costs, and improving patient outcomes. Its capability to process and analyze vast datasets at unprecedented speeds can assist in early diagnosis, personalizing treatment plans, optimizing administrative workflows, and reducing human error. For instance, AI-driven predictive analytics have shown promise in managing patient care effectively, reducing hospital readmission rates and unnecessary testing, which translates to significant cost savings<sup>2</sup>.

Furthermore, labor-intensive administrative tasks, such as billing, coding, and scheduling can be automated with AI freeing up clinicians to focus on patient care. A study by McKinsey & Company estimated that AI could potentially reduce administrative costs in the U.S. healthcare system by approximately \$265 billion annually<sup>4</sup>. This financial relief is critical as providers face escalating operational expenses while grappling with workforce shortages.

The integration of AI into healthcare can also promote a more resilient system capable of adapting to new challenges. By leveraging AI solutions to identify efficiencies and support clinical decision-making, the healthcare sector can move toward a more value-based model that prioritizes quality and affordability. However, as these tools become more embedded in clinical practice, quality professionals must ensure the applications are implemented thoughtfully, maintaining alignment with evidence-based care and safeguarding patient trust.

## Quality Challenges in AI Implementation

It is well known that consumer-grade AI solutions are not driven to have a high degree of accuracy and are subject to bias, error, as well as hallucination. All of these are unacceptable for health use cases.

### Validation of AI-Driven Decisions

Prospective clinical trials are essential for validating AI tools in healthcare, ensuring they meet rigorous safety and efficacy standards before widespread adoption. The primary responsibility for conducting these trials often falls on AI developers and technology companies. These developers typically initiate and fund trials to demonstrate that their products align with clinical best practices and meet regulatory requirements. Government or funding agencies can also provide financial support.

Collaboration with hospitals, academic medical centers, and clinical research organizations is common, offering access to diverse patient populations and real-world healthcare environments. Such partnerships help to ensure comprehensive testing and add credibility to the validation process.

These trials represent an opportunity for health systems and providers to validate the real-world applicability of the AI solution and ensure that it integrates seamlessly into existing patient care and operational workflows.

### Proactive Risk Assessment

The validation process should incorporate the design and testing of proactive risk assessment protocols that help identify potential issues early by simulating different scenarios and stress-testing outputs. This approach ensures that potential risks, such as performance variability or data biases, are recognized and addressed before full deployment.

## Incident Response Plans

As part of the validation process, it is critical to establish incident response plans that outline procedures for managing AI failures and unexpected outcomes, including steps for rapid intervention, investigation, and patient communication. This ensures that healthcare providers are prepared to handle any disruptions, which is key to maintaining patient trust and safety.

Hospitals, large healthcare systems, and academic and research institutions may also conduct their own pilot studies or validation trials to assess the tool's performance before full integration. This approach allows these institutions to tailor the tool's use to their specific needs and patient demographics.

Regulatory bodies, such as the FDA or EMA, oversee the standards and frameworks for these trials but do not conduct them directly. Instead, they require clinical trial data as part of the approval process.

## Integration with Clinical Guidelines

AI models must be designed and validated to follow established clinical protocols. This ensures that AI outputs support decision-making that is consistent with accepted medical practices. Regular updates to the models are necessary to reflect new medical evidence and guidelines – as well as new software – enabling the AI to remain effective and compliant over time.

Two notable examples include PathAI's histopathology tools and Aidoc's radiology AI for stroke detection.

## Diagnostic Tools

One of the early uses of the Internet was for individuals to look up their own or a loved ones' symptoms to learn what was wrong. The era of AI has taken this to an entirely new level.

- PathAI has conducted pilot programs and validation studies to enhance the accuracy of cancer diagnoses through AI-assisted analysis of pathology slides. Early results indicated that PathAI's tools could support pathologists in achieving more consistent and precise diagnoses<sup>5</sup>.
- There are growing reports that even non-healthcare focused AI tools, such as ChatGPT can outperform human physicians in patient diagnosis<sup>6</sup>.

