Evidence Cloud™

[cite] med

AI in Medtech Clinical Evaluation

A Practical Approach & Implementation Guide

Managing The Technology Hype Cycle - *A path to selecting AI solutions that work, deliver value, and provide a demonstrable return-on-investment.*

[cite] med

EXECUTIVE SUMMARY

Medical device companies face a critical decision: how to evaluate and implement AI solutions that deliver real value while maintaining regulatory compliance. Despite vendor promises, 70% of AI projects fail to deliver expected ROI due to hidden costs, poor implementation strategies, and misaligned expectations.

This whitepaper provides a practical framework for evaluating AI solutions across four critical dimensions: technical capabilities, security and compliance, implementation readiness, and measurable ROI.

We identify literature review automation as the highest-value starting point, offering proven 75% time savings with minimal risk. Key evaluation criteria include understanding the true total cost of ownership (3-4x the licensing fees), ensuring vendor AI models can demonstrate clear reasoning for regulatory audits, and selecting platforms that leverage existing enterprise infrastructure to accelerate deployment.

Organizations should prioritize solutions built on foundation models from major providers (OpenAI, Anthropic, Google) that offer continuous improvements, while ensuring all enterprise AI agreements protect against data reuse.

Success requires treating AI as an enhancement to human expertise rather than a replacement, with clear boundaries between automated assistance and professional judgment. By following this framework and starting with proven use cases, medical device companies can achieve significant efficiency gains while maintaining the quality and compliance standards their industry demands.

AUDIENCE FOR THIS WHITEPAPER

This whitepaper is written for medical device literature teams, regulatory and compliance organizations, IT, and executive leaders who have been tasked with researching, purchasing and implementing AI solutions with real business value within their organizations. Regulatory industry and life sciences professionals seeking to learn more about AI in their domain, will also gain thoughtful insight, industry applicability, and concepts helpful for career development.

TABLE OF CONTENTS

SECTION 1: MEDTECH AI - PROMISE VS. REALITY

1.1 Hype Cycle

- Broad-based vendor claims on "AI-powered" capabilities
- Material gap between marketing promises vs. technical reality
- 70%+ AI projects fail to define or deliver ROI

Industry-specific AI for the regulatory domain is current going through the very early stages of the traditional new technology hype cycle as described below in **Table I**.

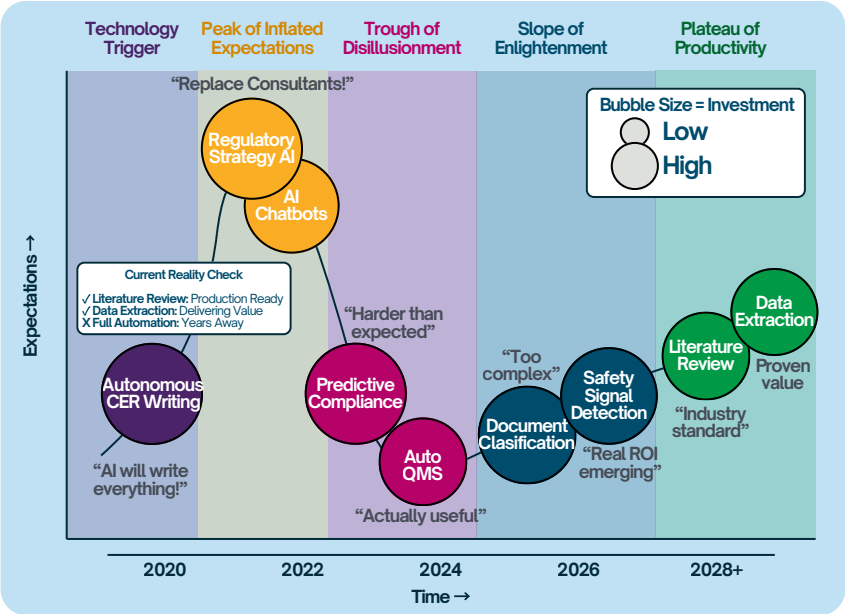


Table I - Regulatory Affairs - AI Technology Hype Cycle

AI has taken over... the attention of Marketing departments.

While we have not seen day to day adoption within Medtech organizations aside from basic use cases, it seems everyone is pursuing AI solutions.

This whitepaper presents a framework for the evaluation of AI based solutions. It will provide a non-technical introduction to the core concepts of AI so decision makers can understand what they are buying.

We will present a rubric for grading AI solution vendors, and a detailed list of questions to ask (and answers to expect). Finally we will close with detailed case studies and discussion of real world solutions beyond the "theory" of AI capabilities.

1.2 Why Traditional Evaluation Methods Fail for AI

- **Software demos don't reveal AI limitations**

For AI solutions this demos are not sufficient. Efficacy of AI systems is dependent on real world data, your data.

