

Exploring Multimodal AI for Early Cancer Detection

Integrating imaging, genomics, and
clinical records for precision oncology

White paper



Executive Summary

Cancer remains one of the leading causes of mortality worldwide, largely because it is frequently detected at advanced stages rather than at onset (World Health Organization, 2023). Although healthcare systems generate vast amounts of patient data, diagnostic workflows still operate in disconnected silos. Radiology images, pathology findings, genomic tests, and patient history are interpreted separately, leaving clinicians to combine them.

This fragmented process delays diagnosis and leads to inconsistent interpretations among specialists. Artificial intelligence has already demonstrated strong performance in detecting abnormalities in imaging and pathology data; however, most deployed systems analyze only one type of medical data at a time (Esteva et al., 2017; Litjens et al., 2017). Because cancer is inherently multifactorial, involving structural changes, molecular mutations, and clinical context simultaneously, single-input AI systems provide incomplete understanding.

This white paper proposes a multimodal AI diagnostic framework that integrates:

- Radiology imaging
- Histopathology slides
- Genomic mutations
- Electronic health records (EHR)

The system produces early detection probability, tumor classification, and risk assessment while providing explainable reasoning to support clinical decision-making.

The objective is not to replace physicians but to augment diagnostic confidence and detect cancer earlier.

Industry context

Table 1: Healthcare is shifting from reactive treatment to predictive care. Early detection significantly improves survival rates across major cancers (Siegel et al., 2024).

Cancer type	Survival if detected early	Survival if detected late
Breast cancer	>90%	<30%
Lung cancer	~65%	<10%
Colorectal cancer	~80%	~15%

Modern hospitals collect data from multiple sources:

- Imaging departments generate scans daily
- Laboratories provide molecular reports
- EHR systems store longitudinal patient history

However, these datasets are rarely analyzed collectively. Each department functions independently, and integration depends on manual clinical review. Precision oncology requires combining multiple evidence streams, which is difficult without computational support (Topol, 2019).

Problem Statement

Fragmented Diagnosis Workflow

This workflow leads to four major limitations.

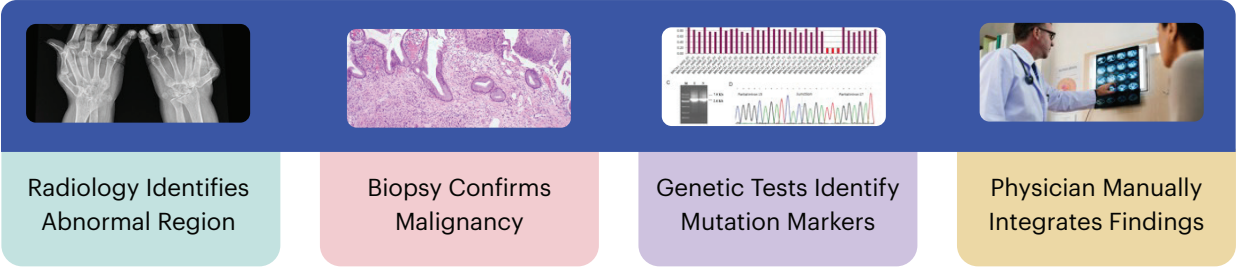


Figure 1: Current diagnostic practice involves sequential evaluation

1. Information loss

When clinicians review data sources separately, they miss relationships between them. Clinicians focus on their specialty-specific evidence, which may hide early disease indicators.

For example, a small pulmonary nodule might appear clinically insignificant in imaging alone. However, when combined with smoking history and mutation markers, it becomes high-risk. Without computational integration, clinicians must rely on individual experience to recognize these patterns.

Studies show that combining imaging and genomic information significantly improves diagnostic accuracy compared to using either alone (Huang et al., 2020).

In practice, early cancer signals are often weak individually but strong collectively. Failing to combine them misses opportunities for early detection.

2. Subjective interpretation

Different clinicians may interpret the same report differently because of their training background and experience level. Variability in radiology and pathology interpretation is well documented in oncology diagnostics (Elmore et al., 2015).