## Medical AI for Diagnostic Radiology

Leveraging artificial intelligence for diagnostic radiology has almost been considered “low-hanging fruit” in terms of the technology's ability to perform the service and achieve implementation. And there have been a number of offerings in this area, partially including:

- Aidoc has developed AI solutions for radiologists, for the identification and triage of both ischemic and hemorrhagic stroke in CTs, including an AI module for flagging and prioritizing intracranial hemorrhage, potentially speeding time to treatment<sup>5</sup>.
- The INSIGHT MMG mammography model has been shown to detect breast cancer up to six years earlier than traditional screening methods, with reduction in false positives and false negatives<sup>10</sup>.

That AI hasn't taken over radiology underscores the importance of collaborative efforts between AI developers and healthcare institutions to validate AI tools, build confidence in results, integrate in existing workflows, ensure compliance with clinical standards, reduce costs, and improve patient outcomes. These criteria must be met for AI solutions to gain wide adoption in healthcare.



## Maintaining Evidence-Based Standards

Maintaining evidence-based standards is an ongoing challenge for medical AI solutions, requiring alignment with current medical practices and adaptability to new research that may change the standard of care. Continuous monitoring and compliance checks are crucial for maintaining high standards of care.

### Ongoing Monitoring and Compliance

AI tools should undergo regular performance assessments to ensure they adhere to the latest clinical guidelines. This includes periodic validation tests, recalibrations, and audits that align with updated medical protocols. For example, implementing dashboards that track performance metrics such as precision, recall, and accuracy can help monitor AI tools effectively. Additionally, continuous data input from real-world cases should be analyzed to confirm that the AI model maintains its reliability and compliance over time. This should be aligned with existing internal quality control procedures.

### Explainable AI and Transparency

Maintaining evidence-based standards requires AI systems to be transparent in their decision-making processes. Much of modern AI operates as a “black box” system without providing clarity into how the system arrived at its output. This also does not work in healthcare where visibility and trust are of paramount importance. Medical AI solutions must be “explainable” – in other words, there must be transparency in how and why an AI process arrived at its output, whether a patient diagnosis or operational decision. Explainability is key for building trust and facilitating the verification process against evidence-based standards. For instance, in diagnostic radiology, an AI solution should create radiology reports that provide clear visual cues and decision paths that radiologists can cross-reference with clinical guidelines.

Standard monitoring approaches, such as the following, can help provide quality oversight:

- **Performance Dashboards:** Display real-time metrics on the accuracy and consistency of AI outputs.
- **Automated Alerts:** Trigger notifications when the AI deviates from evidence-based practices.
- **Periodic Revalidation:** Regularly scheduled updates and testing phases ensure the AI adapts to new medical standards.

- **Audit Logs:** Maintain detailed records of AI decision paths to support retrospective reviews and compliance checks.

These tools will have to fall within an overall process that manages patient care early enough to intervene when a case is going south. They must fit within clinical workflows with the option to create alerts when an AI deviates from accepted practice. This requires the medical AI solution to know the accepted practices and have the necessary comprehension and skill to document why it is deviating in a particular case.

Well designed, these solutions can help ensure AI solutions align with evolving medical knowledge and continuously provide high-quality patient care.

### Educational Initiatives

It is equally important to train healthcare professionals on how to interpret and verify AI outputs. Educational programs should include the strengths and limitations of AI tools to ensure that clinicians can effectively integrate them into patient care while remaining vigilant about maintaining evidence-based practices.

These comprehensive approaches help ensure that AI solutions continue to align with evidence-based standards and adapt to advancements in medical knowledge, thereby maintaining the quality and safety of patient care.

### Monitoring and Evaluation

Effective monitoring and evaluation of AI systems extend beyond initial validation to include a framework for continuous oversight. One such approach is the **Translational Evaluation of Healthcare AI (TEHAI)** framework which offers a comprehensive approach to evaluating AI applications, focusing on their capability, utility, and adoption within clinical settings. This framework emphasizes the importance of assessing the functional, utility, and ethical aspects of AI systems to ensure their effective integration into healthcare workflows<sup>6</sup>.