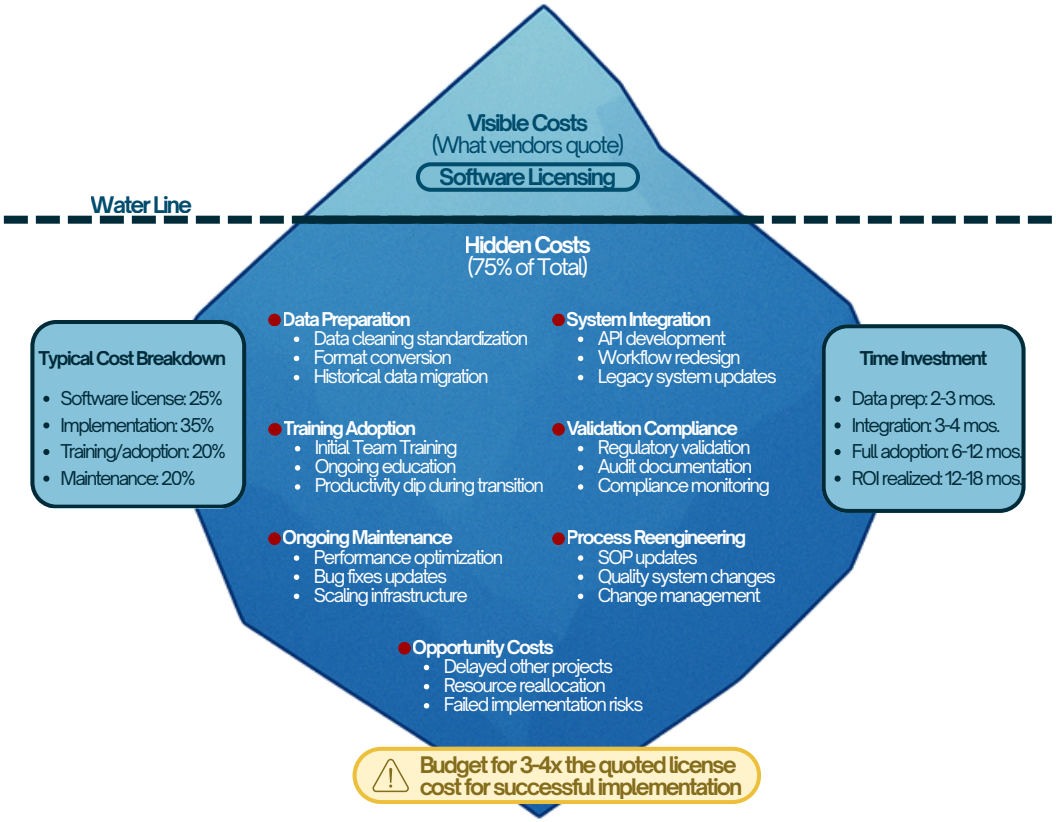
- **Procurement teams lack AI-specific evaluation criteria**

AI is an emerging industry with specialist level knowledge requirements. Most teams lack knowledge.

- **The "black box" problem in vendor presentations**

The quality of the solution behind the software is not easy to evaluate from vendor led demos.

1.3 The Hidden Costs of Poor AI Implementation



Compliance Gaps from Over-Reliance on Unvalidated AI

The most dangerous hidden cost comes from compliance failures:

Validation shortcuts create regulatory time bombs. Under pressure to show ROI quickly, teams may skip proper validation of AI outputs.

False confidence in AI recommendations without understanding their limitations. Teams may trust AI conclusions without maintaining appropriate oversight, leading to submission errors that damage regulatory relationships.

Audit trail gaps emerge when AI systems don't properly document decision rationale. Regulatory authorities increasingly ask not just what AI concluded but how it reached those conclusions.

Early Warning Signs of Impending Hidden Costs

Watch for these red flags during vendor evaluation:

- Vendors who minimize discussion of data preparation requirements
- Unrealistic implementation timelines (less than 3 months for enterprise deployment)
- No clear validation pathway or regulatory guidance
- Lack of change management support or training resources
- Vague pricing structures that could hide usage-based charges
- No discussion of integration requirements with your existing systems

1.4 What's Different About AI for Regulatory Affairs

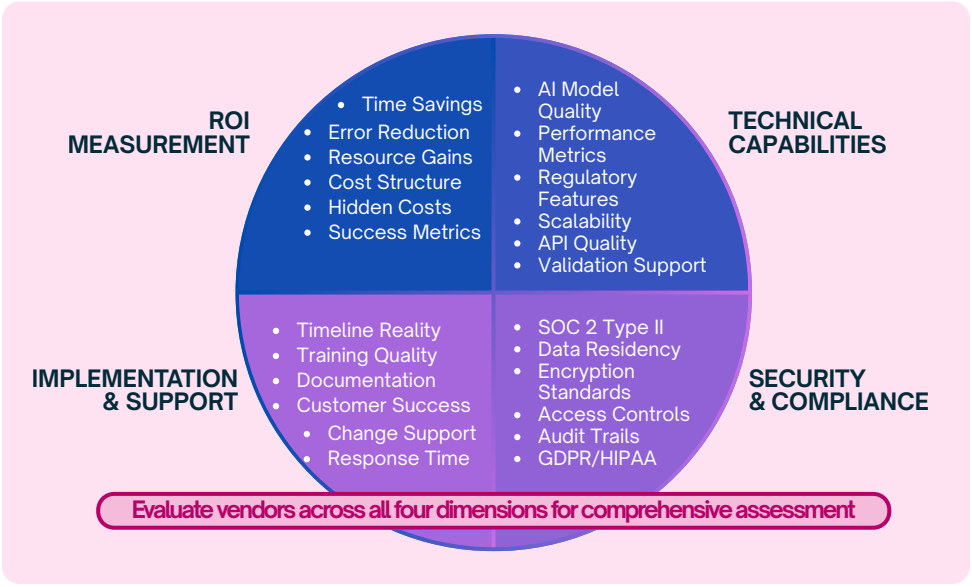
AI for regulatory affairs, clinical, quality, and medical writing are some of the most contentious applications in the industry. This is due to:

- Higher stakes than typical business software
- Audit trail and validation requirements
- Need for explainable AI in regulatory contexts
- Integration with existing quality management systems

Understanding Chain of Thought Reasoning

Where a black box solution might be acceptable for a consumer product customer support chatbot, all solutions related to the placing of medical devices on the market must have detailed reasoning and audit trails. Equally as important to the suggestions made by AI, are the reasoning chains of how the model arrived at that result. We will discuss this in depth later on.

SECTION 2: CRITICAL EVALUATION FRAMEWORK



2.1 Technical Capabilities Assessment

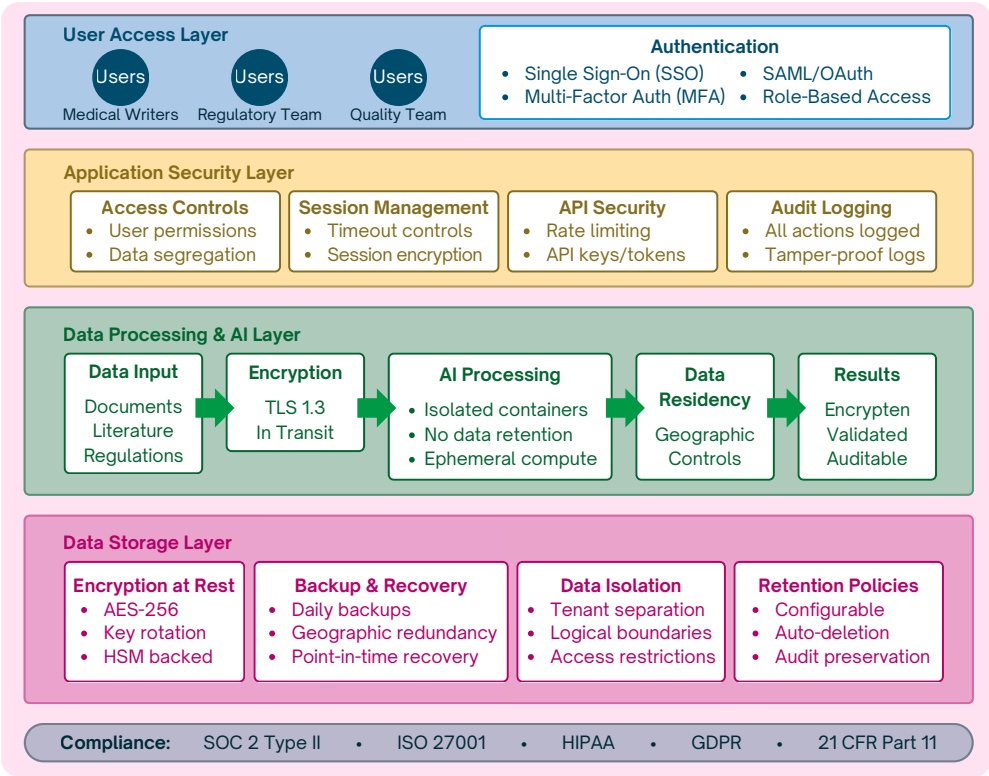
Core AI Functionality
- Natural language processing accuracy for medical literature
- Data extraction and synthesis capabilities
- Pattern recognition in regulatory documents
- Performance benchmarks and validation studies

Reliability and Consistency
- Error rates in different document types
- Performance across various medical device categories
- Handling of edge cases and novel scenarios
- Model versioning and update management

Interpretability and Explainability
- Ability to show reasoning behind AI recommendations
- Audit trail capabilities for regulatory compliance
- Human oversight and intervention mechanisms
- Transparency in AI decision-making processes

Technical capabilities form the foundation of any AI solution evaluation. Without robust performance in your specific use cases, even the most secure and well-integrated system will fail to deliver value. This dimension focuses on three core areas: functionality, reliability, and interpretability.

2.2 Data Security and Compliance Evaluation



Regulatory Compliance

FDA guidance alignment for AI/ML systems continues to evolve. Ensure vendors understand and follow Good Machine Learning Practices (GMLP), provide predetermined change control plans, and support your validation requirements.

Validation documentation and support determines implementation feasibility. Vendors should provide IQ/OQ/PQ protocols, support your validation activities, and maintain validation status through updates.

The EU AI Act

Determining the risk rating of vendor AI system now carries legal implications. AI systems used for medical device regulatory decisions likely fall into high-risk categories, requiring extensive documentation, human oversight, and transparency measures.

You can't trust your vendor's assessment of your use case. Conduct your own risk assessment based on how you'll use the AI, not generic vendor classifications. High-risk categorization triggers requirements for risk management systems, data governance, human oversight, and accuracy monitoring that vendors may not fully support.