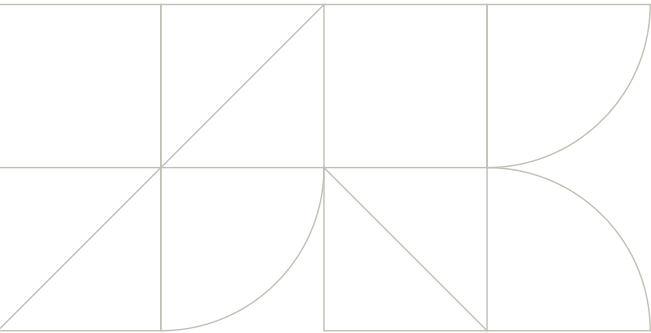
For example:

Two radiologists may classify a borderline lesion differently; one recommending biopsy, another recommending observation.

This variability leads to:

- Inconsistent treatment plans
- Patient anxiety
- Repeated testing

AI systems trained on large datasets can provide standardized assessment, reducing inter-observer variability (Litjens et al., 2017).



3. Delayed Decisions

Cancer diagnosis requires multiple reports, often generated days apart. Treatment decisions are postponed until all evidence is available. During this waiting period, the disease may progress.

Typical delays include:

- Imaging scheduling
- Pathology processing
- Genetic testing turnaround

Studies indicate diagnostic delay is associated with worse outcomes in aggressive cancers (Neal et al., 2015). Automated integrated analysis can produce preliminary risk assessment immediately after data availability, accelerating intervention.

4. Single-modality AI limitations

Table 2: Many current medical AI systems specialize in only one modality.

AI Type	Limitation
Radiology AI	Detects structure but not mutation
Genomic AI	Detects mutation but not location
EHR AI	Predicts risk but not visible tumors

Because cancer manifests across biological levels (anatomical, cellular, and molecular), analyzing one modality produces incomplete assessment (Acosta et al., 2022).

As a result, AI may generate confident but incomplete predictions. Integrating modalities lets models learn clinically meaningful relationships rather than isolated patterns.

Proposed Solution

Multimodal clinical AI framework

We propose an integrated AI system that simultaneously analyzes heterogeneous medical data and constructs a unified patient representation.

Rather than multiple disconnected predictions, the system produces a holistic clinical assessment.

System Overview

Input data sources

1. Medical imaging → CT, MRI, mammography
2. Histopathology → cellular morphology
3. Genomic data → mutation markers
4. Clinical records → demographics and risk factors

Each source contributes a different perspective of the disease.

Architecture design

Step 1 - Feature extraction

Table 3: Each modality is processed by specialized models.

Data type	Purpose
Imaging	Structural abnormality detection
Pathology	Malignant cell identification
Genomics	Mutation risk evaluation
Clinical text	Context extraction

Step 2 – Cross-modal fusion

Features are merged into a shared representation space.
The model learns relationships across modalities rather than treating them independently.