Additionally, the **Health Equity Assessment of Machine Learning (HEAL)** framework provides a structured method to evaluate the performance of AI tools across different demographic groups, ensuring equitable healthcare outcomes. HEAL introduces a four-step process to assess whether an AI tool performs consistently across diverse populations, addressing potential biases and disparities in healthcare delivery<sup>7</sup>.

Advanced Monitoring Techniques include:

- **Adaptive Learning Algorithms:** Utilize real-time data to modify and improve AI outputs automatically, ensuring they remain aligned with current standards.
- **Continuous Feedback Loops:** Collect data from end-users and patients to identify potential issues and enhance system performance.
- **Robust KPI Tracking:** Key performance indicators such as false positive/negative rates, specificity, and sensitivity should be measured and compared to benchmark levels regularly.

KPIs should also include the metrics against which health systems and hospitals are currently measured against, such as readmissions, hospital acquired infections, falls, etc. This represents an opportunity for healthcare to work with technologists to develop solutions that address their existing pain points.



Structured frameworks like TEHAI and HEAL are helping in developed monitoring and evaluation processes for the safe and consistently effective use of AI tools.

## Safety Standards and Risk Management

Any use of AI in healthcare must address its recognized bias, error, data privacy, patient safety, and hallucination risk. Each of these challenges poses distinct risks that must be managed to uphold patient trust and system reliability. To address bias, it is critical to use diverse training data and conduct regular bias audits that help identify and rectify disparities affecting patient outcomes.

Mitigating errors involves robust validation, continuous real-time monitoring, and human oversight, especially for high-risk applications, to identify and correct discrepancies swiftly. Ensuring data privacy requires adherence to regulatory standards such as HIPAA and GDPR, secure data protocols, and data anonymization practices to protect patient information during training and analysis phases. Enhancing patient safety calls for well-defined risk management protocols and incident response plans that include steps for rapid intervention, investigation, and patient communication when unexpected outcomes occur. Preventing hallucinations involves limiting AI outputs to validated and specifically curated data sources, refining models through high-quality training, and using multi-source verification to ensure AI-generated insights are accurate. Addressing these issues holistically guarantees that AI systems in healthcare maintain safety, transparency, and patient confidence.



## The Role of Quality Professionals

Quality professionals play an essential role in ensuring the safe and effective deployment of medical AI. Their responsibilities include overseeing the validation and continuous monitoring of AI systems, ensuring alignment with clinical best practices, and maintaining patient safety and high standards of care. They must actively track the impact of AI on patient outcomes, monitor diagnostic and treatment recommendations, and ensure that the AI-driven decisions align with patient-specific needs and clinical guidelines.

Quality professionals are also responsible for developing robust documentation that captures the AI tool's workflow, compliance to standards, and update history to support transparency and readiness for audits. They lead efforts in conducting risk assessments to identify and mitigate potential biases and unintended consequences. Additionally, they evaluate the relevance and appropriateness of AI-suggested tests and treatments, verifying that they contribute positively to patient care without unnecessary procedures or risks.

### Examples of Effective Quality Assurance Practices:

- **Patient Outcome Tracking:** Implementing systems to monitor patient outcomes linked to AI-assisted diagnostics and treatments ensures that the AI tool contributes to improved clinical results.

- **Routine Clinical Validation:** Regular comparison of AI outputs with expert clinician reviews helps to maintain consistency and accuracy in diagnostic and treatment decisions.
- **Review Panels for AI Decisions:** Multidisciplinary panels, including clinicians and quality professionals, assess AI-driven recommendations for clinical appropriateness and adherence to medical guidelines.

These measures must be taken after diagnosis and during treatment as well. Assessing outcomes is important and easier since it is more likely to be a part of existing processes. However, if the AI has deviated from the standard of care or made an error – waiting for the outcome may be too late. Post diagnosis and during treatment, such as during a hospital stay, is helpful to implement. Such workflow may be new and therefore such mechanisms should be a part of the AI tool itself and/or its trial deployment.