Workflow Adaptation

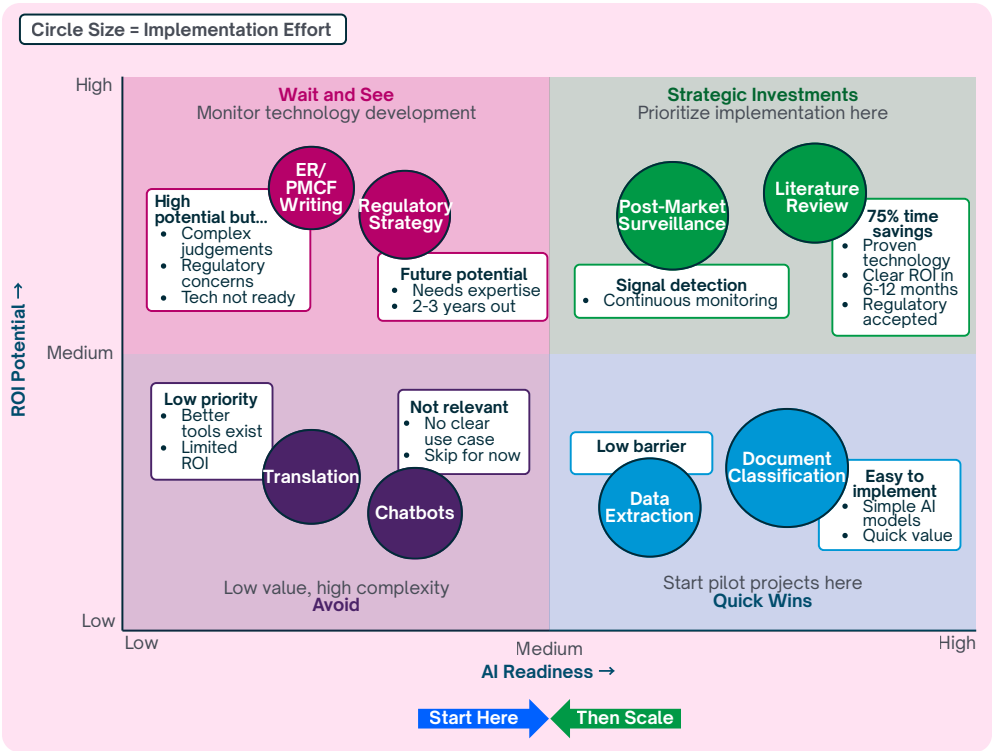
Learning curve and training requirements directly impact adoption timelines. While AI should simplify work, initial training is inevitable. Evaluate whether training is role-specific, how long competency takes to achieve, and what ongoing support is available.

Change management support and resources separate successful implementations from failures. Vendors should provide implementation playbooks, template communications, and change management expertise. This soft support often matters more than technical features for driving adoption.

Customization options for company-specific processes acknowledge that no two companies work identically. Can search strategies be modified for your therapeutic areas? Can output templates match your submission formats? Limited customization forces process changes that teams resist.

Scalability across different team sizes and functions ensures long-term viability. The system that works for a 5-person medical writing team must scale to support 50 users across multiple departments. Consider concurrent user limits, performance degradation with scale, and licensing models that accommodate growth.

2.3 ROI Measurement Methodology



Demonstrating return on investment requires moving beyond vendor promises to measure actual impact. This dimension provides frameworks for building a compelling business case and tracking value delivery over time.

Quantifiable Benefits

Time savings per regulatory process provides the clearest ROI metric. Measure specific tasks: literature review time reduction, adverse event processing acceleration, or submission compilation improvements. Convert time savings to dollar values using fully loaded employee costs.

Error reduction and quality improvements translate to risk mitigation. Track metrics like citations missed in manual review but caught by AI, data extraction accuracy improvements, and reduction in submission deficiencies. Each prevented error represents avoided rework and potential compliance issues.

Resource reallocation opportunities unlock strategic value. If AI handles routine literature screening, medical writers can focus on complex analysis and strategic recommendations. Quantify how many additional submissions your team can support without adding headcount.

Faster time-to-market for submissions drives revenue impact. If AI-assisted literature reviews shave weeks off submission preparation, calculate the value of earlier market entry. For high-revenue devices, even small acceleration generates significant returns.

Remember: These four evaluation dimensions work together like quadrants on a wheel. Technical excellence means nothing without security compliance. Perfect integration fails without demonstrable ROI. Strong security with poor user experience leads to rejection. Evaluate holistically across all four dimensions to ensure successful AI implementation.

This evaluation framework provides the structure for systematic AI solution assessment. The following sections will explore specific vendor questions and implementation strategies to apply this framework effectively.

SECTION 3: REGULATORY-SPECIFIC AI APPLICATIONS

Within Medtech, Pharma and Life Science organizations there are far too many useful applications of AI to be discussed in full. However, narrowing the focus to compliance-related Medtech solutions can be a great starting point for this whitepaper and illustrative of the overall approach.

3.1 Systematic Literature Review Automation - The Biggest ROI Opportunity

The collection, review and historical management of clinical literature is the biggest opportunity for AI use cases in your Medtech business units. The procedures relating to AI are often specialized, repeatable and driven by manual (human) labor.

Current Manual Process Pain Points

The traditional literature review process represents one of the most time-consuming and error-prone activities in regulatory affairs:

90+ hours per literature review consumed by manual searching, screening, and data extraction. For complex devices requiring quarterly updates, this translates to over 160 hours annually per device—often spread across multiple team members.