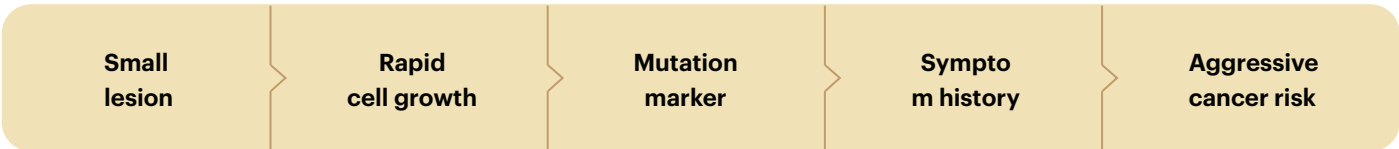


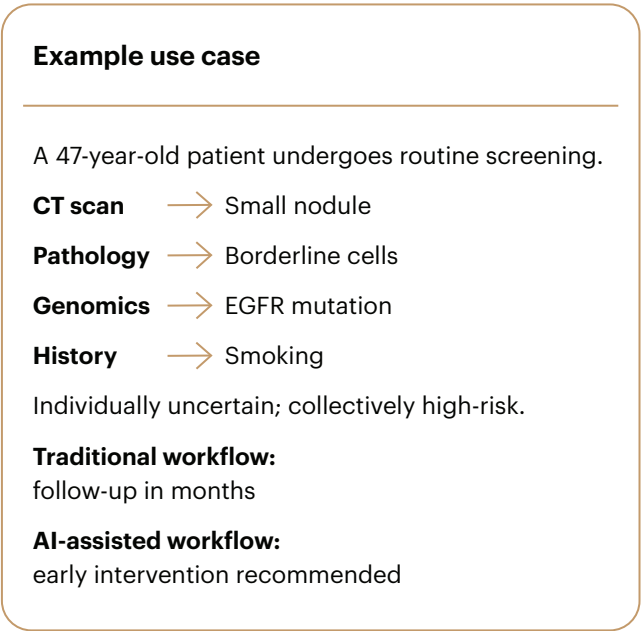
Figure 2: Example learned pattern

Multimodal learning has shown superior performance compared to single-modality analysis in recent clinical AI research (Zhang et al., 2021).

Step 3 - Prediction layer

Output	Clinical use
Cancer probability	Structural abnormality detection
Tumor type	Malignant cell identification
Progression	Mutation risk evaluation
Risk explanation	Context extraction

Figure 3: The system outputs



Clinical workflow integration

- 1. Hospital systems upload patient data
- 2. AI generates structured report
- 3. Doctor validates recommendation
- 4. Decision recorded with explanation

The physician remains the final authority.

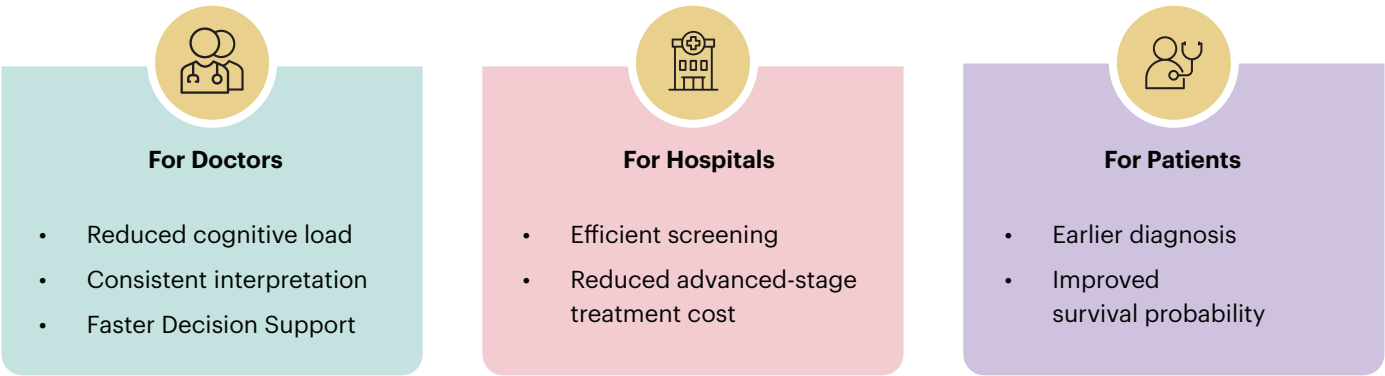
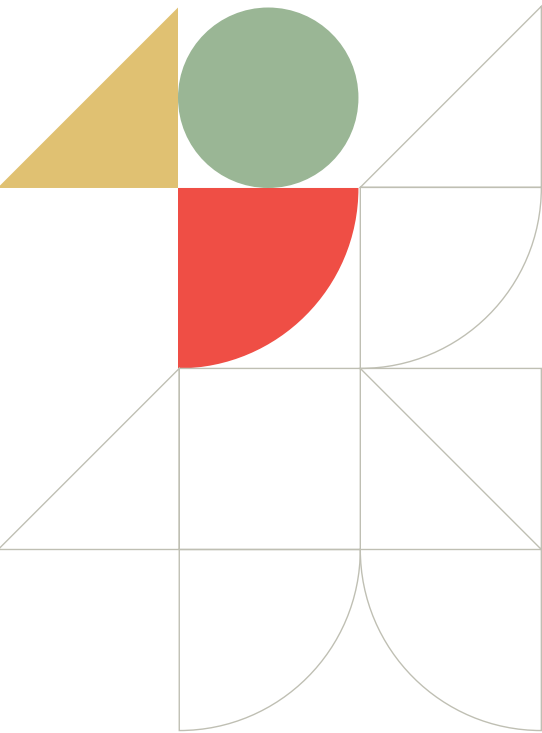