### Benchmarks for Evaluation

Quality assurance practices should be measured against benchmarks such as clinical accuracy, regulatory compliance, patient outcome metrics, and safety incident frequency. Other benchmarks include bias detection scores, interoperability with existing health IT systems, and user trust levels, which gauge adoption among healthcare professionals.

Quality professionals can also gain significant insights through active participation in the validation process. Engaging in validation not only familiarizes them with how AI systems function but also equips them to become more adept at monitoring and providing the necessary oversight. This hands-on experience serves as a practical training method, enabling quality teams to build the skills required for effective management of AI-driven solutions.

### Recommendations for Best Practices

For quality professionals in healthcare, staying informed about AI advancements and maintaining effective oversight requires practical, time-efficient strategies. While formal courses and certifications offer value, more immediate and accessible options are essential for busy professionals.

### Guidelines for Implementing and Scaling AI Safely in Healthcare

Quality professionals should engage with concise, impactful resources like industry reports, research summaries, and webinars. Attending medical AI-focused conferences or virtual panels provides quick, relevant insights without demanding significant time. Furthermore, the validation process itself can be leveraged as a practical training tool. By participating in validation activities, quality teams gain hands-on experience that helps them become familiar with monitoring procedures and confident in providing oversight for AI solutions. Quality professionals should also reach out to initiate partnerships with regulatory agencies and professional organizations, such as Joint Commission and the National Association of Healthcare Quality to help develop guidance documents.

### Ongoing Education and Training

Rather than extensive training programs, subscribing to reputable newsletters, following AI thought leaders, and participating in brief, targeted workshops or panel discussions can keep professionals informed. This continuous learning approach integrates seamlessly into daily routines, ensuring professionals stay current with minimal disruption.

## Learning through the Validation Process

An essential component of professional development for quality teams is active involvement in the validation process. This stage allows teams to observe AI behavior in real-world scenarios, test its adherence to safety standards, and refine their understanding of monitoring requirements. By integrating these insights, quality professionals become well-equipped to provide the necessary oversight and maintain high standards of patient care.

## Conclusion

As AI continues to evolve and transform the landscape of healthcare, quality professionals must remain proactive in adapting their practices to oversee these technologies effectively. Although many foundational elements for monitoring and validating medical AI are being established, significant work remains to be done to build and adopt comprehensive frameworks supporting sustained quality assurance and safety that fit within existing clinical workflows.

Future considerations should focus on the integration of adaptive learning practices that enable AI systems to evolve alongside clinical guidelines, ensuring that patient care continues to be informed by the latest, validated medical advancements. Quality professionals must anticipate the challenges associated with implementing these adaptive systems, including maintaining transparency and mitigating biases that may emerge as AI models learn from new data.

Looking ahead, the role of quality professionals will involve not only maintaining existing oversight mechanisms but also embracing new technologies and methodologies that enhance their ability to monitor AI-driven solutions. By participating in validation and pilot studies, quality teams can build their knowledge and comfort with AI oversight, ultimately reinforcing their ability to ensure safe and effective patient outcomes.

As medical AI becomes more prevalent, quality professionals will be well served through active involvement in the implementation and use of these solutions. The future of AI in healthcare holds immense potential for improving patient outcomes and operational efficiency, but this potential can only be fully realized with the diligent oversight of quality professionals who uphold the highest standards of safety and efficacy. By embracing the validation process as a training tool and leveraging adaptive monitoring frameworks, healthcare institutions can ensure that AI-driven care remains trustworthy, equitable, and beneficial for all patients.