High error rates in manual screening stem from reviewer fatigue and the sheer volume of literature. Studies show human reviewers miss 5-10% of relevant studies and inconsistently apply inclusion criteria, especially when processing hundreds of abstracts.

Inconsistent reviewer decisions create quality concerns and rework. When two reviewers disagree on 20-30% of articles (common in practice), resolution requires additional senior reviewer time.

Version control and audit trail challenges plague spreadsheet-based tracking systems. Teams struggle to maintain clear documentation of search strategies, screening decisions, and extraction rationale—all required for regulatory compliance.

AI Solution Capabilities

Modern AI transforms literature review from a manual marathon into a streamlined, auditable process:

Automated database searching across multiple sources eliminates the need to manually construct searches for each database.

AI-powered relevance screening and ranking applies natural language processing to understand study abstracts beyond simple keyword matching.

Duplicate detection and removal uses sophisticated algorithms to identify the same study published in different venues or with slight variations.

PRISMA-compliant report generation automates the creation of flow diagrams, evidence tables, and bibliographies.

Implementation Considerations

Successful implementation requires thoughtful planning beyond just purchasing software:

Validation requirements for systematic reviews demand demonstrating that AI-assisted reviews are as thorough as manual processes. This typically involves parallel processing during initial implementation—running both manual and AI-assisted reviews to verify comparable results. Document your validation approach, acceptance criteria, and ongoing performance monitoring.

Integration with existing review protocols should enhance, not replace, your established procedures. Map current workflows to identify where AI adds value (searching, screening, extraction) versus where human expertise remains essential (quality assessment, synthesis, conclusions). Modify SOPs to incorporate AI tools while maintaining compliance.

Training AI on company-specific inclusion criteria improves performance over time. While modern LLMs work well out-of-the-box, feeding back your screening decisions helps the system understand your unique requirements. For example, your definition of "long-term follow-up" might be 5 years for joint replacements but 1 year for wound dressings.

Maintaining human oversight for critical decisions ensures regulatory compliance and clinical validity. Establish clear boundaries: AI suggests relevance rankings, but humans make final inclusion decisions. AI extracts data points, but humans verify accuracy and clinical significance. This graduated approach maintains quality while maximizing efficiency.

Case Study Example

A mid-sized orthopedic device manufacturer reduced literature review time by 75% while maintaining 100% audit compliance with their notified body. Their systematic review process for a Class III knee implant previously required 6 weeks (240 hours) across two medical writers. After implementing AI-assisted literature review:

- Initial searching across 5 databases: Reduced from 8 hours to 15 minutes
- Duplicate removal: Automated process eliminated 6 hours of manual Excel work
- Abstract screening (500 articles): Reduced from 20 hours to 5 hours using AI ranking
- Full-text review and data extraction: Reduced from 80 hours to 20 hours with AI pre-population
- Report generation: Reduced from 16 hours to 1 hour using automated templates

Total time: 60 hours (1.5 weeks) with improved consistency and complete audit trail. The notified body praised their documentation during the next audit.

3.2 Post-Market Surveillance Enhancement

AI Applications

- Automated safety signal detection
- Adverse event classification and trending

- Literature monitoring for device-related publications
- Regulatory database monitoring and alerts

Technical Requirements
- Real-time data processing capabilities
- Multi-language support for global surveillance
- Integration with complaint management systems
- Regulatory reporting automation

Post-market surveillance represents the second major opportunity for AI implementation, addressing the growing regulatory emphasis on continuous safety monitoring throughout device lifecycle.

AI Applications

Automated safety signal detection transforms reactive surveillance into proactive risk management. AI continuously monitors multiple data streams—literature, regulatory databases, social media, and internal complaints—to identify patterns humans might miss. For instance, three case reports mentioning "unusual wear patterns" combined with a subtle increase in revision rates could trigger early investigation of a potential issue.

Adverse event classification and trending standardizes inconsistent reporting across sources. The AI recognizes that "device malfunction," "implant failure," and "prosthesis breakage" may represent the same issue, enabling accurate trending. Natural language processing extracts severity, timing, and contributing factors from narrative reports, converting unstructured complaints into analyzable data.

Literature monitoring for device-related publications ensures no critical safety study goes unnoticed. Unlike periodic manual searches, AI provides continuous monitoring with intelligent filtering. The system distinguishes between your specific device model and similar products, preventing alert fatigue while maintaining comprehensive coverage.

Regulatory database monitoring and alerts tracks global safety communications relevant to your devices. When FDA issues a warning letter to a competitor or EU authorities publish new guidance, AI alerts identify potential impacts on your products. This early warning system enables proactive response rather than reactive scrambling.

Technical Requirements

Effective post-market surveillance AI demands specific technical capabilities:
- Real-time data processing capabilities
- Multi-language support for global surveillance
- Integration with existing complaint management systems

3.3 Clinical Evaluation Report (CER) Support

Clinical Evaluation Reports represent the most complex regulatory documents, requiring synthesis of diverse evidence into coherent arguments for device safety and performance. AI can significantly support this process while respecting the need for clinical judgment.