Figure 4: Benefits of early cancer detection

Evaluation Strategy

Table 4: Performance metrics

Metric	Purpose
Sensitivity	Detect cancer
Specificity	Avoid false alarms
AUC	Overall accuracy
Time-to-diagnosis	Workflow improvement



Clinical metric

Agreement rate between AI and oncologists

Ethical and Safety Considerations

- AI functions as decision support
- Physician approval required
- Anonymized patient data
- Bias monitoring across demographics

Responsible AI deployment is essential for clinical acceptance (Topol, 2019).

Future scope

Potential extensions include:

- Treatment selection
- Therapy response prediction
- Recurrence monitoring

This leads toward continuous patient monitoring systems.

Conclusion

Cancer diagnosis requires synthesizing multiple types of medical evidence. Current workflows rely on manual integration across disconnected systems, causing delays and inconsistencies.

A multimodal AI framework can combine imaging, pathology, genomics, and clinical history into a unified understanding of the patient. Such a system improves early detection, enhances clinician confidence, and supports precision oncology without replacing medical expertise.

This approach represents a practical step toward data-driven cancer care, where diagnosis is informed by complete patient context rather than isolated observations.



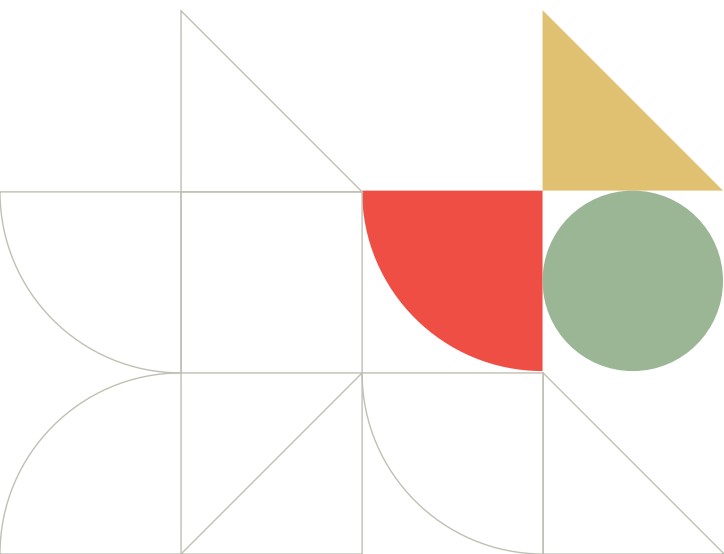
References

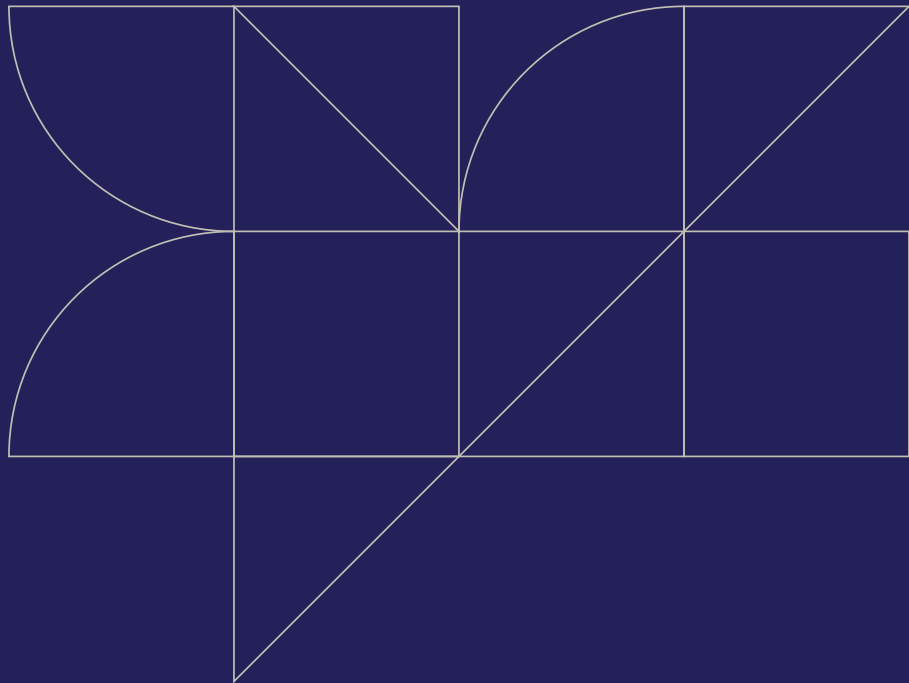
- Acosta, J. N., Falcone, G. J., Rajpurkar, P., & Topol, E. J. (2022). Multimodal biomedical AI. *Nature Medicine*, 28(9), 1773–1784. <https://doi.org/10.1038/s41591-022-01981-2>
- Elmore, J. G., Longton, G. M., Carney, P. A., et al. (2015). Diagnostic concordance among pathologists interpreting breast biopsy specimens. *JAMA*, 313(11), 1122–1132. <https://doi.org/10.1001/jama.2015.1405>
- Esteva, A., Kuprel, B., Novoa, R. A., et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115–118. <https://doi.org/10.1038/nature21056>