## References

1. CMS, Office of the Actuary, [National Health Spending Grew Slightly in 2021](#), December 14, 2022.
2. Peter G Peterson Foundation, [Healthcare Spending Will Be One-Fifth Of The Economy Within A Decade](#), September 16, 2024.
3. Huang, Yinan, et al., [Application of machine learning in predicting hospital readmissions: a scoping review of the literature](#), BMC Medical Research Methodology, 21, Article number 96, 2021.
4. McKinsey & Company, McKinsey Center for US Health System Reform, [Administrative simplification: How to save a quarter-trillion dollars in US healthcare](#), October 2021.
5. Path AI, <https://www.pathai.com/resources/pathai-and-concertai-partner-to-create-first-in-class-histopathology-and-clinical-real-world-data-solutions/>.
6. New York Times, [A.I. Chatbot Defeated Doctors at Diagnosis](#).
7. Aidoc, <https://www.aidoc.com/about/news/ai-saving-brain-fda-clears-aidocs-complete-ai-stroke-package/>.
8. Reddy, Sandeep, et al., BMJ Health & Care Informatics, [Evaluation framework to guide implementation of AI systems into healthcare settings](#), October 12, 2021.
9. Schaeckermann, Mike, et al., [HEAL: A framework for health equity assessment of machine learning performance](#), March 15, 2024.
10. Gjesvik, J., Moshina, N., Lee, C. I., et al. (2024). Artificial Intelligence Algorithm for Subclinical Breast Cancer Detection. *JAMA Network Open*. doi:10.1001/jamanetworkopen.2024.37402.
11. Thomas, Liji, [AI detects breast cancer years before diagnosis from mammograms](#), News Medical, October 8, 2024.

## Appendix A. Suggestions for Quality Leaders in the Planning and Implementation of Medical AI Solutions

Below is a list of suggestions for Healthcare Quality Leaders engaged in the development and deployment of medical AI systems. Any or all of these should play a role in such efforts.

### 1. Identify a Clear Use Case and Define Objectives:

- Determine a critical challenge that can be addressed with medical AI. This includes automating medical billing or scheduling, the identification of high-risk patients for a specific disease or condition (e.g., pressure injuries).
- Set clear, measurable goals and objectives for the AI implementation. Examples include improved patient outcomes, patient satisfaction, reduced treatment costs.

### 2. Convene a Multidisciplinary Team:

- Membership should include key stakeholders as well as those with expertise in the disciplines relevant for the specific medical AI solution being developed. This may include data analysts, clinicians, IT specialists, compliance, senior leadership, patient advocates, case managers, and of course quality personnel.
- To avoid the team becoming too large, it can sometimes be advisable to include subject matter experts on an on-call basis.

### 3. Data Acquisition and Preparation:

- Data collection may not be the role of quality, however, measures should be in place to ensure high-quality, relevant data should be collected from various sources, partially including electronic health records (EHRs), medical imaging, and patient registries.
- In addition, measures to ensure accuracy, completeness, and consistency of the data along with processes for handling missing data, correcting data and standardizing data formats.
- Provenance.

### 4. Selecting and/or Developing AI Models:

- The team must have confidence that the particular model(s) used are specifically suited to the project at hand.
- This typically involves evaluating the AI developer's efforts in model development, selection, training, and validation – such as through review of technical papers and past pilots and implementations.



- It is also possible that project team becomes involved in model selection and development depending on the stage of the overall solution.
- In either case, the project team should play a role in training and validating the AI model(s) using the data prepared for this implementation.
- This may require several rounds of review and testing. Be sure to implement security measures for patient privacy and compliance with regulatory requirements, such as HIPAA.

5. Support Pilot Testing and Validation:

- Conduct pilot testing of the AI system in a controlled environment to evaluate its performance and identify potential issues.
- This also involves ensuring the AI system integrates well with existing workflows – or supports desired workflow changes – as well as user acceptance testing.

6. Provide Insight into Ethical Considerations and Oversight:

- While ethical considerations should be considered by all project team members, quality personnel can play a role in proactively addressing ethical considerations, such as algorithmic bias and hallucinations.
- Mechanisms for ongoing monitoring, evaluation, and oversight of the AI system are also critical to ensure the solution continues to serve patient needs and meet its goals and objectives.