What AI Can Do Well

Literature search and initial screening follows similar principles to systematic reviews but with broader scope. CER literature searches must capture not just clinical studies but also registry data, standards, guidance documents, and competitive device information. AI excels at this comprehensive searching and initial relevance filtering.

Data extraction from clinical studies accelerates the most tedious aspect of CER preparation. AI can accurately extract study designs, patient populations, outcome measures, adverse events, and follow-up durations. For a CER covering 50+ studies, this automation saves days of manual work while ensuring consistent data capture.

Gap analysis identification highlights where additional evidence may be needed. By comparing extracted evidence against regulatory requirements and competitor claims, AI identifies gaps in your clinical evidence. For example, if MDR requires 10-year follow-up data but your longest study is 7 years, AI flags this gap for strategic planning.

What AI Cannot (Yet) Do Reliably

Clinical judgment and interpretation remains firmly in human domain. While AI can extract that Study A showed 95% success rate and Study B showed 85%, only clinical experts can determine whether this difference reflects study populations, surgical technique, or genuine device performance differences.

Risk-benefit analysis synthesis requires nuanced understanding of clinical context. AI cannot weigh whether a 2% infection rate is acceptable given the severity of the treated condition or whether alternative treatments carry greater risks. These judgments require clinical expertise and understanding of patient perspectives.

Regulatory strategy recommendations demand experience with authority expectations and precedent. Choosing whether to claim superiority, equivalence, or non-inferiority based on available evidence requires regulatory expertise AI cannot replicate.

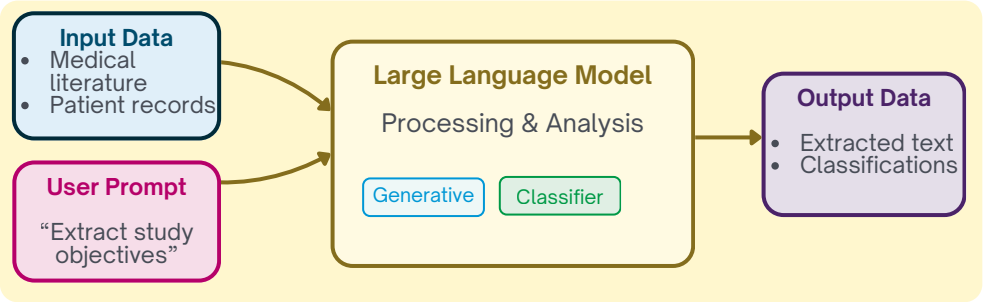
Final clinical conclusions and decisions carry professional and legal responsibility that AI cannot assume. Medical writers and clinical evaluators must personally stand behind conclusions about device safety and performance.

SECTION 4: VENDOR EVALUATION - THE RIGHT QUESTIONS TO ASK

4.1 Technical Capabilities Deep Dive

AI Model and Training

The AI models used in your solution are the "brain" of the functionality and need to be handled and maintained properly. Think of them like input → output machines. Your tasks and business data go in, and (hopefully) your answers and relevant outputs come out.



Models are updated through either an initial training process, or an ongoing incremental fine-tuning process. Below are some questions that should be easily answered by any AI vendor.

> What AI models do you use and how were they trained?

This question will tell you a lot about the type of solution you are evaluating. Commercially trained models large language models (like from OpenAI, Anthropic, Google and Meta) are trained to be comprehensive solution providers. Private (and smaller) models in contrast can be highly specific in their focus and application.

> Show me validation studies comparing your AI to manual processes

Validation studies compare AI generated results to that of a human. Vendors that have built thoughtfully should have some comparison data to share. If they do not (or cannot share), a pilot period may become more essential to conduct your own version of this type of study.

> How do you handle updates to underlying AI models?

Like with any enterprise piece of technology or software an update cycle can introduce significant risk, but also incremental improvements to a solution over time. When using commercial models, the update cycle may not be completely transparent, and thus a plan should be put in place for managing updates and distributing them to the models your solution would be using specifically.

Performance and Reliability

How do you measure and report AI performance over time?

While we feel that many metrics related to the performance o AI solutions are not necessary, some simple and functional reporting should be a pre-requisite to deploying any new solution.

4.2 Integration and Implementation

What APIs are available for custom integrations?

Understanding API capabilities helps you assess integration complexity and future customization options. The vendor should provide comprehensive API documentation and support for custom integration development.

What's required from our IT team during implementation?

Some AI solutions require significant IT infrastructure changes while others work as hosted services with minimal IT involvement. Understand the technical requirements upfront to ensure your IT team can support the implementation.

How do you handle data migration from our current systems?

Moving existing document libraries and historical data into new AI systems can be complex and time-consuming. The vendor should provide migration tools, data validation processes, and support for handling legacy data formats.

Change Management

What training is required for our team?

Training requirements vary significantly between AI solutions. Some require extensive technical training while others are designed for immediate use by non-technical professionals. Understand the time commitment and ongoing training needs for your team.

How long is the typical implementation timeline?

Implementation timelines depend on organizational complexity, data migration requirements, and integration needs. Get realistic timelines based on organizations similar to yours, not best-case scenarios.

What ongoing support do you provide post-implementation?

Post-implementation support should include technical support, user training refreshers, performance optimization, and guidance on new feature adoption. Understand support response times and escalation procedures.

How do you measure user adoption success?