- Huang, S., Yang, J., Fong, S., & Zhao, Q. (2020). Artificial intelligence in cancer diagnosis and prognosis: Opportunities and challenges. *Cancer Letters*, 471, 61–71. <https://doi.org/10.1016/j.canlet.2019.08.019>
- Litjens, G., Kooi, T., Bejnordi, B. E., et al. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88. <https://doi.org/10.1016/j.media.2017.07.005>
- Neal, R. D., Tharmanathan, P., France, B., et al. (2015). Is increased time to diagnosis and treatment in symptomatic cancer associated with poorer outcomes? *British Journal of Cancer*, 112(S1), S92–S107. <https://doi.org/10.1038/bjc.2015.48>
- Siegel, R. L., Miller, K. D., Wagle, N. S., & Jemal, A. (2024). Cancer statistics, 2024. *CA: A Cancer Journal for Clinicians*, 74(1), 17–48. <https://doi.org/10.3322/caac.21820>
- Risheng Wang, Tao Lei, Ruixia Cui, Bingtao Zhang, Hongying Meng, Asoke K. Nandi (2021). Medical Image Segmentation Using Deep Learning: A Survey, arXiv:2009.13120
- Electrical Engineering and Systems Science <https://doi.org/10.48550/arXiv.2009.13120>
- Topol, E. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44–56. <https://doi.org/10.1038/s41591-018-0300-7>
- World Health Organization. (2023). Cancer. <https://www.who.int/news-room/fact-sheets/detail/cancer>

Author

N Mouna

Zensar Applied Research, Zensar Technologies





zensar
An  **RPG** Company

At Zensar, we're 'experience-led everything.' We are committed to conceptualizing, designing, engineering, marketing, and managing digital solutions and experiences for over 145+ leading enterprises. Using our 3Es of experience, engineering, and engagement, we harness the power of technology, creativity, and insight to deliver impact.

Part of the \$4.8 billion RPG Group, we are headquartered in Pune, India. Our 10,000+ employees work across 30+ locations worldwide, including Milpitas, Seattle, Princeton, Cape Town, London, Zurich, Singapore, and Mexico City.

For more information, please contact: info@zensar.com | www.zensar.com