7. Implementation and Integration:

- It will be important to have a detailed implementation plan outlining the steps involved in achieving integration with existing IT infrastructure, such as EHRs and other clinical systems.
- Each step of the implementation plan should have demonstrable metrics assuring its successful completion.

8. Training and Education:

- Assist with training (mandatory) and both initial and ongoing education to healthcare professionals on how to use and interpret the AI system effectively.
- Encourage user adoption and ensure that healthcare professionals are comfortable using the AI system in their daily practice.
- At initial roll-out, consider having tech support at the “elbow” of users to provide real-time reference, guidance, and reinforcement.

9. Continuous Monitoring and Improvement:

- Continuous monitoring of the performance of the AI system is a good way to collect feedback from users that can help improve the solution overall.

- Update and refine the AI system at regular intervals, as indicated, based on new data, user feedback, and any advancements in the AI solution utilized.
- While great care will be given during the implementation project to ensure the solution is effective – there is no substitute for real-world use to guide project delivery.

## Appendix B. Ethical Considerations of AI

While there are ethical considerations that need to be addressed by the legal community when implementing AI, those discussions should not impede the implementation of potentially beneficial medical AI solutions. Such discussions can be held parallel to the planning process. As a quality leader, your input might be solicited on several of those matters to ensure these systems are developed and implemented in a responsible and equitable manner.

First and foremost is the privacy and security of the data used or gleaned in conjunction with the medical AI solution, with the caveat there might be some challenges as clarity and guidance in certain circumstances might not be clear. Robust safeguards must be in place to prevent any breaches, unauthorized use, and misuse of patient information. De-identification is one way to protect patient privacy. However, there could easily be times when data collected from patients might be useful in the development of a separate medical AI solution at a later date. This needs to be a part of any discussions on ethical considerations related to privacy.

Informed consent is another important consideration. The benefits and limitations of the use of AI in the patient's direct care, when its use at the time is very distinct and well-defined, must be clearly explained in a manner the patient understands which will allow the patient to make informed decisions about its use in their care and treatment.

In direct patient care, the use of medical AI is distinct and well-defined and can be more clearly explained to patients and their caregivers in a manner allowing for informed consent. However, when a patient's data is used in the development, training, or validation of a future model, there can be less clarity on how to provide information to patients on which true informed consent can be given.

Equitable access to AI-powered solutions *must be guaranteed for all patient populations regardless of socio-economic status or other factors*. When developing the AI solution, care must be taken to ensure there are no limitations in the system that might exacerbate already existing disparities. Not all solutions meet the needs of every population and, as most hospitals have diverse populations, this element needs to be factored into the development of the system. This is essential when choosing data for training purposes. If the data chosen is specific to only certain demographics, the AI system may perpetuate biases with resultant inequities.

There must be clear and concise guidelines as to who is responsible for monitoring the medical AI solution and the steps that must be taken to address any errors or malfunctions – especially for those that may cause patient harm. Depending on the severity, the error may need to be disclosed to the patient, reported to appropriate regulatory agencies, organizational leadership, and the board of trustees.

While not related to the care of the patient, another factor of ethical concern is the potential for healthcare workers to experience job displacement as a result of automation and other operational and productivity gains that medical AI solutions may offer. Roles as diverse as

laboratory technicians, appointment schedulers, and even potentially nursing may be at risk from such displacement. The organization should determine how it would handle this type of situation, such as training staff to develop new skills to either complement the AI solution or upskill to transition the work force to new positions.

Something that should be on the forefront of everyone is the risk of becoming over-reliant on a medical AI system leading to a dehumanization of patient care by prioritizing data and algorithms over human connection and empathy. These systems need to be utilized in a way to enhance the person-to-person interaction by freeing workers from time consuming, non-patient facing tasks and providing more time for direct patient contact.

As described, the utilization of medical AI solutions in a healthcare setting is infused with a complex ethical landscape. However, organizations can successfully maneuver the terrain with careful planning, thoughtful discussions, clear objectives, as well as continuous collaboration and open dialogue among all key stakeholders.