The vendor should have frameworks for measuring and improving user adoption rates. Look for vendors that provide adoption analytics, identify usage barriers, and offer strategies for increasing engagement.

4.3 Vendor Stability and Longevity

What's your strategy for incorporating new AI advances?

AI technology evolves rapidly. The vendor should have clear strategies for incorporating new AI capabilities, staying current with research developments, and continuously improving their platform.

What's your long-term vision for AI in regulatory affairs?

The vendor's vision should demonstrate deep understanding of regulatory challenges and realistic perspectives on AI capabilities. Be wary of vendors promising unrealistic automation levels or dismissing the importance of human oversight in regulatory processes.

SECTION 5: IMPLEMENTATION STRATEGY

6.1 Pilot Program Design

The Reality of AI Pilots

Most AI pilot programs fail not because the technology doesn't work, but because organizations choose the wrong processes to test or set unrealistic expectations for what constitutes success. The key is selecting pilots that can demonstrate clear value while minimizing organizational disruption.

Literature review processes are ideal first pilots because they're high-volume, clearly defined, and have measurable outcomes. Avoid starting with complex clinical evaluation reports or novel regulatory pathways where success criteria are ambiguous.

Essential Pilot Design Elements

Parallel Processing is Non-Negotiable
Run your existing manual process alongside the AI-assisted workflow for the entire pilot duration. This serves two purposes: it provides validation data for performance comparison, and it ensures business continuity if the AI solution fails to meet expectations.

Define "Success" Before You Start
Establish specific, measurable criteria such as "75% reduction in initial screening time" or "95% agreement between AI and human reviewer decisions." Vague goals like "improved efficiency" will lead to inconclusive pilots and difficult vendor negotiations.

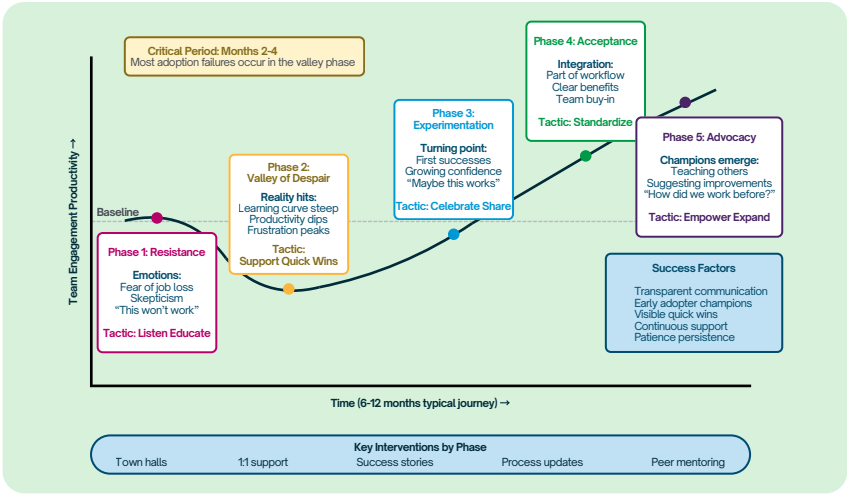
Plan for Data Quality Issues
Most organizations discover their existing data isn't as AI-ready as they assumed. Budget additional time for data cleaning, format standardization, and quality validation. This preparation work often takes longer than the actual AI implementation.

Checkpoint Reviews Every 2-4 Weeks
Regular reviews allow for course corrections and help maintain stakeholder engagement. Use these sessions to address user concerns, refine processes, and document lessons learned for broader implementation.

Realistic Timeline Expectations
A 3-6 month pilot duration provides adequate time for meaningful evaluation without dragging out the decision-making process:

- **Month 1:** Setup, configuration, and initial training
- **Month 2-3**: Active pilot with parallel processing
- **Month 4**: Data analysis and decision-making
- **Month 5-6**: Optimization and scaling preparation (if proceeding)

6.2 Change Management for AI Adoption



6.3 Measuring Success Metrics

Establish Your Baseline Before Implementation

Most organizations skip this critical step and later struggle to demonstrate AI value. Document current process performance: how long literature reviews actually take (not how long they're supposed to take), error rates in manual screening, cost per regulatory deliverable, and team satisfaction levels.

User satisfaction and stress levels are often overlooked but critical metrics. If AI reduces processing time by 50% but increases user frustration, adoption will fail.

Ongoing Performance Tracking

Weekly performance reports should focus on leading indicators: user adoption rates, system uptime, and immediate quality metrics. Monthly reports can address broader trends in efficiency gains, cost savings, and user satisfaction changes.

Quality metrics require careful definition in regulatory contexts. Track not just speed improvements but accuracy maintenance. If AI-assisted reviews are 60% faster but miss 5% more relevant studies, the net value may be negative.

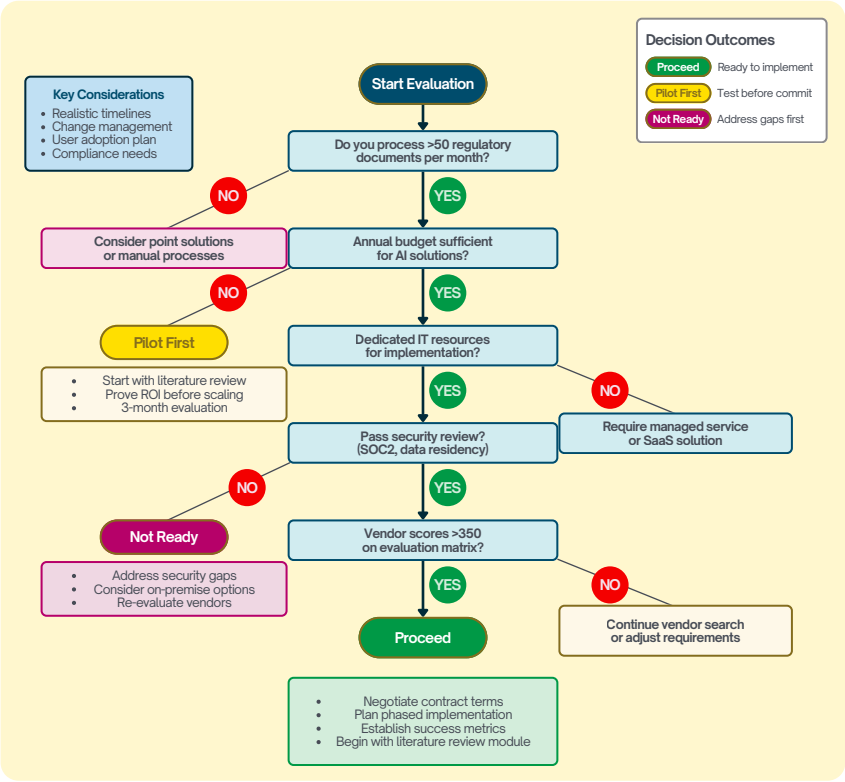
ROI calculations should include hidden costs such as additional training time, system maintenance overhead, and process modification expenses. Many organizations focus only on time savings while ignoring implementation and maintenance costs.

Long-term Value Assessment

Annual business impact reviews should evaluate strategic positioning, not just operational metrics. Has AI adoption enabled your team to take on more complex projects? Has it improved your competitive position in submission timelines?

Team capability development often provides unexpected value. Teams using AI tools frequently develop better understanding of regulatory patterns and more systematic approaches to problem-solving.

SECTION 6: MAKING THE DECISION - A PRACTICAL FRAMEWORK



SECTION 7: COMMON CONCERNS AND MISUNDERSTANDINGS OF MANUFACTURERS

7.1 Major Concerns with AI

Will Our Proprietary Data Be Safe?

There are disputes in the news daily around major providers improperly training their AI models on public data. While these may be valid, the concern **does not translate to enterprise AI.** Any enterprise provider will be able to guarantee that your requests, data, or results are not used in the training or improvement of AI models. If they can't, do not work with them.

What if the AI Models Hallucinate?

In our industry the accuracy of information is critical to a valuable AI system. This is where selection of your AI features matters, and creating strong human feedback loops in all processes powered by AI. In addition, any worthy vendor will have configurable and extensive "guardrails" for their solutions to minimize or eliminate misinformation.

CONCLUSION

You should now have awareness into the main components needed to evaluate an AI system (and vendor). We hope that you enjoyed the presentation format, and will continue to reference this white paper and evaluation rubrics provided in the Appendices.

The team at CiteMed that powers Evidence Cloud cares about AI governance and manufacturers receiving relevant information on AI. If you have additional questions or would like to connect I can be reached:

Email: edrower@citemed.com
Linkedin: https://linkedin.com/in/ethandrower
CiteMed.com

APPENDICES

Visit citemed.com/ai to download all templates and appendices

AI Solution Evaluation Checklist for Medical Device Companies

Scoring:
- ✓ = Yes (2 pts)
- ~ = Partial (1 pt)
- ✗ = No (0 pts)

Min Score: 35/50
Critical items (*) must be ✓

Technical Capabilities		
Criteria	**✓/~/✗**	**Notes**
*Uses major AI providers (OpenAI, Anthropic, Google)		
*Chain of thought reasoning visible		
*>95% accuracy on validation studies		
Comprehensive audit trail		
FDA/EU AI Act compliance ready		

Security & Compliance		
Criteria	**✓/~/✗**	**Notes**
*SOC 2 Type II certified		
*Data won't be used for model training		
HIPAA/GDPR compliant		
Encryption (AES-256/TLS 1.3)		
Validation documentation (IQ/OQ/PQ)		

AI Performance & Optimization		
Criteria	**✓/~/✗**	**Notes**
*Fine-tuning on your company's data/criteria		
*Continuous prompt engineering improvements		
Customizable for your therapeutic areas		
Regular model updates without disruption		
*Fine-tuning on your company's data/criteria		

Measurable ROI & Value		
Criteria	**✓/~/✗**	**Notes**
*>75% time savings for literature review		
*Clear metrics dashboard/reporting		
Reduces error rates vs manual process		
Enables resource reallocation		
Faster submission timelines demonstrated		

Red Flags (Any = Disqualify)	Decision Framework	
	Total Score	**Action**
• Implementation <3 months promised • Vague/usage-based pricing hidden • No validation pathway • Can't guarantee data privacy • No human oversight options	40 - 50	Proceed w/ implementation
	35 - 39	Run 3-month pilot first
	30 - 34	Address gaps, re-evaluate
	<30	Do not proceed

[cite] med